Higgs Hunting 2014

Young Scientist Forum

Search for H→WW in Atlas

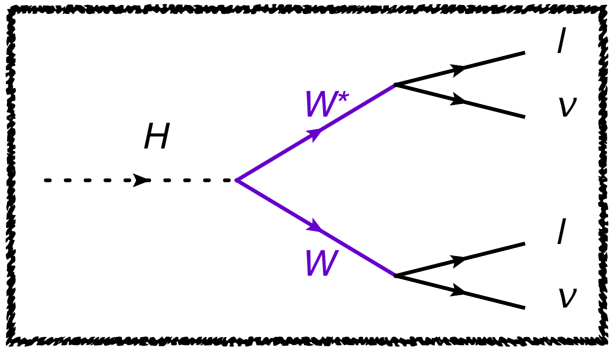
Top Background in HWW VBF Analysis

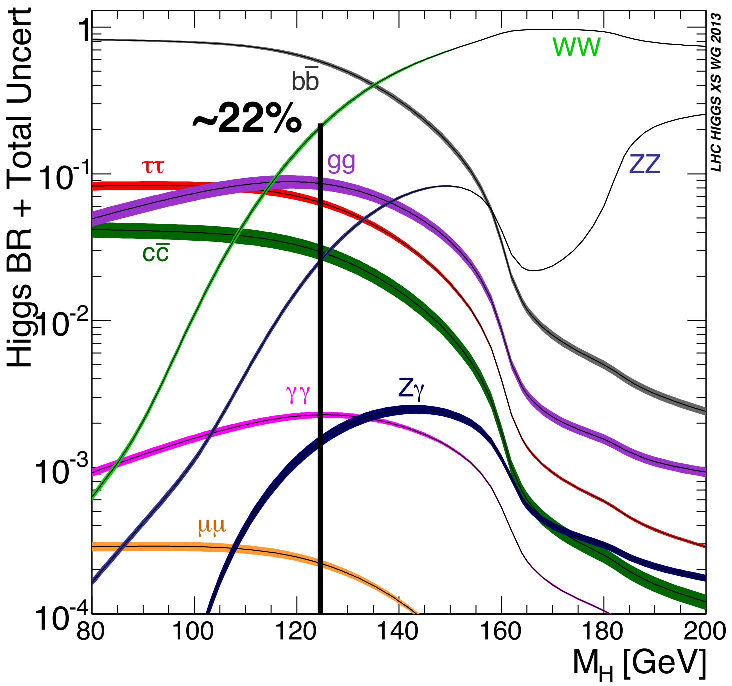
Nika Valenčič

July 21-23, Orsay

H→WW → lνlν

- HWW has a large BR and a clean di-lepton signature

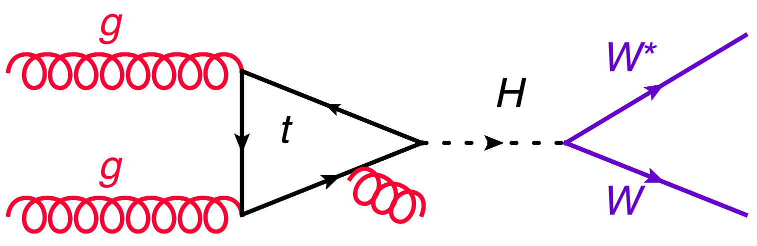




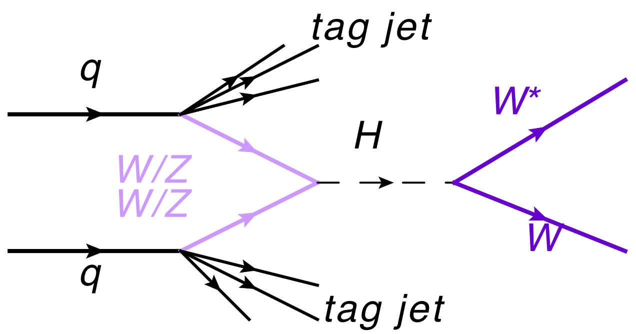
- Due to missing energy not suitable for mass measurements but provides a direct access to coupling measurement.

- HWW analysis separated into two categories depending on Higgs production modes:

- ggF : 0 and 1 jet



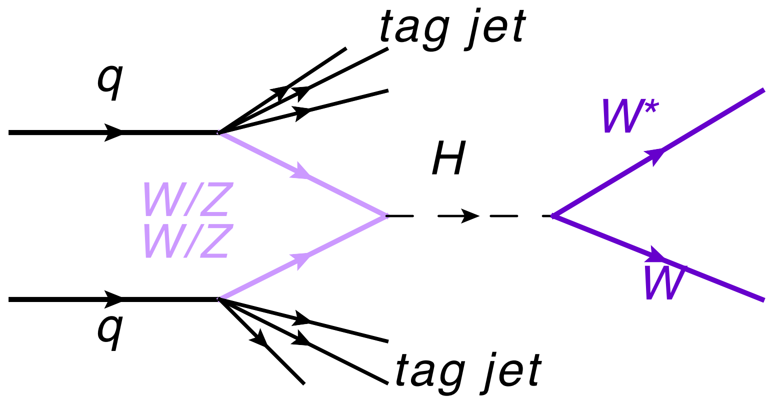
- VBF: 2 jets



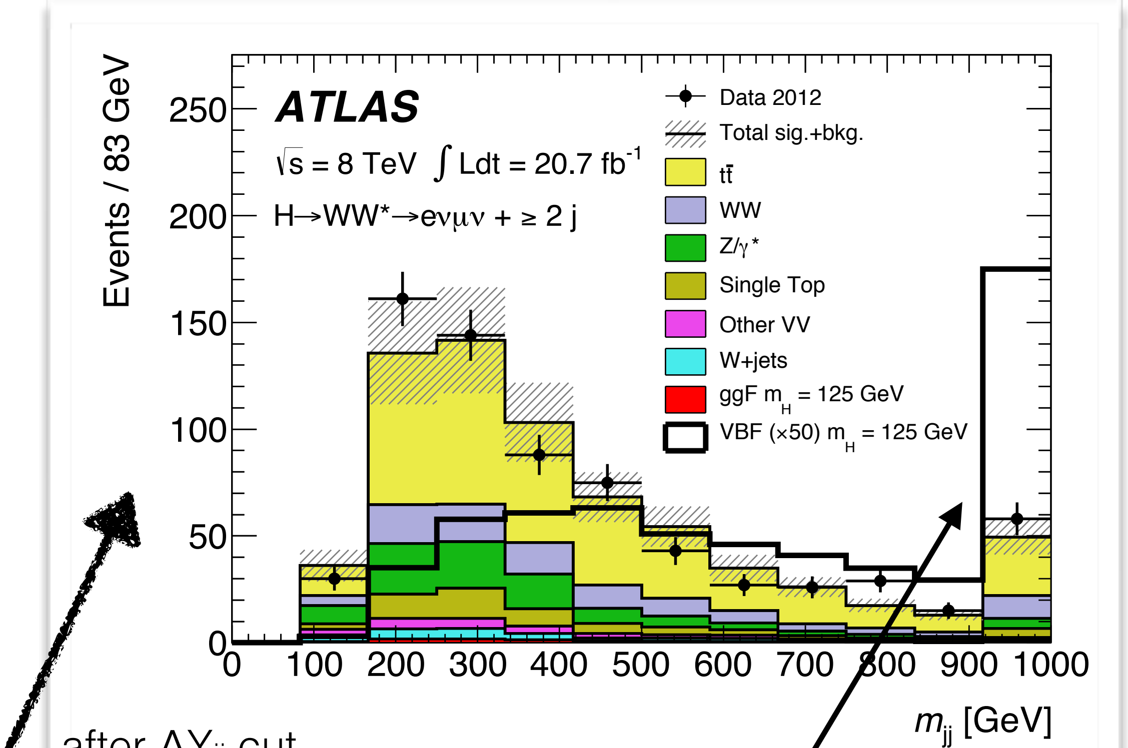
VBF HWW

plots taken from 1307.1427

- VBF - second largest production mode, purely EW process.
- Two vector bosons radiated from the initial-state quarks producing a Higgs boson at the tree level.

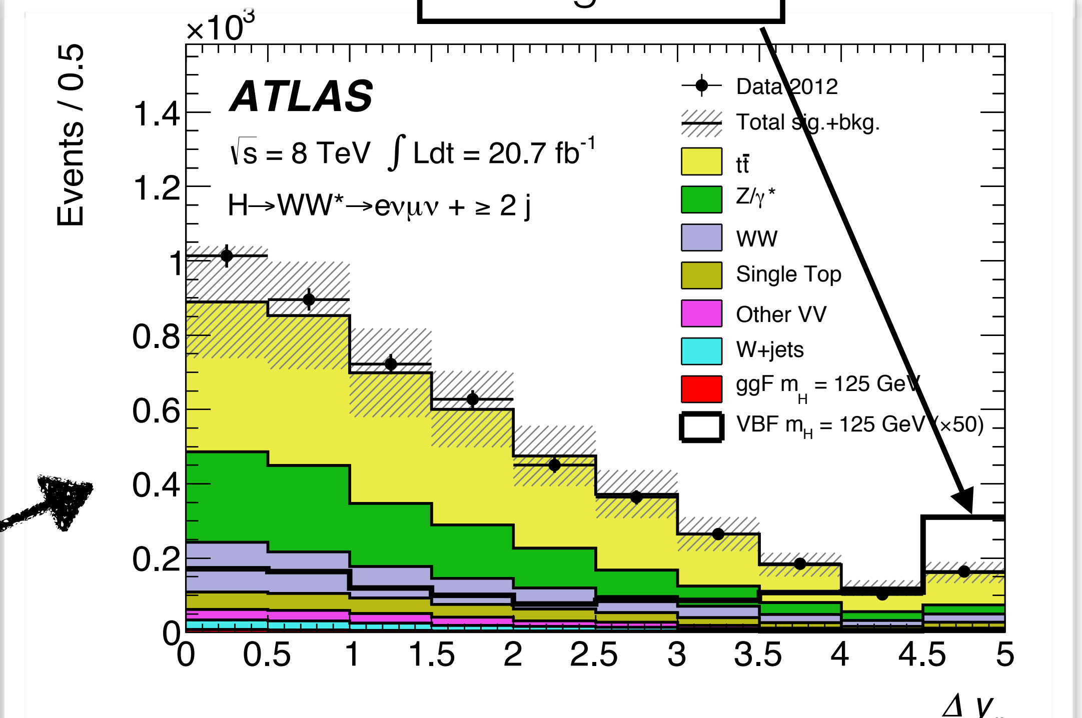


- Specific signature:
  - 2 tag jets
  - large invariant mass of the two tag jets m_{jj}
  - high rapidity gap between the two tag jets ΔY_{jj}
  - no hadronic activity between the two tag jets (CJV)



