# Overview of VH WG1 experimental/theory status

A. Calandri (ETH Zürich), H. Arnold (NIKHEF)
C. Williams (Buffalo), G. Ferrera (University of Milan) T. Calvet (CPPM), C. Pandini (UniGe), P. Windischhofer (University of Oxford), S. Jigging (University of Edinburgh)

WG1 - VH group meeting - November 8, 2021

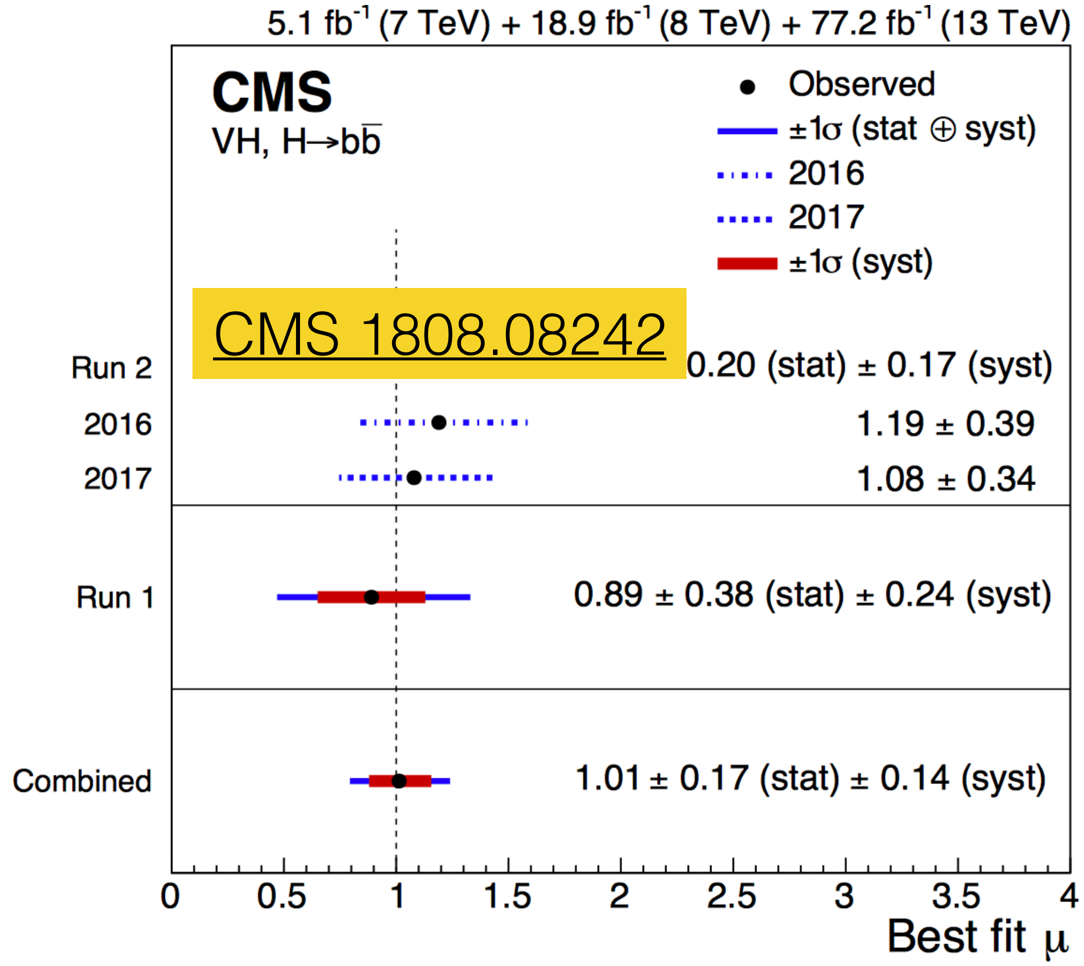
# Outline of the talk

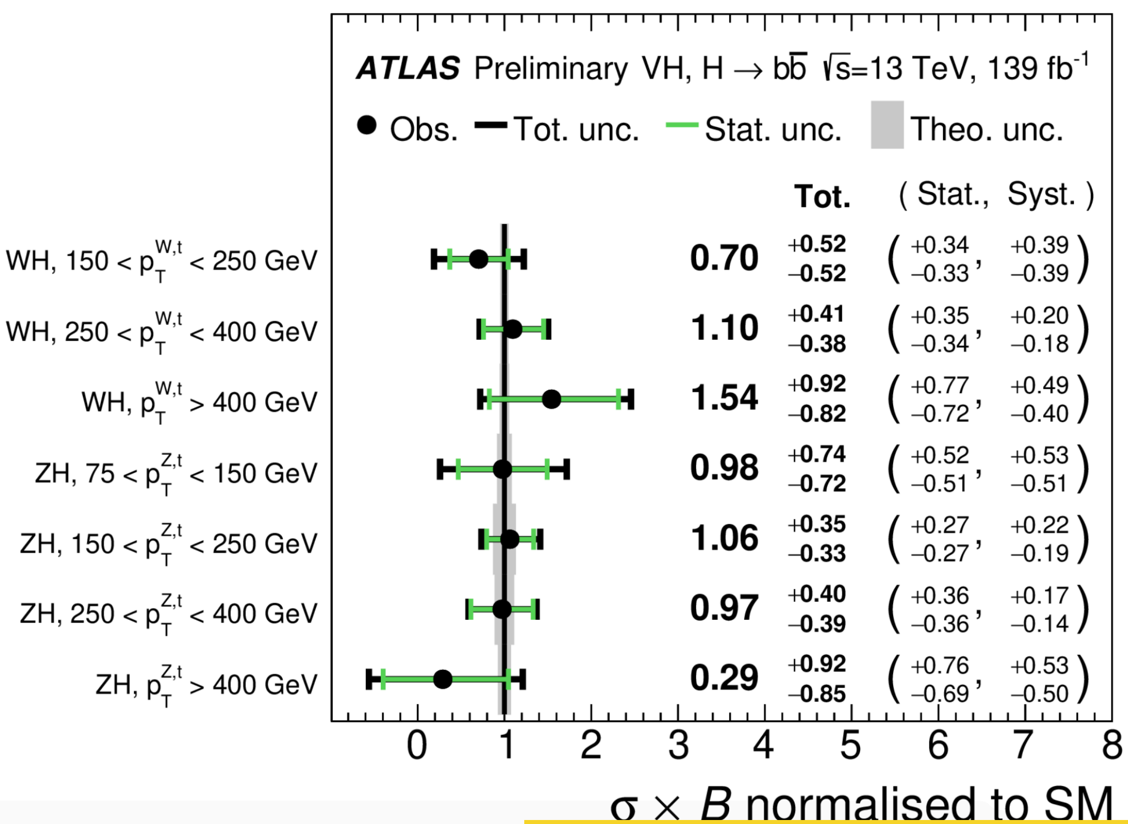


- ➡ LHC Higgs WG1 - VH sub-group [twiki], WG1 fall meeting at this link
- ➡ Quick overview of the state-of-the-art for ATLAS&CMS VHbb measurements
  - new developments wrt last LHC Higgs WG workshop [Nov 2020] marked as NEW
- ➡ VH signal uncertainties - theoretical developments and feedbacks from the analyses
  - signal uncertainties on STXS measurement [backup slides]
  - ggZH merged predictions
  - review of qqZH processes
  - signal modelling of Hbb branching ratio and decay
- ➡ Background uncertainties - theoretical developments and feedbacks from the analyses
  - a quick glance at the non V+X backgrounds
  - V+jets (V+heavy-flavour) modelling - ATLAS/CMS comparisons and state-of-the art of the investigations
- ➡ Wrapping-up and conclusions