after ΔY_{jj} cut

VBF signal x 50



after p_T^{tot} cut

VBF HWW analysis

- Apply a set of cuts after requiring two opposite sign leptons with $p_T > 25(15)$ GeV - more info. on slide 14, backup.

- Select events with ≥ 2 jets (VBF production) and split the selection by di-lepton flavours

DF: eμ - most sensitive

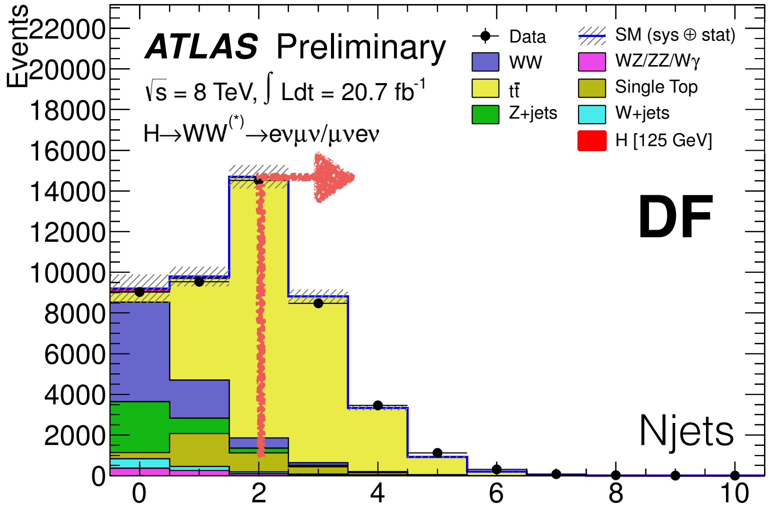
SF: μμ / ee - large DY bkg.

- Introduce control regions for background normalisation - slide 16, backup.

- Transverse mass m_T is used to extract the signal strength.

$$m_T = \sqrt{(E_T^{\ell\ell} + E_T^{\text{miss}})^2 - |\boldsymbol{p}_T^{\ell\ell} + \boldsymbol{E}_T^{\text{miss}}|^2} \qquad E_T^{\ell\ell} = \sqrt{|\boldsymbol{p}_T^{\ell\ell}|^2 + m_{\ell\ell}^2}.$$

SR, after preselection



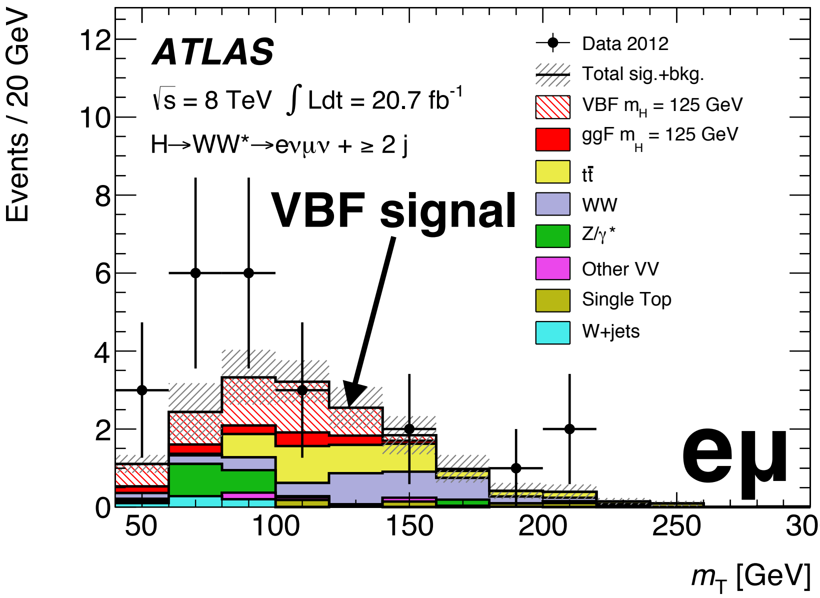
Event yields at the end of event selection:

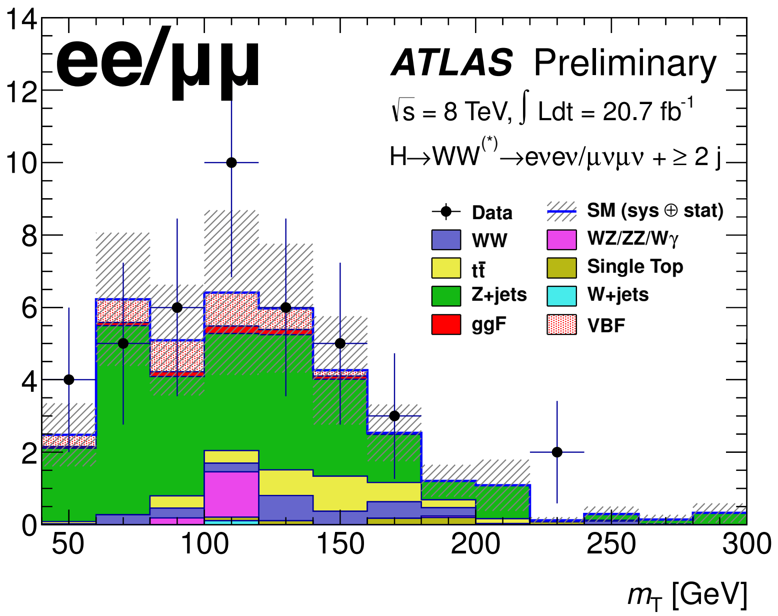
	eμ	μμ + ee
WW	3.5 ± 0.4	2.8 ± 0.3
top	**4.4 ± 0.7**	4.0±0.6
Z/γ	1.9 ± 0.5	**25±2**
W+jets	0.6 ± 0.3	0.1 ± 0.2
VV	0.6 ± 0.2	1.6 ± 1.1
ggF	1.3 ± 0.1	0.7 ± 0.1
Total bkg.	12±1	34±2
VBF signal	5.1 ± 0.1	3.7 ± 0.1
Observed	23	42

m_T in SR

ATLAS-CONF-2013-030

m_T in SR





Top background

- The most dominant background, especially for DF final states.
- Define a top-enriched region (CR), requiring **exactly one** b-tagged jet since VBF phase space cuts prefer ttbar events with 1 true b-jet and one additional (ISR/FSR?) jet as leading jets.
- Normalise top background using data in order to scale MC predictions in SR.

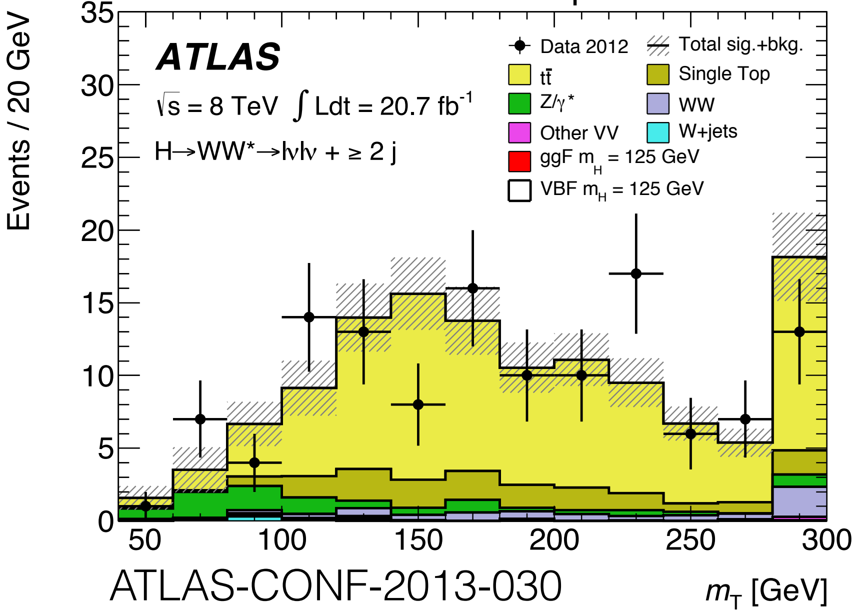
$$N^{\text{SR, est.}}_{\text{top}} = N^{\text{SR, MC}}_{\text{top}} \cdot \underbrace{\frac{N^{\text{CR}}_{\text{data}} - N^{\text{CR}}_{\text{other}}}{N^{\text{CR, MC}}_{\text{top}}}}_{\text{NF}_{\text{top}}} = \underbrace{\frac{N^{\text{SR, MC}}_{\text{top}}}{N^{\text{CR, MC}}_{\text{top}}}}_{\text{CR to SR extrap'n}} \cdot \left(N^{\text{CR}}_{\text{data}} - N^{\text{CR}}_{\text{other}}\right)$$

15% theory uncertainty - comparing MC@NLO (baseline) and Alpgen.

NF= 0.59 ± 0.07(stat.)

mis-modelling of MC@NLO in high m_{jj} and ΔY_{jj} region.

m_T distribution in topCR w. NF



ATLAS-CONF-2013-030

b-tagging, JER

uncertainties on top background:

uncertainties	**%**
theory	15
statistical	10
experimental	29
non-Top bkg	19
total	**39**

Additional jets in top CR

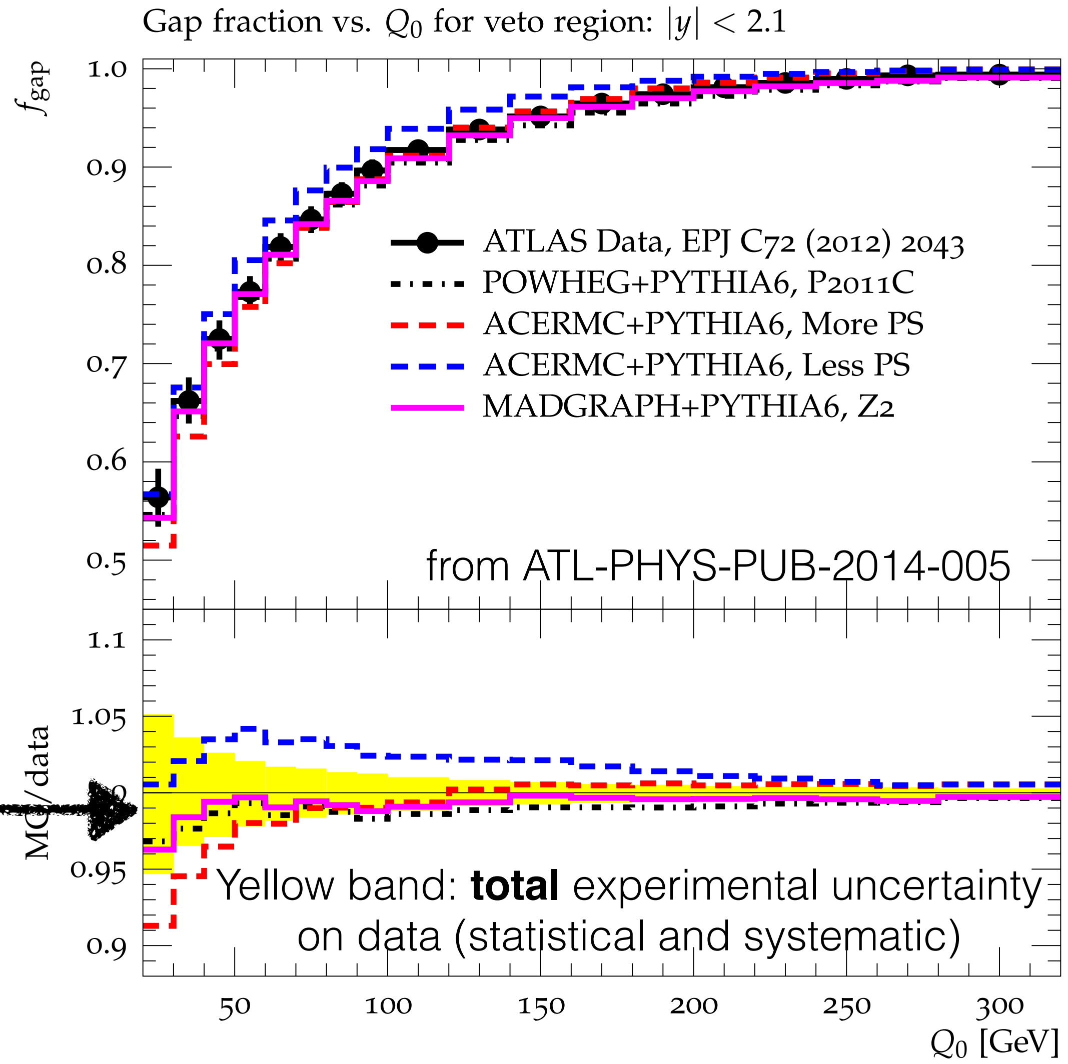
- Top bkg. estimation is sensitive to modelling of additional jets.
- How important is ISR/FSR for VBF HWW? Should we include it in our top systematics?
- Our ISR/FSR strategy:
  - use two samples with increased (morePS) and decreased (lessPS) ISR/FSR - tuning I/FSR parameters in Pythia - slide 21/22, backup.
  - central value is produced by AcerMC +Pythia - **a leading order** generator
  - estimate the I/FSR systematics as:

    0.5* (morePS - lessPS)
- ISR/FSR systematic limited by the systematic and statistical uncertainty from the data-driven gap fraction analysis.

Gap fraction analysis:

Studies the fraction of events which **do not** contain an additional jet with $p_T > Q_0$ (x-axis in the bottom graph) in different η regions.

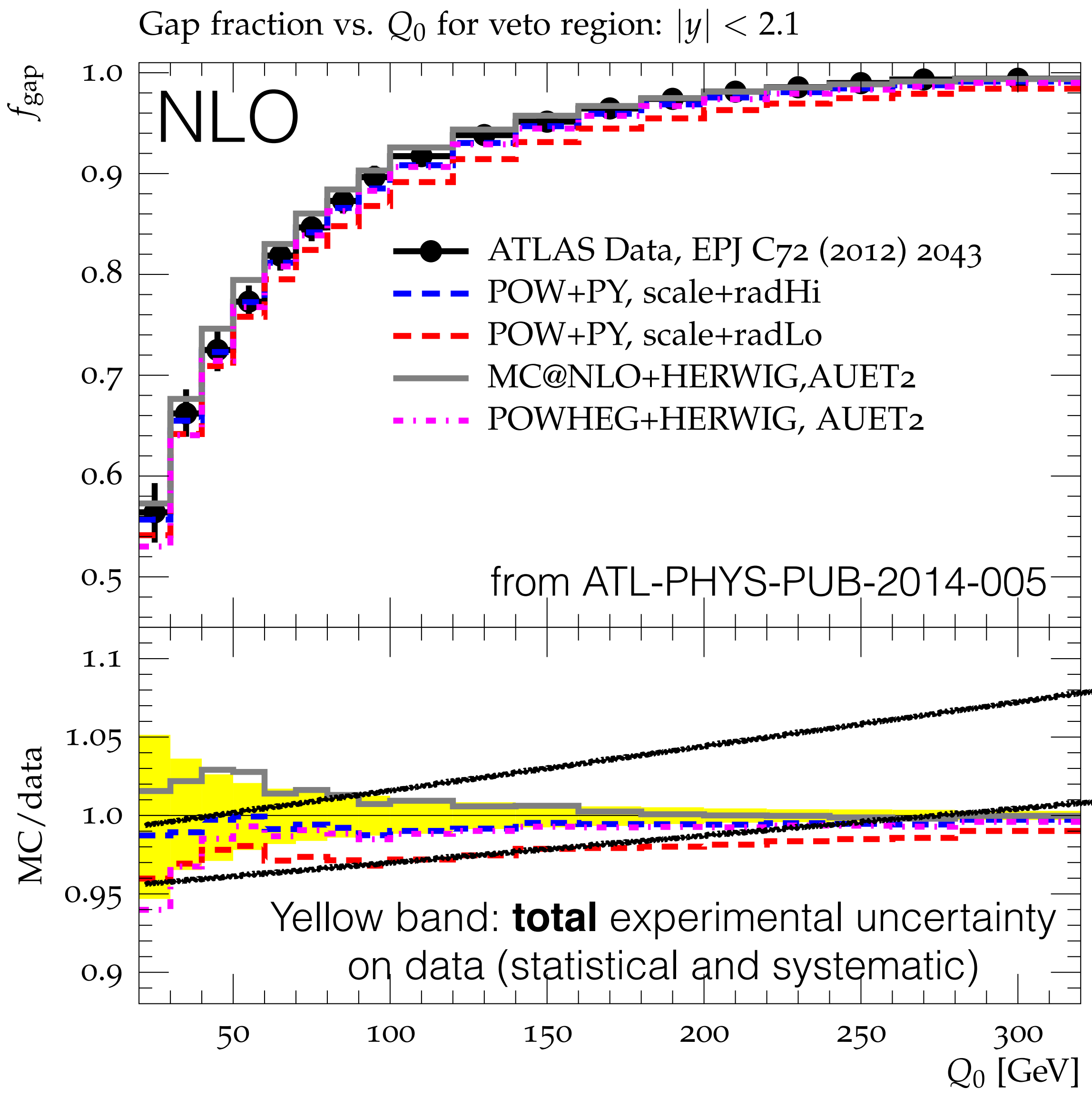
$$f_{gap}(Q_0) = \frac{n_{gap}(Q_0)}{N_{t\bar{t}}}$$