# Run 2 VHbb measurements - the state of the art

CMS 1808.08242

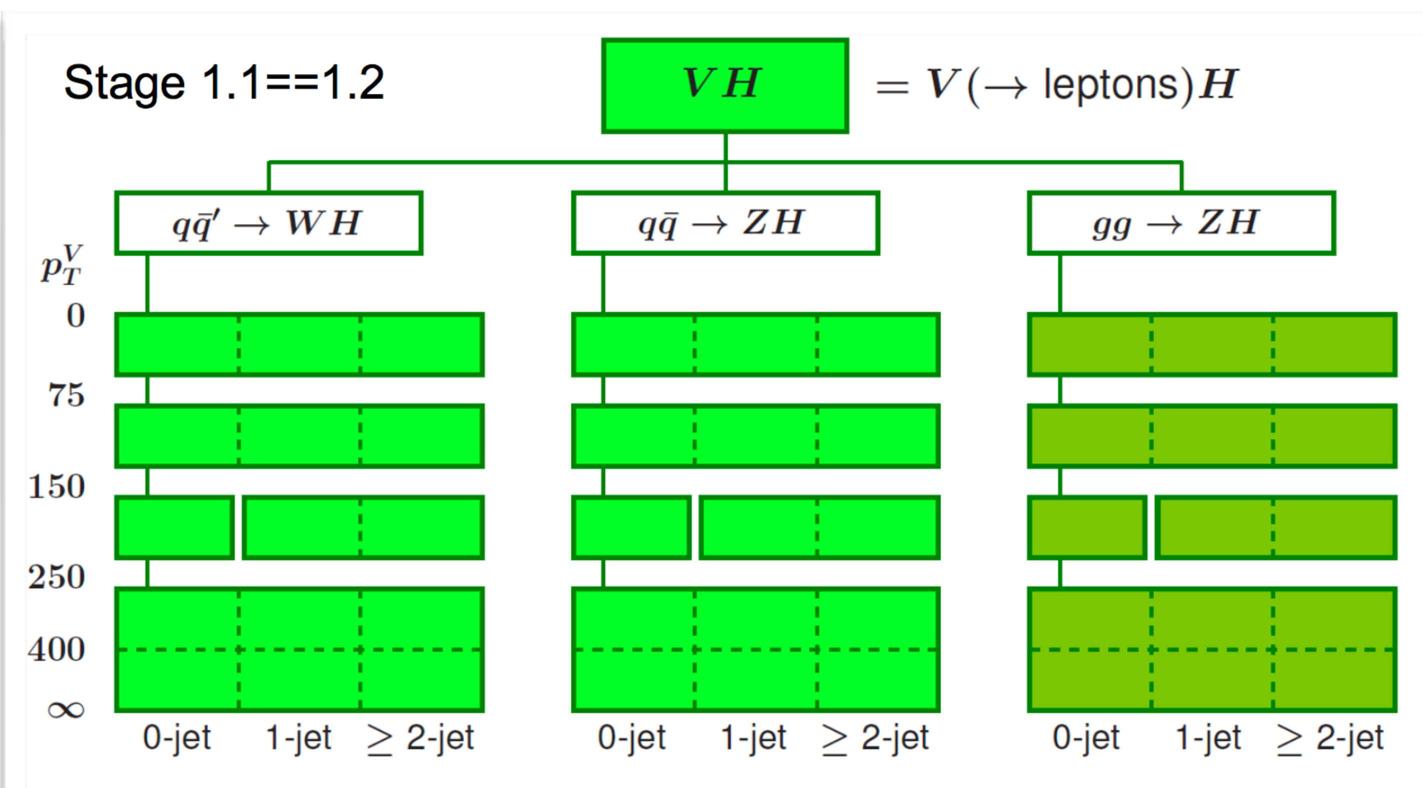



ATLAS-CONF-2021-051

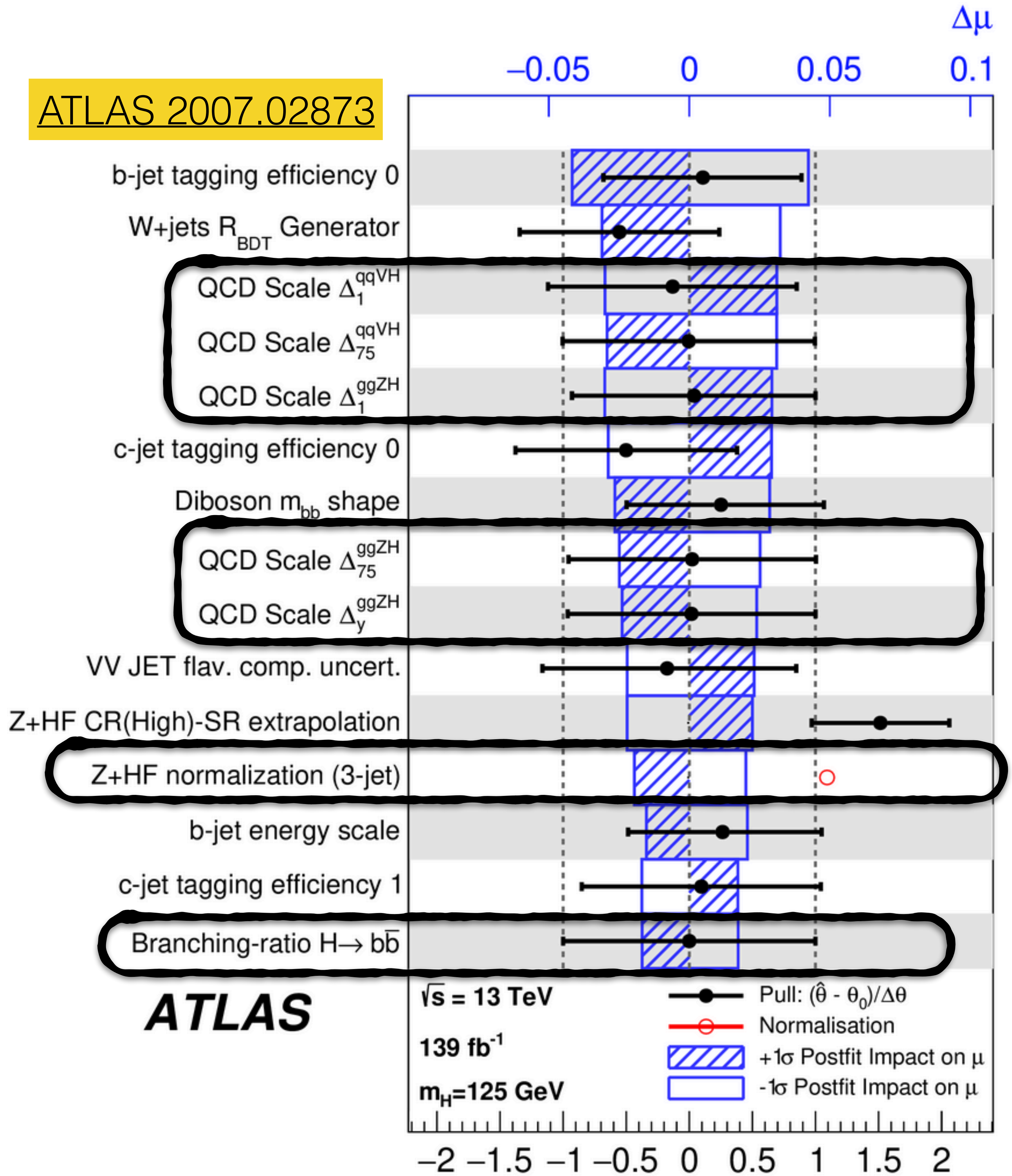
- Evolution of inclusive measurements - STXS approach categorises events at gen-level using analyses observables (ptV, ptH, nJet, …)
- Signal extraction is optimised for kinematic features of specific bin
- Several points of interest for analysis sensitivity (bins to target, what to do with non-sensitive bins, define dedicated bins to be sensitive to NP effects, … )