from ATL-PHYS-PUB-2014-005

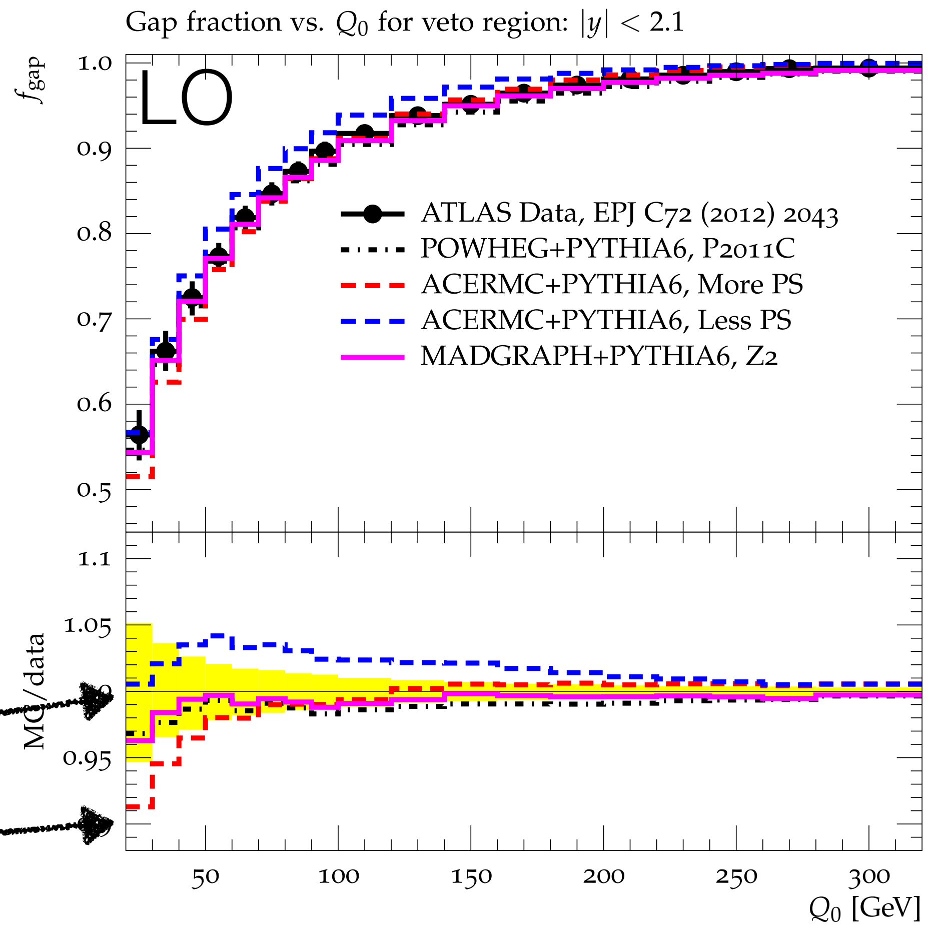
Yellow band: **total** experimental uncertainty on data (statistical and systematic)

ISR/FSR at NLO



from ATL-PHYS-PUB-2014-005

Yellow band: **total** experimental uncertainty on data (statistical and systematic)



Powheg+Pythia6 morePS`(radHi)` and lessPS`(radLO)`Perugia 2012 tune. With varied renormalisation and factorisation scales.

When using an NLO generator (already has 1 parton in ME) the spread between morePS`(radHi)` and lessPS`(radLO)` variations is smaller!

ISR/FSR - a truth level study

- To verify the results, perform a **generator level** study using MC@NLO(v4.10) + Herwig(v6.521,fortran) ttbar sample.

- Control the amount of ISR/FSR by modifying parameters in Herwig (private communication with prof. Bryan Webber):

  ➡ varying the veto on emissions in the parton showers with p_T > SCALUP.

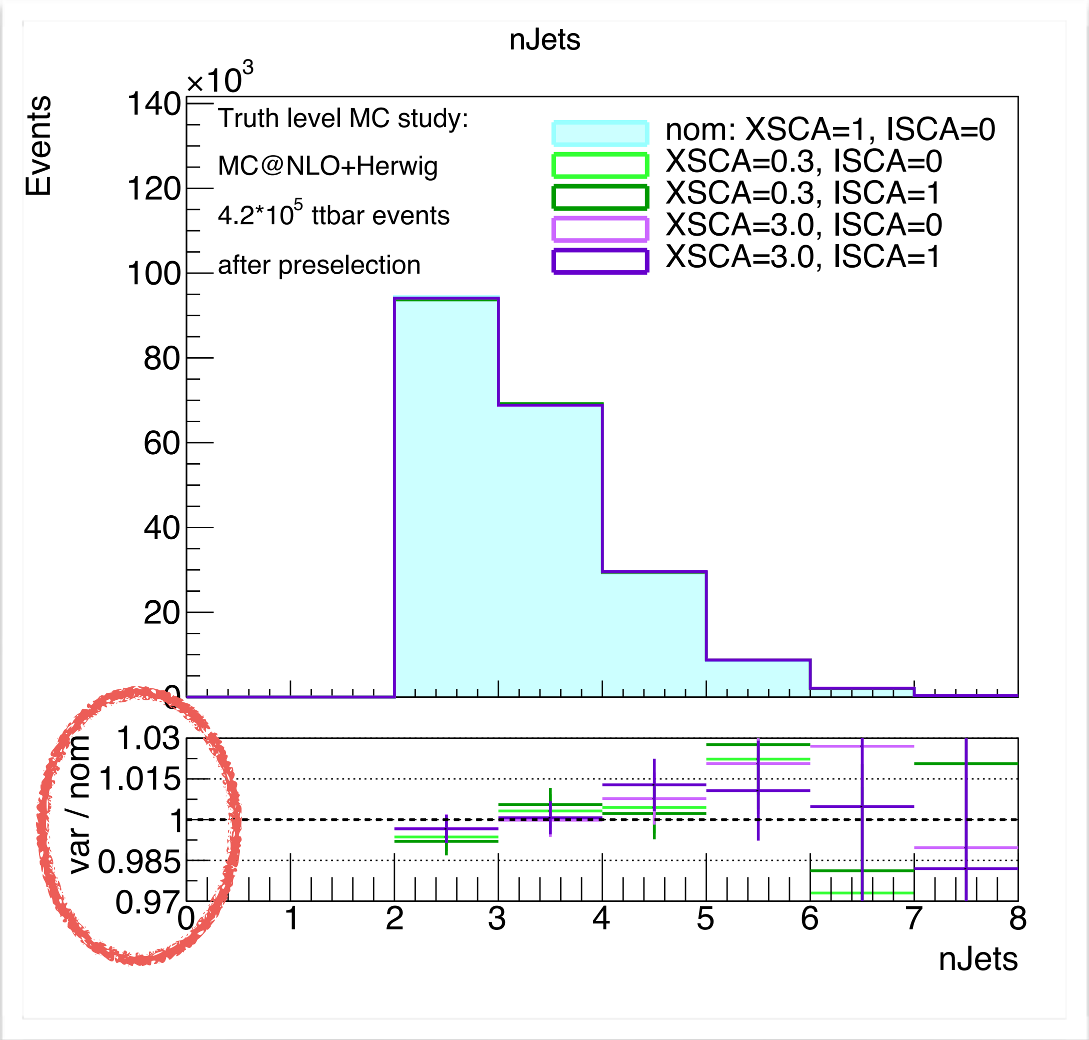
SCALUP = SCALUP * **XSCA**

SCALUP = A ← **ISCA** = 0
SCALUP = B ← **ISCA** = 1

more on slide 22 - backup.

- Variations apply only to the events with a real emission in the NLO part (produced by an NLO ME).

- By construction a beyond NLO effect and expected to be small.



- Maximal variations of XSCA and ISCA have only a small effect on kinematic distributions - more on slide 23, backup.
- When using an NLO generator the effect of ISR/FSR is only **at the percent level.**

Since ISR/FSR systematic is **negligible** compared to other leading sources it is **not included** in top theory systematic.

Summary

- A brief overview of HWW analysis was presented, with a greater focus on VBF production mode and top background.
- The top background theory systematic in VBF HWW analysis is estimated by comparing extrapolation factors for MC@NLO (baselineNLO) and Alpgen (multi-leg) generators and is found to be 15% (39% total uncertainty on the top background).
- **What is the magnitude of ISR/FSR systematic? Should it be included in the top theory systematic?**
- Estimating ISR/FSR systematic by a LO generator gives a larger estimation than when using an NLO generator (already models one additional jet).
- Additional checks were performed by privately produced MC@NLO + Herwig (NLO) ttbar sample at the truth level with varied ISR/FSR contribution.
- Maximal variations of the parameters controlling ISR/FSR in Herwig show discrepancies only at the percent level.
- **ISR/FSR effects on the ttbar background are found to be small** and therefore **do not constitute as a relevant source of systematic uncertainty** on this crucial background in the HWW VBF measurement.

new results coming soon!

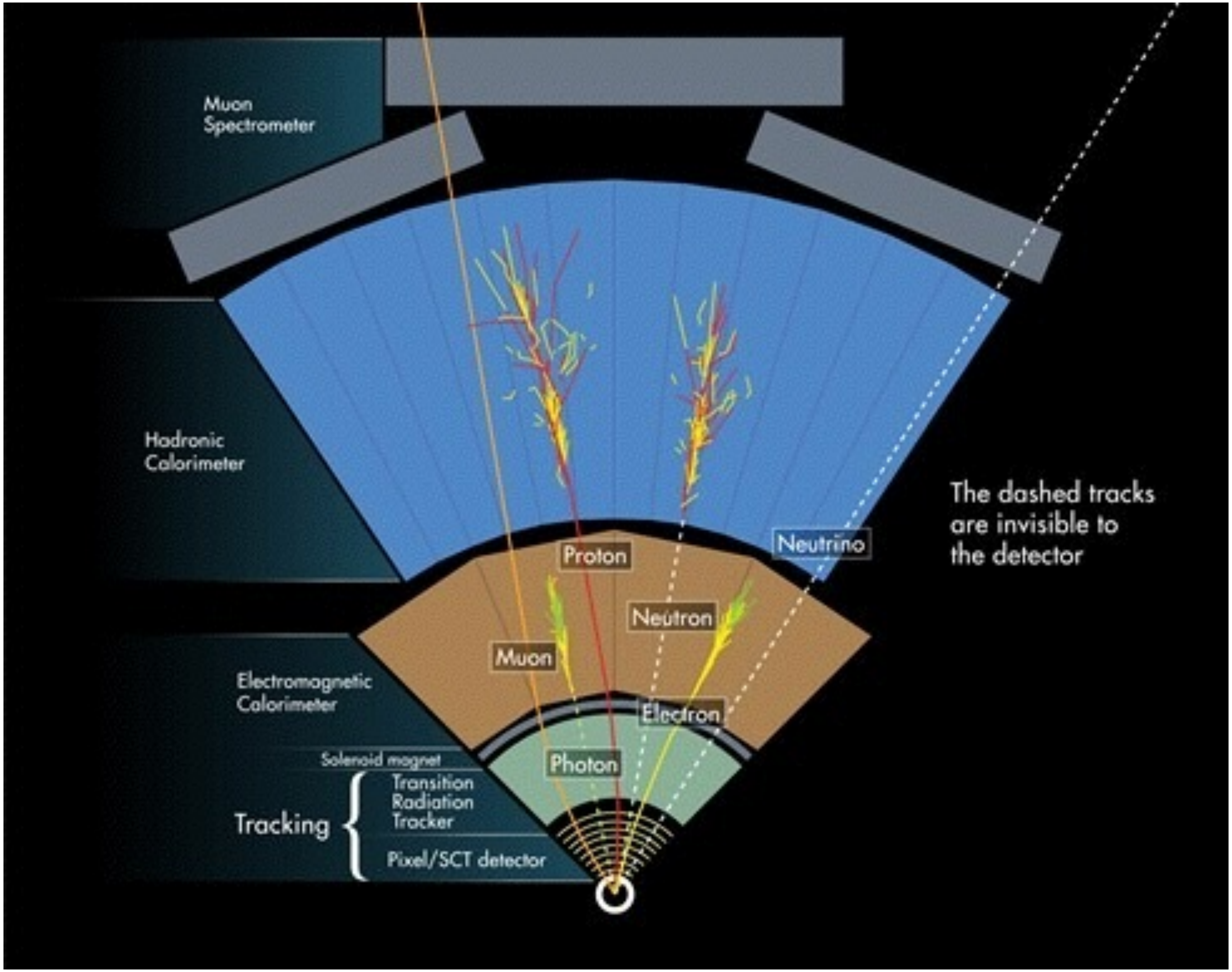
Thank you for your attention!

Backup

References

- ATLAS Higgs Coupling, July 2013 (PLB, arxiv)
- ATLAS HWW note, March 2013 (ATLAS-CONF-2013-030)
- ATLAS gap fraction analyses:
  - May 2014 (ATL-PHYS-PUB-2014-005)
  - May 2013 (ATLAS-PHYS-PUB-2013-005)
  - June 2013 (Eur.Phys.J. C72 (2012) 2043)

ATLAS detector



The dashed tracks are invisible to the detector

N.Valencic, Higgs Hunting 2014, Orsay

Cuts in VBF HWW

VBF cutbased analysis stage	cut	value
preselection	lepton pT	> 25(15) for leading (subleading) lepton
	m	> 10(12) GeV for DF(SF)
	lepton charge	opposite
Z/DY rejection	MET	> 20(45,[35]) GeV for DF(SF [STVF])
	Z/γ-veto (SF only)	$\|m_{ll}$
	Z/γ-ττ veto (DF only)	$\|m_{\tau\tau}$
top bkg. rejection	b-veto	✓
	p_T	< 45GeV
VBF decay topology	m	> 500 GeV
	|ΔY	> 2.8
	CJV	✓ (< 20 GeV)
	OLV	✓
H->WW->lvlv decay topology	|Δϕ	< 1.8
	m	< 60 GeV
	m	discriminating variable

Object selection

leptons:

Single lepton trigger	$p_T > 24$ GeV
Leading lepton	$p_T > 24$ GeV
Subleading lepton	$p_T > 15$ GeV
Electrons	$\|\eta\| < 2.47$
Muons	$\|\eta\| < 2.5$
Tracks	Isolation and impact parameter cuts

jets:

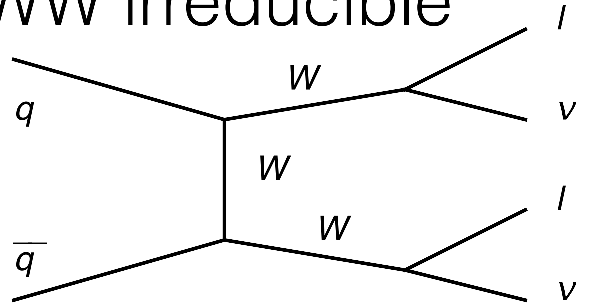
Jets	Anti-k_T with R=0.4
Central Jets in VBF	Anti-k_T with R=0.4, $p_T > 20$ GeV
B-tagging	Neural Network at 85% operating point.

MET:

MET	Calorimeter based
$\mathrm{MET}_{\mathrm{STVF}}$	Calo. weight the unassociated clusters by f_{JVF}
$\mathrm{MET}_{\mathrm{trk}}$	Track based

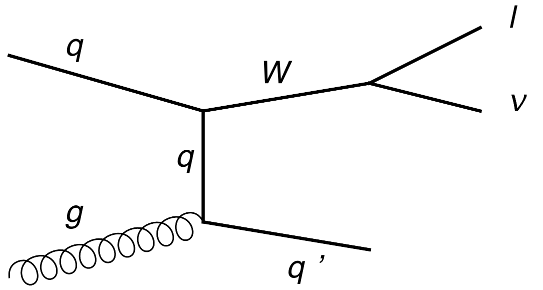
Backgrounds in VBF HWW

WW irreducible



estimated by MC only

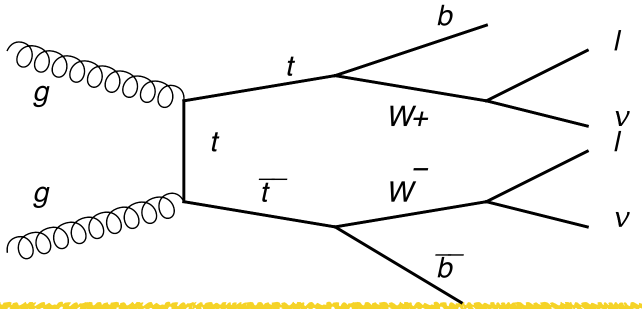
W+jets - faking leptons



estimated by CR

CR: 2nd lepton fails isolation/ID
data-driven estimation -
using di-jet data to
estimate fake-factor

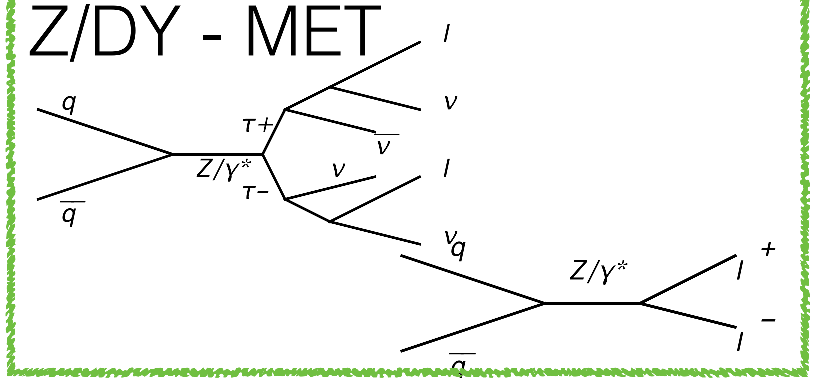
top - b-jets



estimated by CR

CR: 1-bjet, no di-lep cuts
using data-driven
normalisation factor

Z/DY - MET



$Z \rightarrow \tau\tau$ (DF): estimated by CR

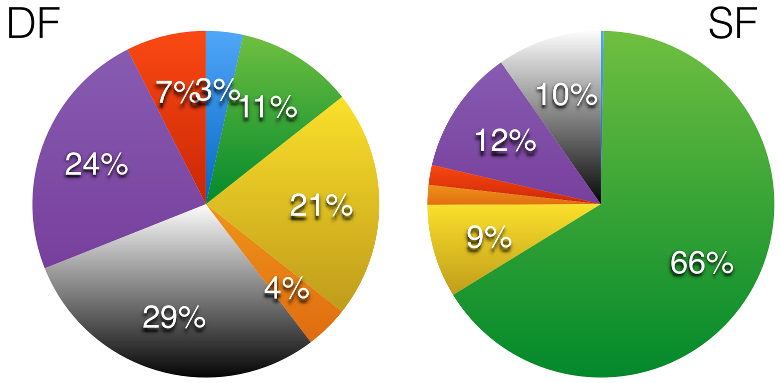
CR: $|\Delta\phi_{ll}|>2.8$+b-jet, p_T^{tot},m_{ll}

$Z \rightarrow ll$ (SF): data-driven estimation

using MET-mll distribution in data to extrapolate to SR

background	eμ	μμ + ee
WW	3.5 ± 0.4	2.8 ± 0.3
top	**4.4 ± 0.7**	4.0±0.6
Z/γ	1.9 ± 0.5	**25±2**
W+jets	0.6 ± 0.3	0.1 ± 0.2
VV	0.6 ± 0.2	1.6 ± 1.1
ggF	1.3 ± 0.1	0.7 ± 0.1
VBF signal	5.1 ± 0.1	3.7 ± 0.1
Total bkg.	12±1	34±2
Observed	23	42