CERN YR4



# Limitations of the current measurements - the role of the uncertainties

ATLAS 2007.02873




- ➡ Large uncertainties relate to theory modelling
  - VH signal, Hbb decay
  - V+jets (especially V+HF) modeling
- ➡ Will review the status of the investigations for modelling/ predictions as well as the new features currently under study within the VH LHC Higgs sub-group

| Uncertainty source | $\Delta\mu$ | |
|---|---|---|
| Statistical | +0.26 | −0.26 |
| Normalization of backgrounds | +0.12 | −0.12 |
| Experimental | +0.16 | −0.15 |
| b-tagging efficiency and misid | +0.09 | −0.08 |
| V+jets modeling | +0.08 | −0.07 |
| Jet energy scale and resolution | +0.05 | −0.05 |
| Lepton identification | +0.02 | −0.01 |
| Luminosity | +0.03 | −0.03 |
| Other experimental uncertainties | +0.06 | −0.05 |
| MC sample size | +0.12 | −0.12 |
| Theory | +0.11 | −0.09 |
| Background modeling | +0.08 | −0.08 |
| Signal modeling | +0.07 | −0.04 |
| Total | +0.35 | −0.33 |

CMS 1808.08242

ETH zürich

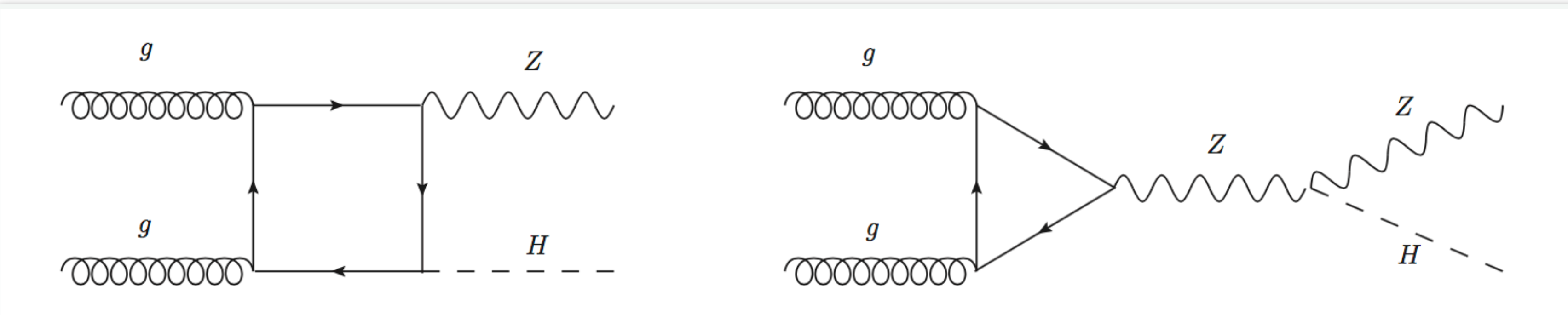
# Signal modeling

# Signal modeling of ggZH process

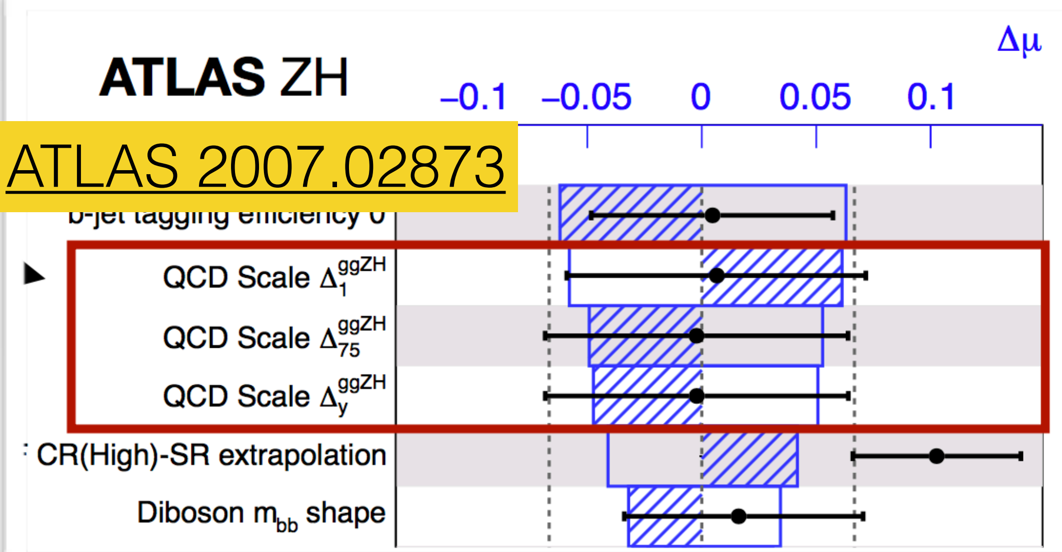
ATLAS 2007.02873

| | Signal |
|---|---|
| Cross-section (scale) | 0.7% ($qq$), 25% ($gg$) |
| $H \to b\bar{b}$ branching fraction | 1.7% |
| Scale variations in STXS bins | 3.0%–3.9% ($qq \to WH$), 6.7%–12% ($qq \to ZH$), 37%–100% ($gg \to ZH$) |
| PS/UE variations in STXS bins | 1%–5% for $qq \to VH$, 5%–20% for $gg \to ZH$ |
| PDF+$\alpha_S$ variations in STXS bins | 1.8%–2.2% ($qq \to WH$), 1.4%–1.7% ($qq \to ZH$), 2.9%–3.3% ($gg \to ZH$) |
| $m_{bb}$ from scale variations | M+S ($qq \to VH$, $gg \to ZH$) |
| $m_{bb}$ from PS/UE variations | M+S |
| $m_{bb}$ from PDF+$\alpha_S$ variations | M+S |
| $p_T^V$ from NLO EW correction | M+S |

**ATLAS & CMS:**
Powheg ggZH@LO in QCD




➡ Scale uncertainties are quite large - full NLO calculation important to mitigate effect associated to scale variations