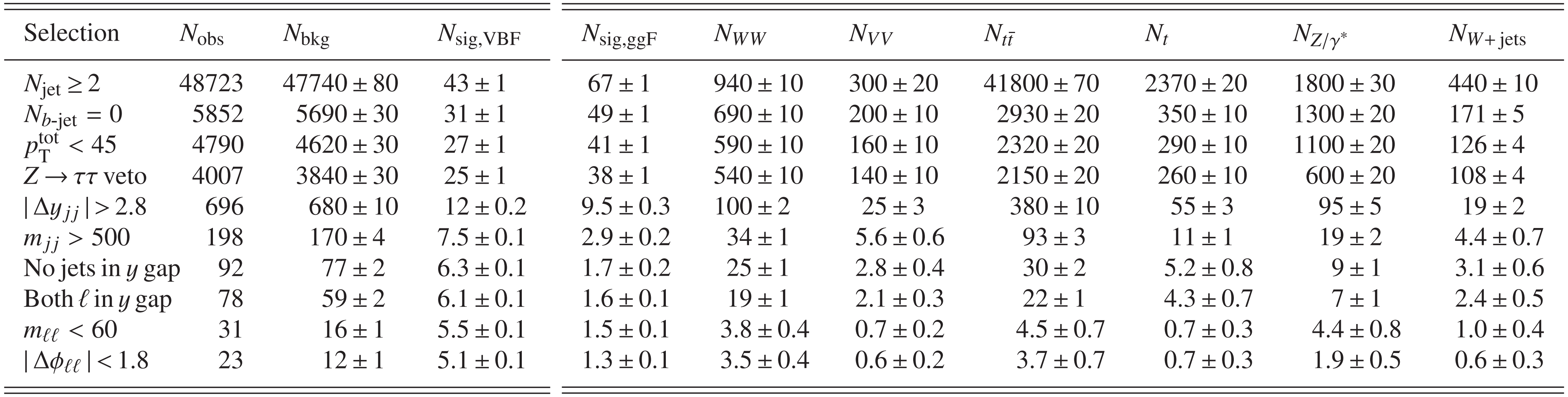
Number of events at the end of the cutflow.



Complete Event Yields

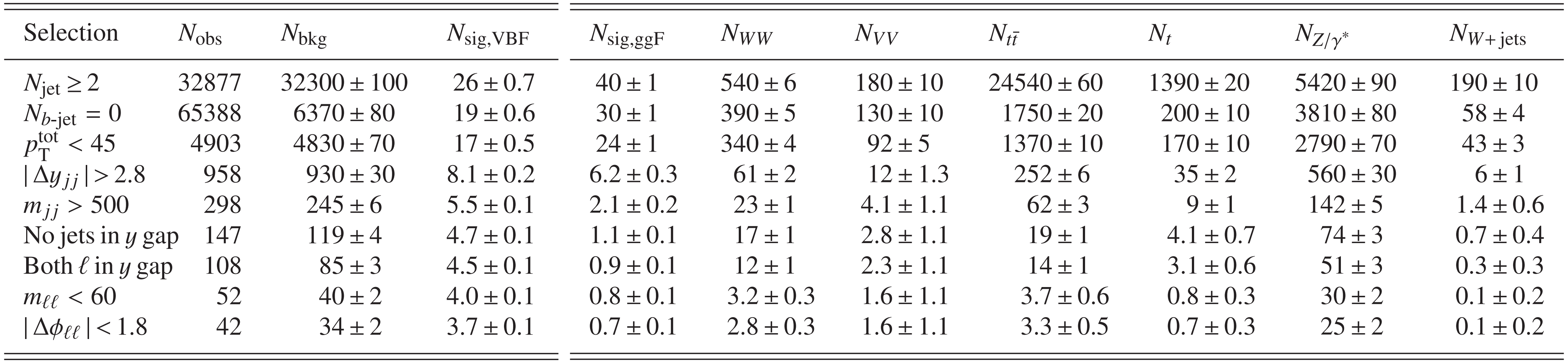
For VBF HWW:

(a) $e\mu + \mu e$ channel



Selection	N_{obs}	N_{bkg}	$N_{\text{sig,VBF}}$	$N_{\text{sig,ggF}}$	N_{WW}	N_{VV}	$N_{t\bar{t}}$	N_t	N_{Z/γ^*}	$N_{W+\text{jets}}$
$N_{\text{jet}} \geq 2$	48723	47740 ± 80	43 ± 1	67 ± 1	940 ± 10	300 ± 20	41800 ± 70	2370 ± 20	1800 ± 30	440 ± 10
$N_{b\text{-jet}} = 0$	5852	5690 ± 30	31 ± 1	49 ± 1	690 ± 10	200 ± 10	2930 ± 20	350 ± 10	1300 ± 20	171 ± 5
$p_{\text{T}}^{\text{tot}} < 45$	4790	4620 ± 30	27 ± 1	41 ± 1	590 ± 10	160 ± 10	2320 ± 20	290 ± 10	1100 ± 20	126 ± 4
$Z \to \tau\tau$ veto	4007	3840 ± 30	25 ± 1	38 ± 1	540 ± 10	140 ± 10	2150 ± 20	260 ± 10	600 ± 20	108 ± 4
$\lvert\Delta y_{jj}\rvert > 2.8$	696	680 ± 10	12 ± 0.2	9.5 ± 0.3	100 ± 2	25 ± 3	380 ± 10	55 ± 3	95 ± 5	19 ± 2
$m_{jj} > 500$	198	170 ± 4	7.5 ± 0.1	2.9 ± 0.2	34 ± 1	5.6 ± 0.6	93 ± 3	11 ± 1	19 ± 2	4.4 ± 0.7
No jets in y gap	92	77 ± 2	6.3 ± 0.1	1.7 ± 0.2	25 ± 1	2.8 ± 0.4	30 ± 2	5.2 ± 0.8	9 ± 1	3.1 ± 0.6
Both ℓ in y gap	78	59 ± 2	6.1 ± 0.1	1.6 ± 0.1	19 ± 1	2.1 ± 0.3	22 ± 1	4.3 ± 0.7	7 ± 1	2.4 ± 0.5
$m_{\ell\ell} < 60$	31	16 ± 1	5.5 ± 0.1	1.5 ± 0.1	3.8 ± 0.4	0.7 ± 0.2	4.5 ± 0.7	0.7 ± 0.3	4.4 ± 0.8	1.0 ± 0.4
$\lvert\Delta\phi_{\ell\ell}\rvert < 1.8$	23	12 ± 1	5.1 ± 0.1	1.3 ± 0.1	3.5 ± 0.4	0.6 ± 0.2	3.7 ± 0.7	0.7 ± 0.3	1.9 ± 0.5	0.6 ± 0.3

(b) $ee + \mu\mu$ channel



Selection	N_{obs}	N_{bkg}	$N_{\text{sig,VBF}}$	$N_{\text{sig,ggF}}$	N_{WW}	N_{VV}	$N_{t\bar{t}}$	N_t	N_{Z/γ^*}	$N_{W+\text{jets}}$
$N_{\text{jet}} \geq 2$	32877	32300 ± 100	26 ± 0.7	40 ± 1	540 ± 6	180 ± 10	24540 ± 60	1390 ± 20	5420 ± 90	190 ± 10
$N_{b\text{-jet}} = 0$	65388	6370 ± 80	19 ± 0.6	30 ± 1	390 ± 5	130 ± 10	1750 ± 20	200 ± 10	3810 ± 80	58 ± 4
$p_{\text{T}}^{\text{tot}} < 45$	4903	4830 ± 70	17 ± 0.5	24 ± 1	340 ± 4	92 ± 5	1370 ± 10	170 ± 10	2790 ± 70	43 ± 3
$\lvert\Delta y_{jj}\rvert > 2.8$	958	930 ± 30	8.1 ± 0.2	6.2 ± 0.3	61 ± 2	12 ± 1.3	252 ± 6	35 ± 2	560 ± 30	6 ± 1
$m_{jj} > 500$	298	245 ± 6	5.5 ± 0.1	2.1 ± 0.2	23 ± 1	4.1 ± 1.1	62 ± 3	9 ± 1	142 ± 5	1.4 ± 0.6
No jets in y gap	147	119 ± 4	4.7 ± 0.1	1.1 ± 0.1	17 ± 1	2.8 ± 1.1	19 ± 1	4.1 ± 0.7	74 ± 3	0.7 ± 0.4
Both ℓ in y gap	108	85 ± 3	4.5 ± 0.1	0.9 ± 0.1	12 ± 1	2.3 ± 1.1	14 ± 1	3.1 ± 0.6	51 ± 3	0.3 ± 0.3
$m_{\ell\ell} < 60$	52	40 ± 2	4.0 ± 0.1	0.8 ± 0.1	3.2 ± 0.3	1.6 ± 1.1	3.7 ± 0.6	0.8 ± 0.3	30 ± 2	0.1 ± 0.2
$\lvert\Delta\phi_{\ell\ell}\rvert < 1.8$	42	34 ± 2	3.7 ± 0.1	0.7 ± 0.1	2.8 ± 0.3	1.6 ± 1.1	3.3 ± 0.5	0.7 ± 0.3	25 ± 2	0.1 ± 0.2