ATLAS 2007.02873

ggZH uncertainties (QCD scale) largely impactful at pre- and post-fit level

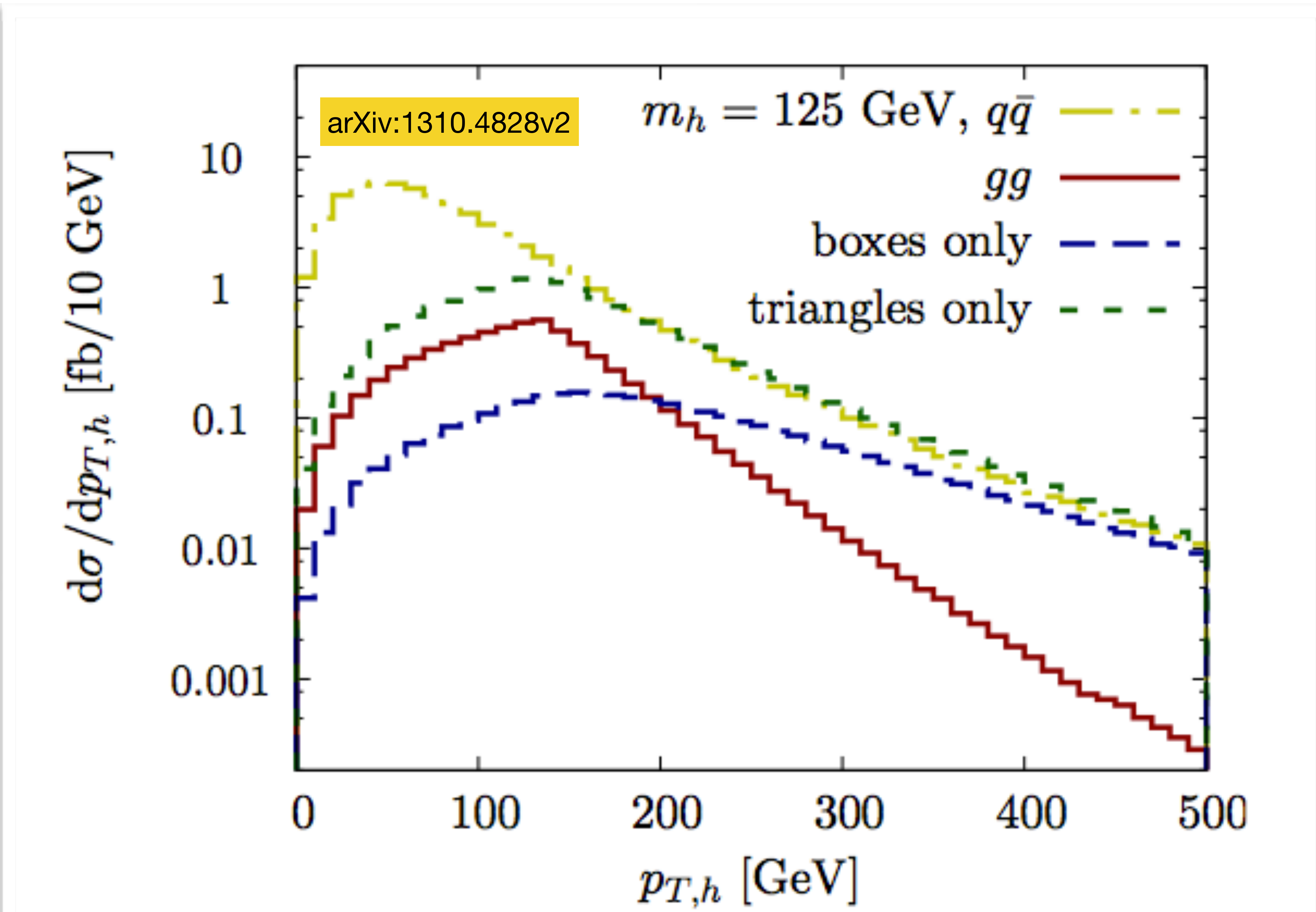
| Process | Cross Section (pb) | +QCD Scale % | -QCD Scale % | ±(PDF+$\alpha_s$) % | ±PDF % | ±$\alpha_s$ % |
|---|---|---|---|---|---|---|
| pp→ZH | 0.8839 | +3.8% | -3.1% | ±1.6% | ±1.3% | ±0.9% |
| qq/qg→ZH, gg→HZ+qqbar (all but gg→ZH) | 0.7612 | +0.5% | -0.6% | ±1.9% | ±1.7% | ±0.9% |
| gg→ZH | 0.1227 | +25.1% | -18.9% | ±2.4% | ±1.8% | ±1.6% |

CERN YR4

# Signal modeling of ggZH process



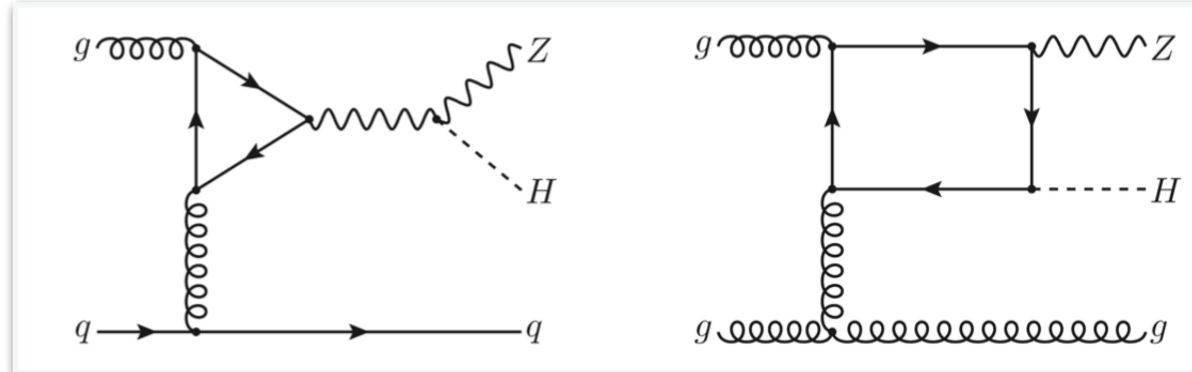
- Important to account for ggZH contribution as part of inclusive ZH signal cross-section
  - XS(ggZH) 15% of inclusive ZH xsec - strong Hpt dependency and enhances contribution in medium VPT range - threshold effect at m(VH)

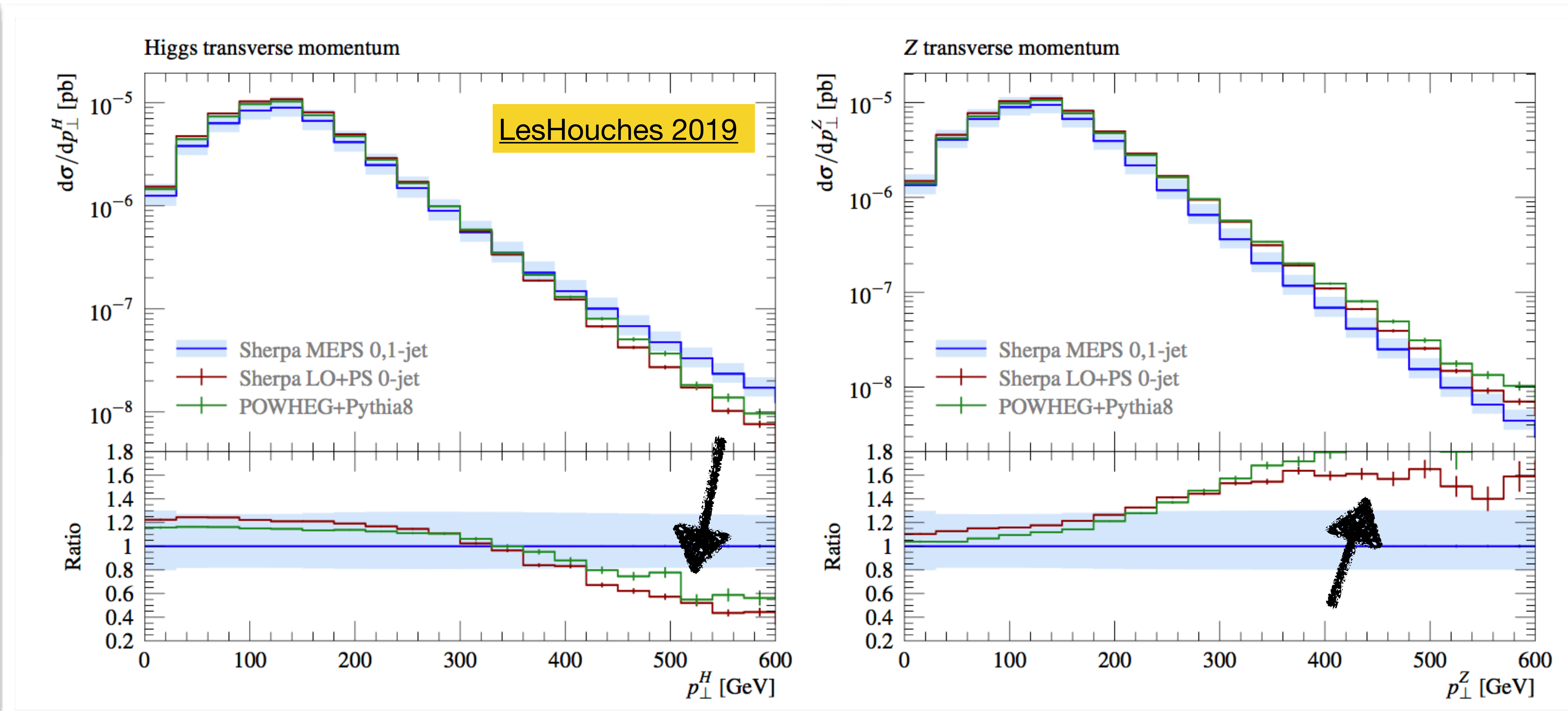


# Signal modeling of ggZH process - hard scattering

➡ Adding 2→3 processes, i.e. gg→ZH+0,1j merged prediction (Sherpa & MC@NLO).

- sizeable modifications in ptH/ptV spectra

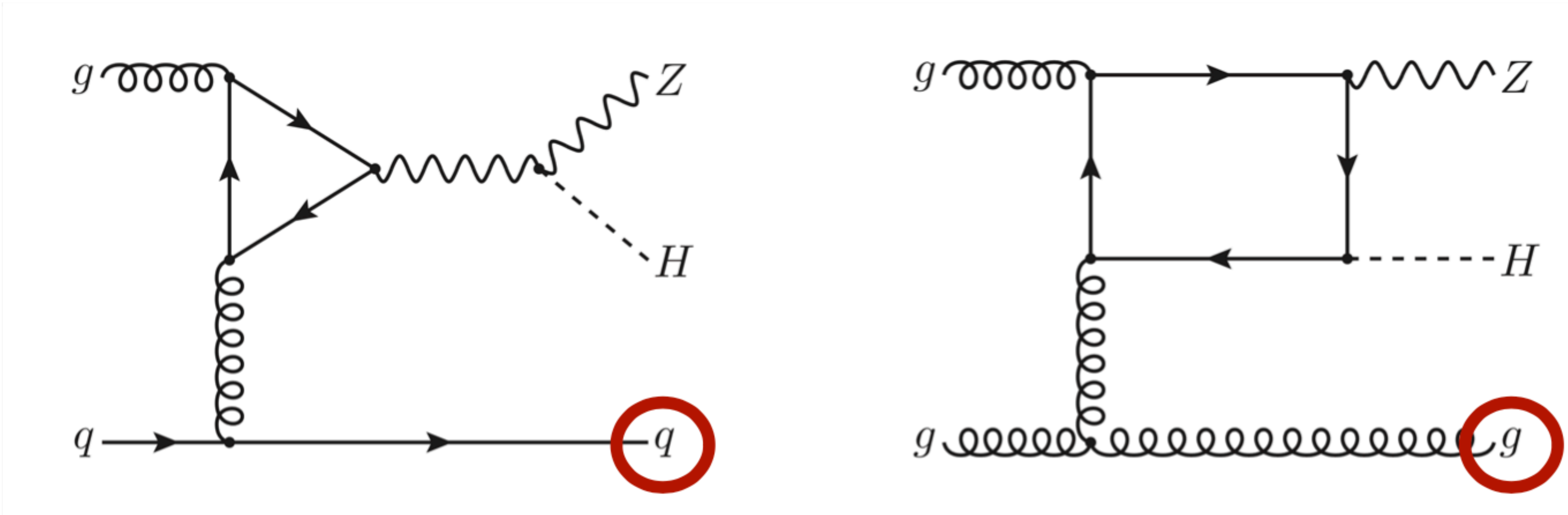



➡ Increase of QCD scale uncertainties in 2→3 processes wrt Powheg+Pythia 2→2 (23% →38% on total ggZH cross-section)

# NEW Signal modeling of ggZH process in ATLAS

➡ Goal: improve modelling of ggZH kinematics: add 2→3 process into matrix element (LO)




➡ Sherpa:

- implemented in ATLAS production environment based on LesHouches setup, performed ATLAS validation
- missing information in truth record, not used for STXS classification paper

➡ MadGraph:

- Developed MadGraph+P8 implementation and full ATLAS validation
- STXS categorisation possible with HXSWG Rivet routine

Ongoing comparison in ATLAS of MadGraph and Sherpa: central values found to be compatible within scale variations but Sherpa assigns larger scale variations than Madgraph (40% vs 25%)

# Signal modeling of qqZH process

ATLAS 2007.02873

| | Signal |
|---|---|
| Cross-section (scale) | 0.7% ($qq$), 25% ($gg$) |
| $H \to b\bar{b}$ branching fraction | 1.7% |
| Scale variations in STXS bins | 3.0%–3.9% ($qq \to WH$), 6.7%–12% ($qq \to ZH$), 37%–100% ($gg \to ZH$) |
| PS/UE variations in STXS bins | 1%–5% for $qq \to VH$, 5%–20% for $gg \to ZH$ |
| PDF+$\alpha_S$ variations in STXS bins | 1.8%–2.2% ($qq \to WH$), 1.4%–1.7% ($qq \to ZH$), 2.9%–3.3% ($gg \to ZH$) |
| $m_{bb}$ from scale variations | M+S ($qq \to VH$, $gg \to ZH$) |
| $m_{bb}$ from PS/UE variations | M+S |
| $m_{bb}$ from PDF+$\alpha_S$ variations | M+S |
| $p_T^V$ from NLO EW correction | M+S |

ATLAS & CMS:
Powheg
qqZH@MiNLO
(QCD)
VPT reweighting
@NLO for EW

- Relatively large variations of ATLAS PS/UE uncertainties originated by two-point systematics difference between Pythia8 and Herwig 7

- Additional predictions available for qqZH modelling:
  - qqVH available for Powheg @NLO QCD+EW
  - ZH @ NNLOPS (reweighting from Powheg to MCFM), WH+jets @ NNLO
  - being tested and included in CMS/ATLAS analyses

CMS 1808.08242

| Uncertainty source | $\Delta\mu$ | |
|---|---|---|
| Statistical | +0.26 | −0.26 |
| Normalization of backgrounds | +0.12 | −0.12 |
| Experimental | +0.16 | −0.15 |
| b-tagging efficiency and misid | +0.09 | −0.08 |
| V+jets modeling | +0.08 | −0.07 |
| Jet energy scale and resolution | +0.05 | −0.05 |
| Lepton identification | +0.02 | −0.01 |
| Luminosity | +0.03 | −0.03 |
| Other experimental uncertainties | +0.06 | −0.05 |
| MC sample size | +0.12 | −0.12 |
| Theory | +0.11 | −0.09 |
| Background modeling | +0.08 | −0.08 |
| Signal modeling | +0.07 | −0.04 |
| Total | +0.35 | −0.33 |

# Signal modeling of branching ratios and Hbb decays



ATLAS 2007.02873

| | Signal |
|---|---|
| Cross-section (scale) | 0.7% ($qq$), 25% ($gg$) |
| $H \to b\bar{b}$ branching fraction | 1.7% |
| Scale variations in STXS bins | 3.0%–3.9% ($qq \to WH$), 6.7%–12% ($qq \to ZH$), 37%–100% ($gg \to ZH$) |
| PS/UE variations in STXS bins | 1%–5% for $qq \to VH$, 5%–20% for $gg \to ZH$ |
| PDF+$\alpha_S$ variations in STXS bins | 1.8%–2.2% ($qq \to WH$), 1.4%–1.7% ($qq \to ZH$), 2.9%–3.3% ($gg \to ZH$) |
| $m_{bb}$ from scale variations | M+S ($qq \to VH$, $gg \to ZH$) |
| $m_{bb}$ from PS/UE variations | M+S |
| $m_{bb}$ from PDF+$\alpha_S$ variations | M+S |
| $p_T^V$ from NLO EW correction | M+S |