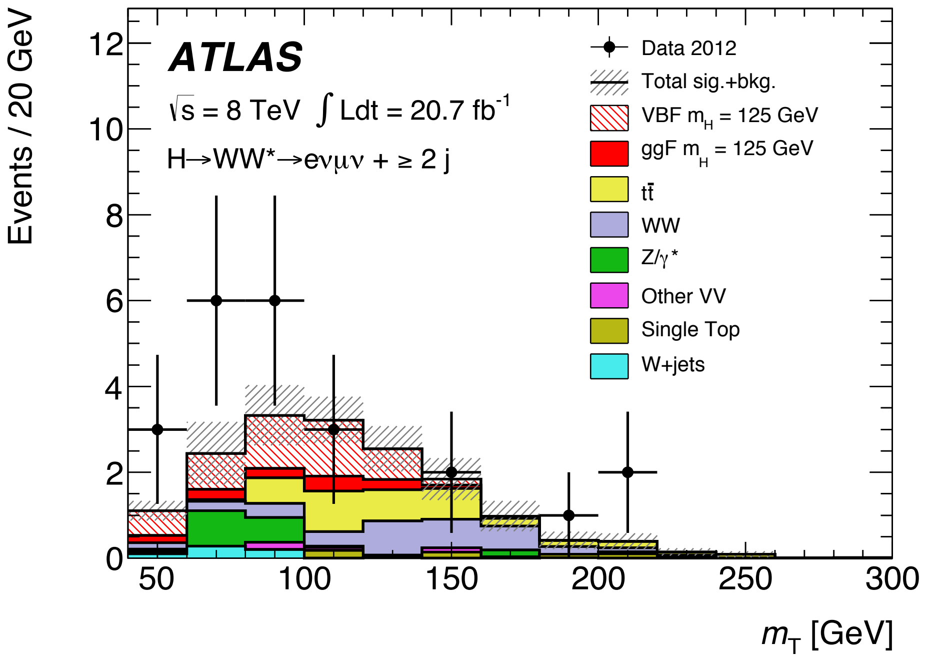
Units for m, p in GeV.

Only statistical uncertainties.

ATLAS-CONF-2013-030

Signal Strength and Systematic Uncertainties for VBF HWW

Fitting m_T to extract signal strength:



Final event yields:

Observed	Signal	Total Bkg
55	10.9±1.4	36 ± 4

The best-fit measured signal strength for m_H=125GeV for 7 and 8TeV data:

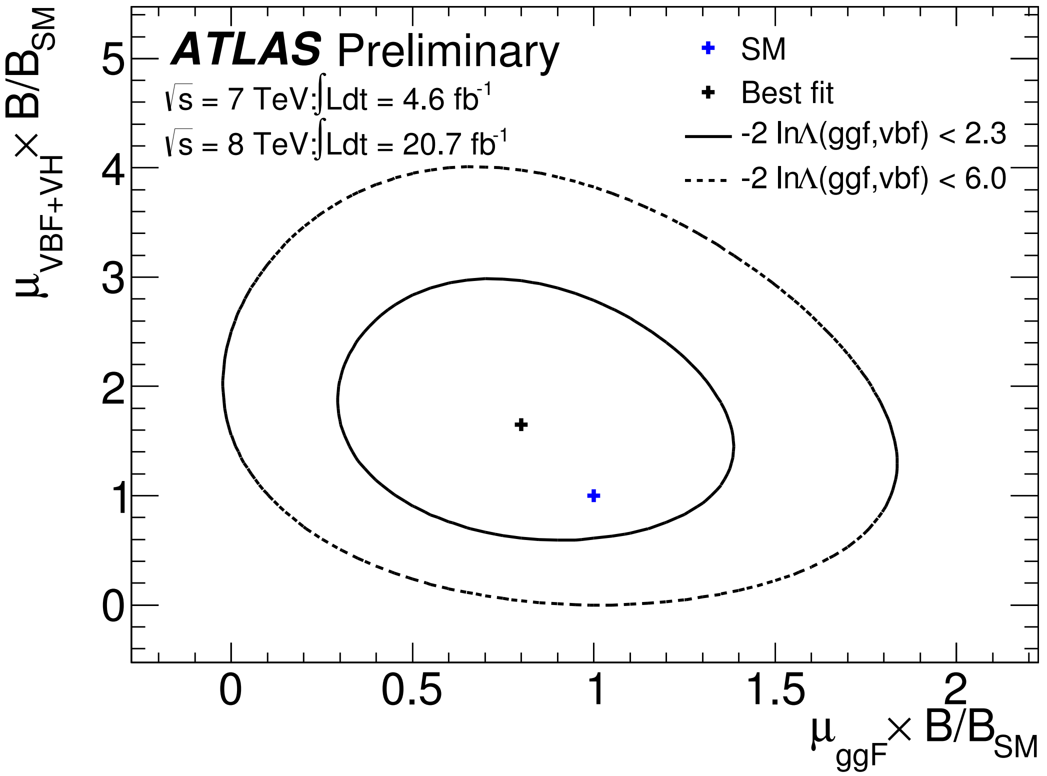
$$\boxed{\mu_{obs}\text{(VBF)}=1.7\pm\underline{0.7\text{(stat)}}\pm 0.4\text{(syst)}}$$

Top 10 leading bkg. relative uncertainties on the total bkg. yields in the combined eμ+μe VBF channel

source	relative uncertainty on bkg. [%]
b-tagging efficiency	7.88
QCD scale VV+2jets	7.61
top bkg. theory syst.	5.43
JER	3.9
top normalisation using CR	3.85
QCD scale VV accept	3.62
JES 2012 modelling	2.72
JES FlavComp HWW	2.32
QCD scale, ggF+2jets H Xsec	2.32
QCD scale ggF+3jets H Xsec	2.3

ATLAS-CONF-2013-030

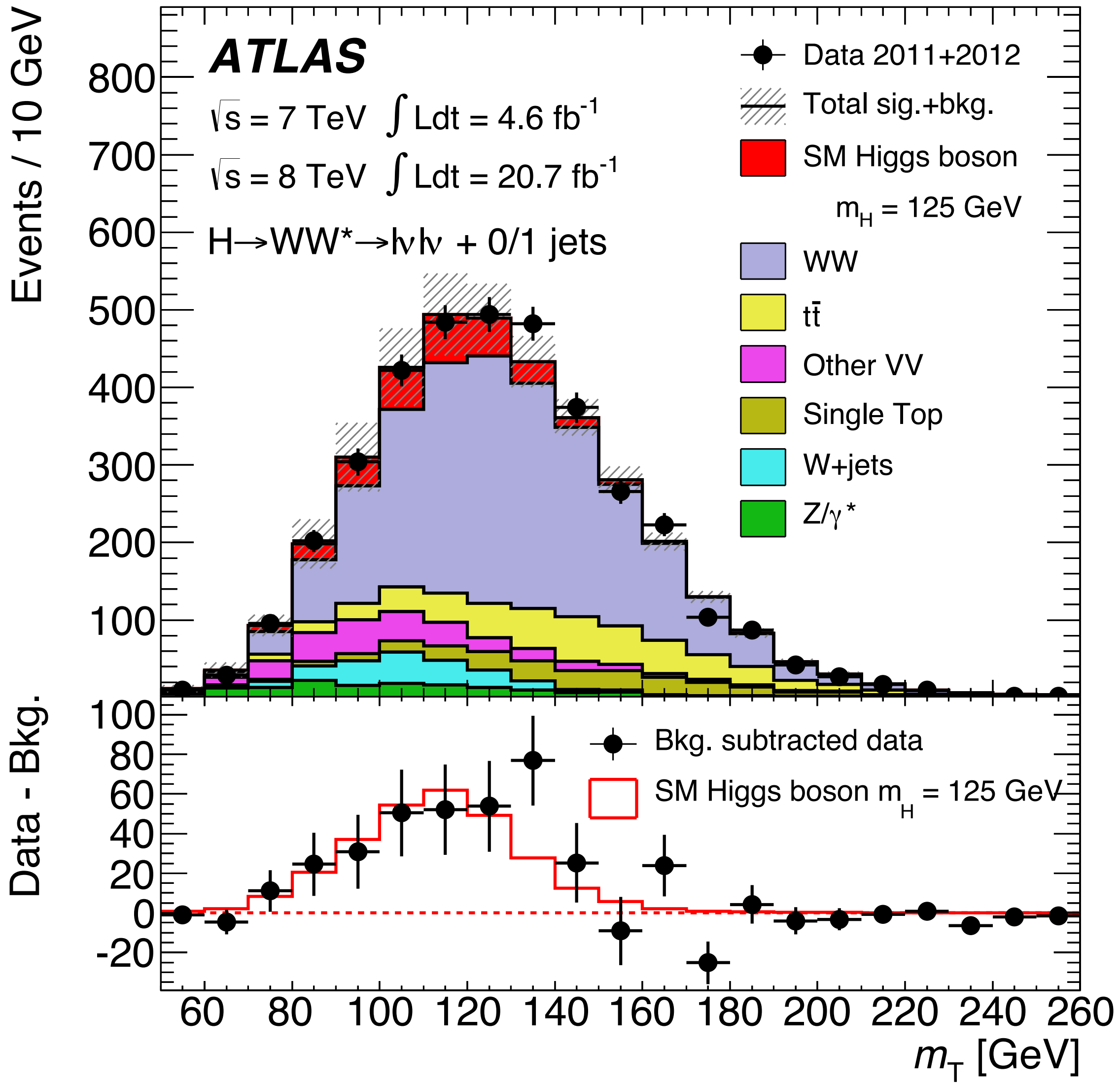
H→WW→ lvlv signal strength



$$\mu_{\text{obs, VBF}} = 1.66 \pm 0.67\,(\text{stat.}) \pm 0.42\,(\text{syst.})$$

$$\mu_{\text{obs, ggF}} = 0.82 \pm 0.24\,(\text{stat.}) \pm 0.28\,(\text{syst.})$$