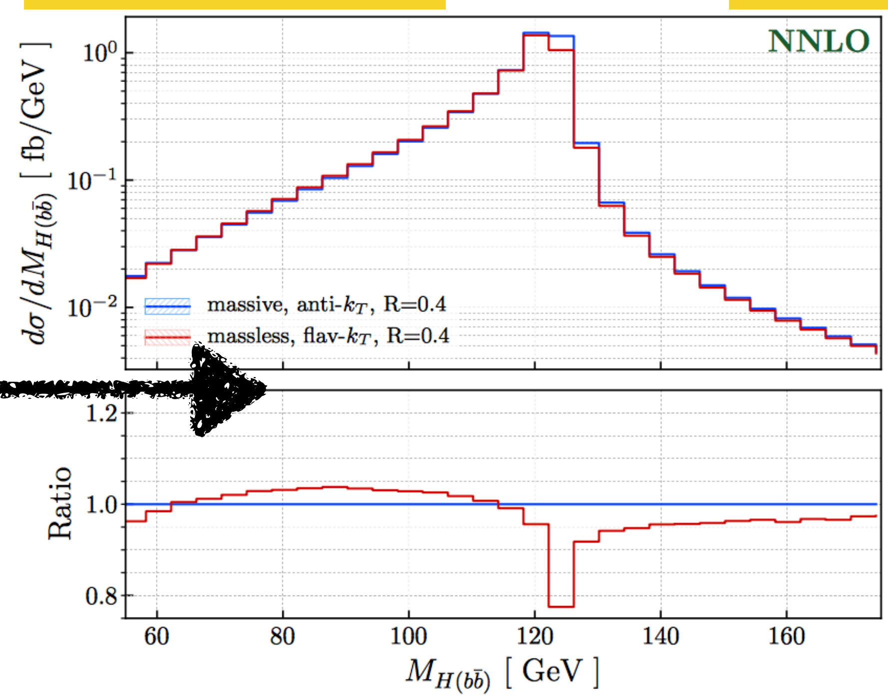
**ATLAS & CMS**: Hbb BR/decay using Pythia8 for the decay. Uncertainty on BR from YR

Decay in Hbb known at N3LO with inclusion of effects due to finite botton-quark mass in NNLO predictions

R. Mondini, C. Williams et al

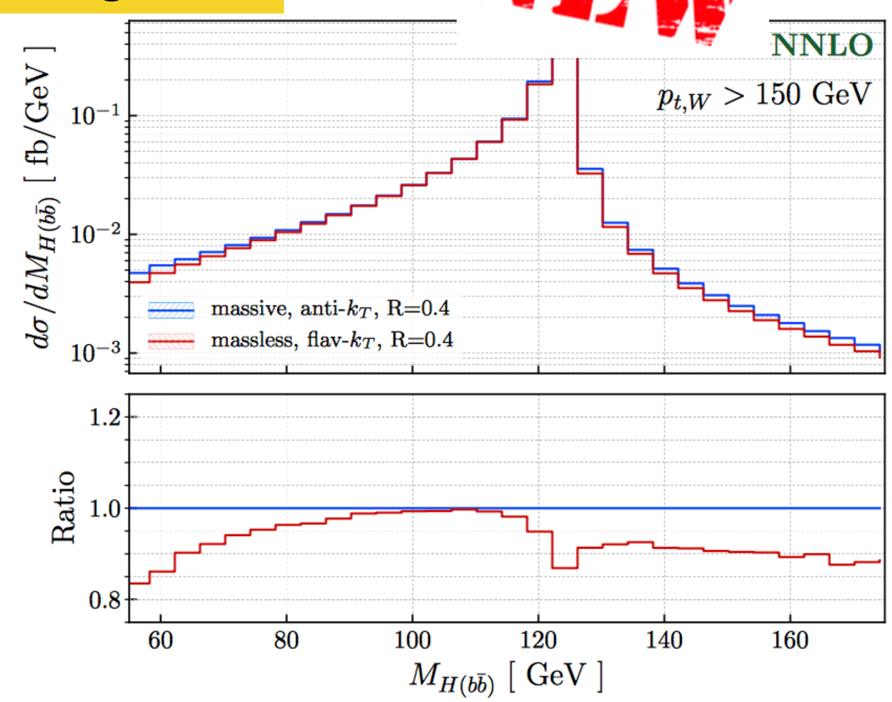
- Finite b-quark mass calculation on NNLO has large impact on m(bb) lineshape modelling especially in high pt
- Interesting to check impact on analysis to test if the effect is already accounted for by PS

W. Astill et al.




A. Behring et al

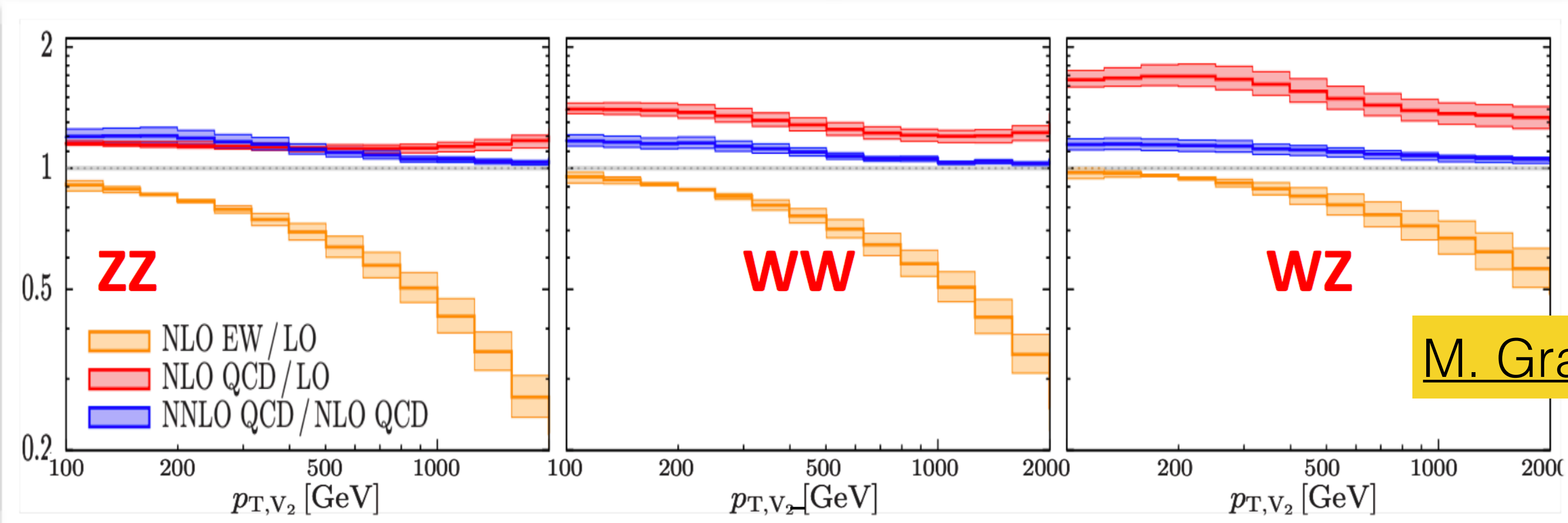
NEW



# Background modeling

# Non V+X backgrounds: VV, ttbar, single-top modeling

- Diboson cross-section constrained with a dedicated cross-check VZ→bb (WZ/ZZ) analysis
  - signal extraction compatible with SM and good modelling of observables, no concern on prediction
  - NLO VV production used by CMS with fully leptonic diboson EWK correctionsn




M. Grazzini et al.

- ttbar and single-top modelling achieved with different approaches in ATLAS and CMS
  - CMS: MC-based analysis (PP8) and dedicated tt-enriched control regions to constrain shape and normalisation with data; ATLAS: MC-based measurement with addition of data-driven techniques in 2lep channel
  - Studies by ATLAS on 0/1lep and boosted Hbb regarding tt ptV spectrum using PP8 [link]
  - PP8 DS found to provide better agreement with data
  - no problematic issues in MC predictions - post-fit process rate parameters for ttbar processes always close to unity and largely constrained due to excellent purity of tt-enriched regions

# V+jets modeling



ATLAS 2007.02873

| Source of uncertainty | | $\sigma_\mu$ VH | WH | ZH |
|---|---|---|---|---|
| Total | | 0.177 | 0.260 | 0.240 |
| Statistical | | 0.115 | 0.182 | 0.171 |
| Systematic | | 0.134 | 0.186 | 0.168 |
| **Statistical uncertainties** | | | | |
| Data statistical | | 0.108 | 0.171 | 0.157 |
| $t\bar{t}$ $e\mu$ control region | | 0.014 | 0.003 | 0.026 |
| Floating normalisations | | 0.034 | 0.061 | 0.045 |
| **Experimental uncertainties** | | | | |
| Jets | | 0.043 | 0.050 | 0.057 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ | | 0.015 | 0.045 | 0.013 |
| Leptons | | 0.004 | 0.015 | 0.005 |
| $b$-tagging | $b$-jets | 0.045 | 0.025 | 0.064 |
| | $c$-jets | 0.035 | 0.068 | 0.010 |
| | light-flavour jets | 0.009 | 0.004 | 0.014 |
| Pile-up | | 0.003 | 0.002 | 0.007 |
| Luminosity | | 0.016 | 0.016 | 0.016 |
| **Theoretical and modelling uncertainties** | | | | |
| Signal | | 0.072 | 0.060 | 0.107 |
| $Z$ + jets | | 0.032 | 0.013 | 0.059 |
| $W$ + jets | | 0.040 | 0.079 | 0.009 |
| $t\bar{t}$ | | 0.021 | 0.046 | 0.029 |
| Single top quark | | 0.019 | 0.048 | 0.015 |
| Diboson | | 0.033 | 0.033 | 0.039 |
| Multi-jet | | 0.005 | 0.017 | 0.005 |
| MC statistical | | 0.031 | 0.055 | 0.038 |

CMS 1808.08242

| Uncertainty source | $\Delta\mu$ | |
|---|---|---|
| Statistical | +0.26 | −0.26 |
| Normalization of backgrounds | +0.12 | −0.12 |
| Experimental | +0.16 | −0.15 |
| b-tagging efficiency and misid | +0.09 | −0.08 |
| V+jets modeling | +0.08 | −0.07 |
| Jet energy scale and resolution | +0.05 | −0.05 |
| Lepton identification | +0.02 | −0.01 |
| Luminosity | +0.03 | −0.03 |
| Other experimental uncertainties | +0.06 | −0.05 |
| MC sample size | +0.12 | −0.12 |
| Theory | +0.11 | −0.09 |
| Background modeling | +0.08 | −0.08 |
| Signal modeling | +0.07 | −0.04 |
| Total | +0.35 | −0.33 |

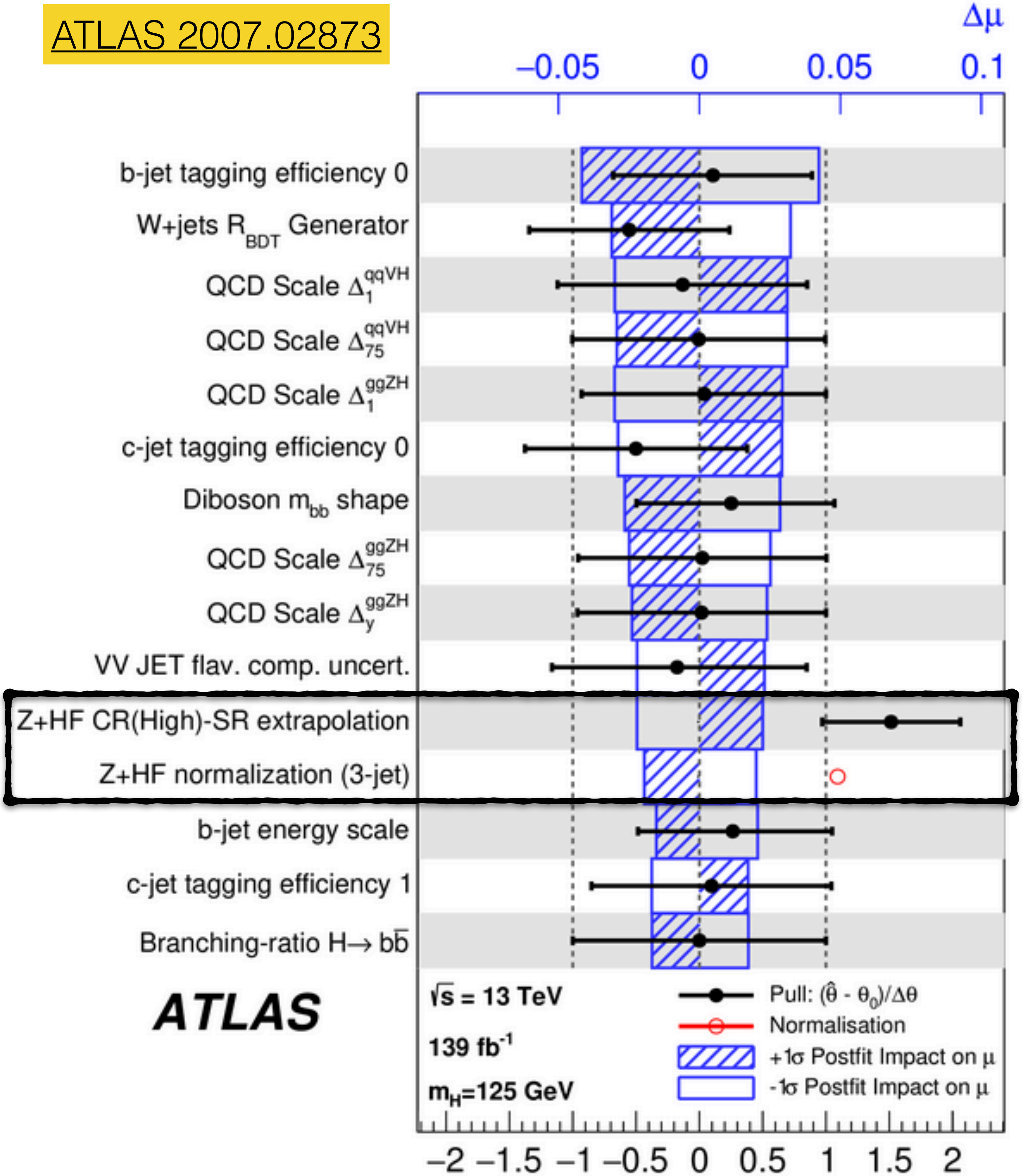
➡ V+heavy-flavour represents the main irreducible background of the VHbb analysis

- theory prediction extremely important for accurate signal extraction
- data constrains prediction of V+jets processes very precisely → MC modelling and choice of systematics variations can impact the measurement significantly

# V+jets modeling (2)

NEW

ATLAS 2007.02873




➡ Different strategies in ATLAS and CMS on V+jets

- modelling of MC prediction
- associated modelling uncertainties
- phase-space for rate-parameter constraints

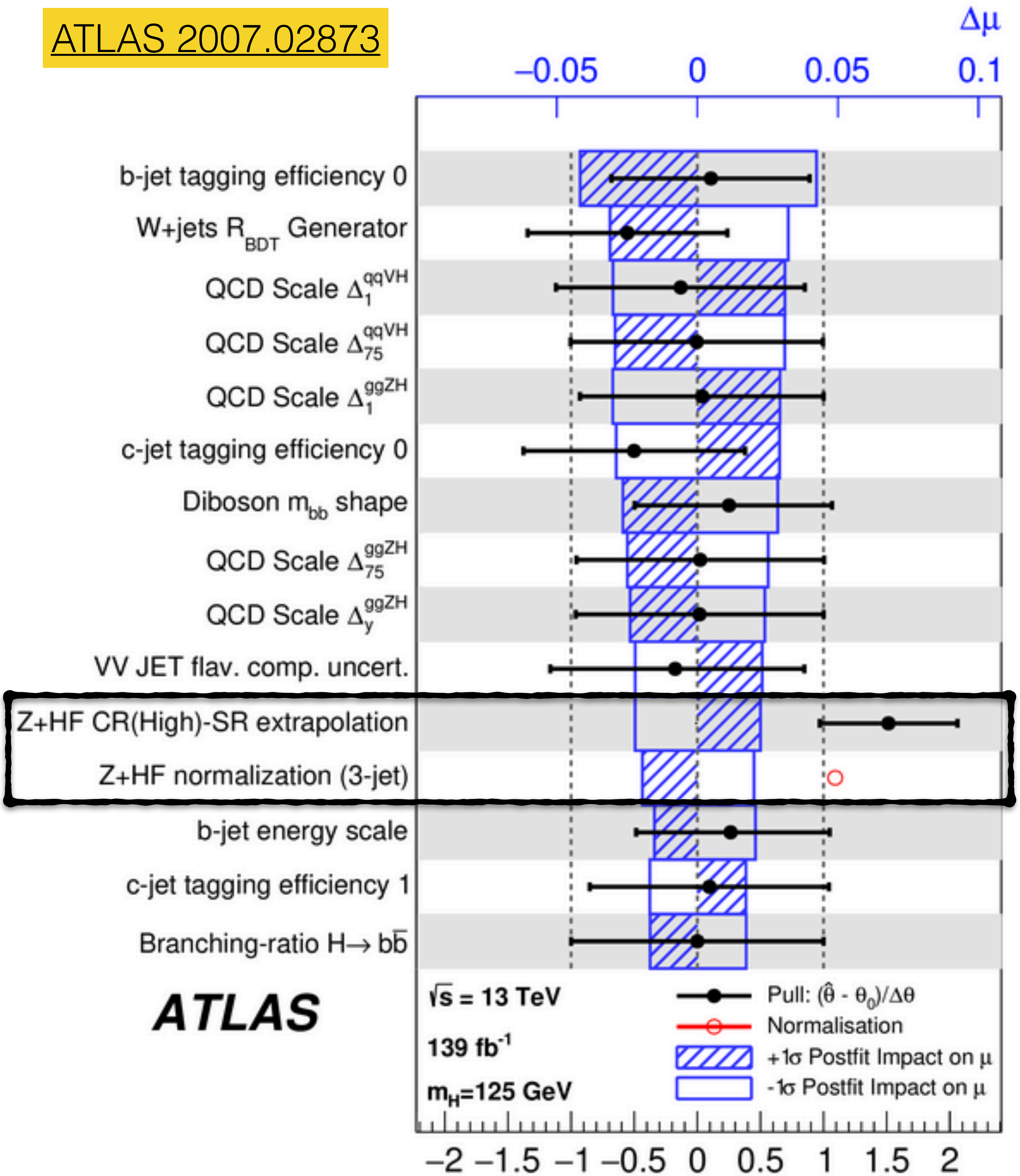
| | CMS | ATLAS |
|---|---|---|
| V+jets nominal | MadGraph V+jets @ LO (HT+bEnriched) | Sherpa V+jets @ NL (0, 1, 2j) +3, 4j@LO |
| Reweightings | VPT EWK corrections+NLO/LO Δη(bb) reweighing + uncertainties | / |

ATLAS/CMS comparison of V+jets predictions is far from trivial!

# V+jets modeling (3)

NEW

ATLAS 2007.02873




➡ Different strategies in ATLAS and CMS on V+jets

- modelling of MC prediction
- associated modelling uncertainties
- phase-space for rate-parameter constraints

| | |
|---|---|
| CMS | HF-enriched CR's based on HFDNN multiclassifer, DeepCSV discriminant in 2lep for separate STXS VPT bins |
| ATLAS | CR's defined using ΔR(bb) for separate jet multiplicities and STXS VPT bins |

ATLAS/CMS comparison of V+jets predictions is far from trivial!

# V+jets modeling (4)

- Normalisation of V+jets background extracted from data by freely-floating the corresponding rate parameters in the simultaneous ML fit of SR and CR
  - significant differences in data wrt MC pre-fit predictions
  - difference in phase-space definition for CR constraints of process scale factors between ATLAS and CMS make the comparison not very trivial
  - need for harmonisation of phase-space definition, objects to ensure meaningful comparison of background process scale factors in ATLAS and CMS

ATLAS 2007.02873

| Process and Category | Normalisation factor |
|---|---|
| $t\bar{t}$ 2-jet | $0.98 \pm 0.09$ |
| $t\bar{t}$ 3-jet | $0.93 \pm 0.06$ |
| $W$ + HF 2-jet | $1.06 \pm 0.11$ |
| $W$ + HF 3-jet | $1.15 \pm 0.09$ |
| $Z$ + HF 2-jet, $75 < p_T^V < 150$ GeV | $1.28 \pm 0.08$ |
| $Z$ + HF 3-jet, $75 < p_T^V < 150$ GeV | $1.17 \pm 0.05$ |
| $Z$ + HF 2-jet, 150 GeV $< p_T^V$ | $1.16 \pm 0.07$ |
| $Z$ + HF 3-jet, 150 GeV $< p_T^V$ | $1.09 \pm 0.04$ |

CMS 1808.08242

| Process | $Z(\nu\nu)H$ | $W(\ell\nu)H$ | $Z(\ell\ell)H$ low-$p_T$ | $Z(\ell\ell)H$ high-$p_T$ |
|---|---|---|---|---|
| W + udscg | $1.04 \pm 0.07$ | $1.04 \pm 0.07$ | | |
| W + b | $2.09 \pm 0.16$ | $2.09 \pm 0.16$ | | |
| W + $b\bar{b}$ | $1.74 \pm 0.21$ | $1.74 \pm 0.21$ | – | – |
| Z + udscg | $0.95 \pm 0.09$ | – | $0.89 \pm 0.06$ | $0.81 \pm 0.05$ |
| Z + b | $1.02 \pm 0.17$ | – | $0.94 \pm 0.12$ | $1.17 \pm 0.10$ |
| Z + $b\bar{b}$ | $1.20 \pm 0.11$ | – | $0.81 \pm 0.07$ | $0.88 \pm 0.08$ |

- Continuing V+HF modeling studies in VH sub-group:
  - support/validation studies for nominal modelling of MC prediction of V+jets - important to check even at pre-fit level the difference in ATLAS/CMS V+jets predictions as done for TT [PUB note]
  - definition of common set of uncertainties associated to MC V+HF modelling

# Wrapping-up and conclusions

➡ Very fruitful interactions between theory and experimental community on several VH(bb)-related processes for signal and background modelling

- STXS-based categorisation and definition of associated uncertainties example of success of WG1 proposal
- VH modeling currently one of the main limitations to VH(bb) precision measurements
- improvement in VH modelling will dramatically impact key uncertainties of the analysis (ggZH signal, V+HF background modeling) and will largely benefit from further developments and studies within the LHC Higgs VH subgroup

➡ VH(bb) analyses have concluded or about to conclude cross-section measurement with full Run 2 dataset

- comparisons of analysis strategies and especially on the treatment of the V+jets background modelling of outmost importance

# Backup slides

# Signal uncertainties on STXS measurement



- QCD scale uncertainties parametrised as overall uncertainty component and migration uncertainties across bin-boundaries (in VPT and jet multiplicity)
    - migration uncertainties calculated as combination of renormalisation and factorisation scale
- Cross-section uncertainty become residual shape uncertainties with coarser STXS bins
- Maximal deviation split presented here