0+1 jet fitting m_T distribution



Combined 8+7TeV for ggF + VBF

$$\mu_{\text{obs}} = 1.01 \pm 0.21\,(\text{stat.}) \pm 0.19\,(\text{theo. syst.}) \pm 0.12\,(\text{expt. syst.}) \pm 0.04\,(\text{lumi.})$$

signal significance: **3.8(obs)** 3.8(exp)

Full Systematic Uncertainties and Event Yields

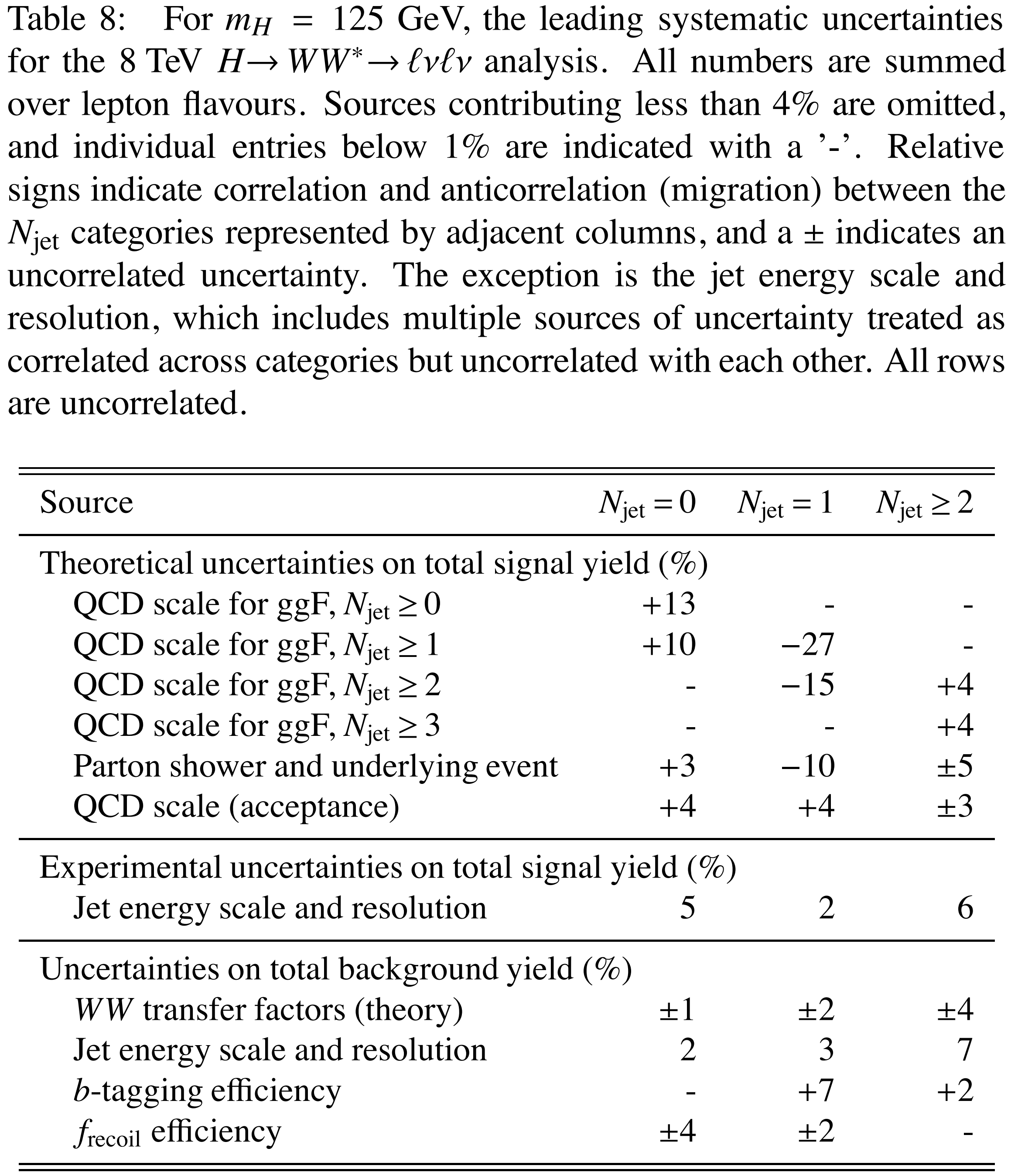


Table 8: For $m_H = 125$ GeV, the leading systematic uncertainties for the 8 TeV $H \to WW^* \to \ell\nu\ell\nu$ analysis. All numbers are summed over lepton flavours. Sources contributing less than 4% are omitted, and individual entries below 1% are indicated with a '-'. Relative signs indicate correlation and anticorrelation (migration) between the N_{jet} categories represented by adjacent columns, and a ± indicates an uncorrelated uncertainty. The exception is the jet energy scale and resolution, which includes multiple sources of uncertainty treated as correlated across categories but uncorrelated with each other. All rows are uncorrelated.

Source	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} \geq 2$
Theoretical uncertainties on total signal yield (%)			
QCD scale for ggF, $N_{\mathrm{jet}} \geq 0$	+13	-	-
QCD scale for ggF, $N_{\mathrm{jet}} \geq 1$	+10	−27	-
QCD scale for ggF, $N_{\mathrm{jet}} \geq 2$	-	−15	+4
QCD scale for ggF, $N_{\mathrm{jet}} \geq 3$	-	-	+4
Parton shower and underlying event	+3	−10	±5
QCD scale (acceptance)	+4	+4	±3
Experimental uncertainties on total signal yield (%)			
Jet energy scale and resolution	5	2	6
Uncertainties on total background yield (%)			
WW transfer factors (theory)	±1	±2	±4
Jet energy scale and resolution	2	3	7
b-tagging efficiency	-	+7	+2
f_{recoil} efficiency	±4	±2	-

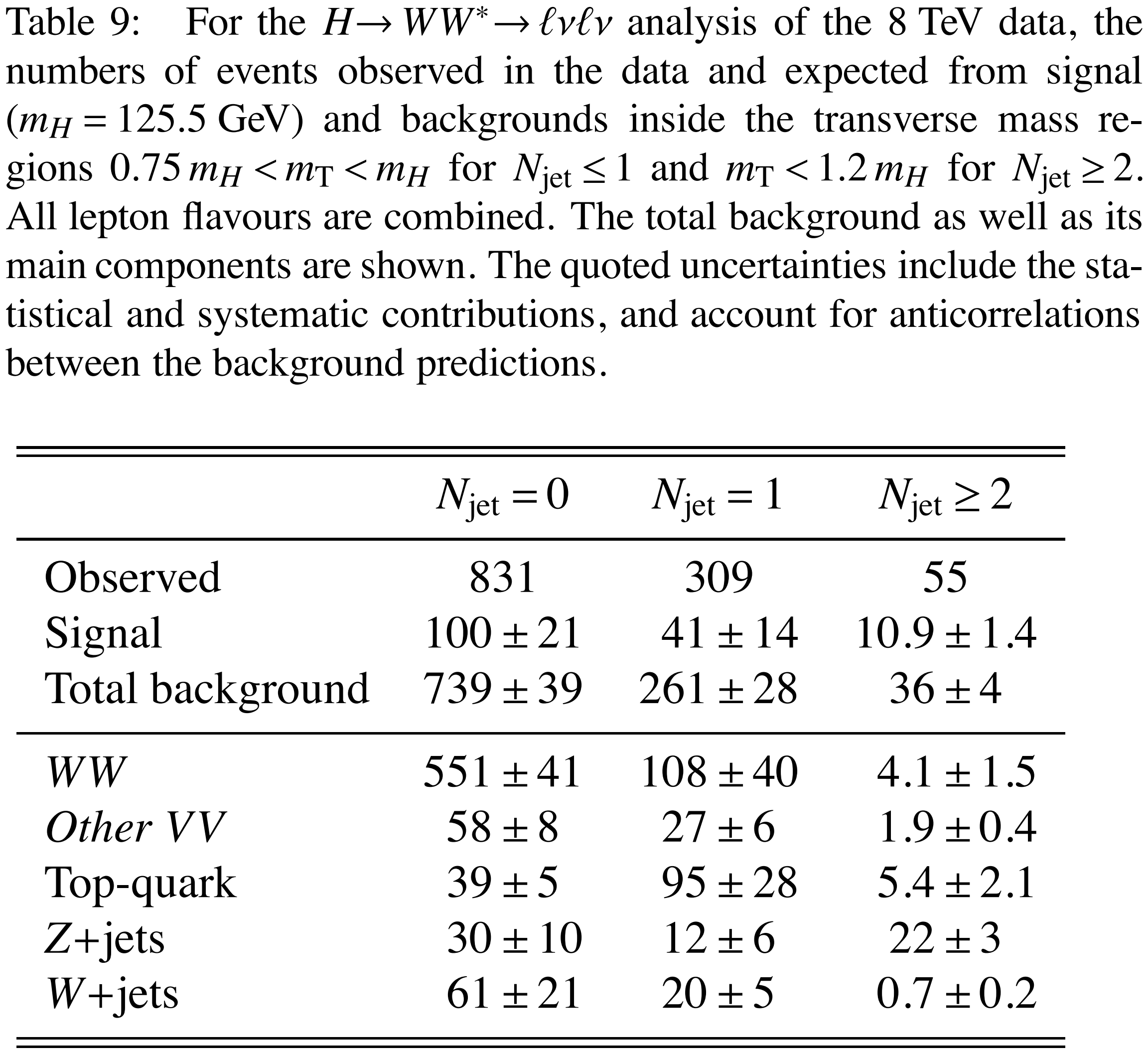


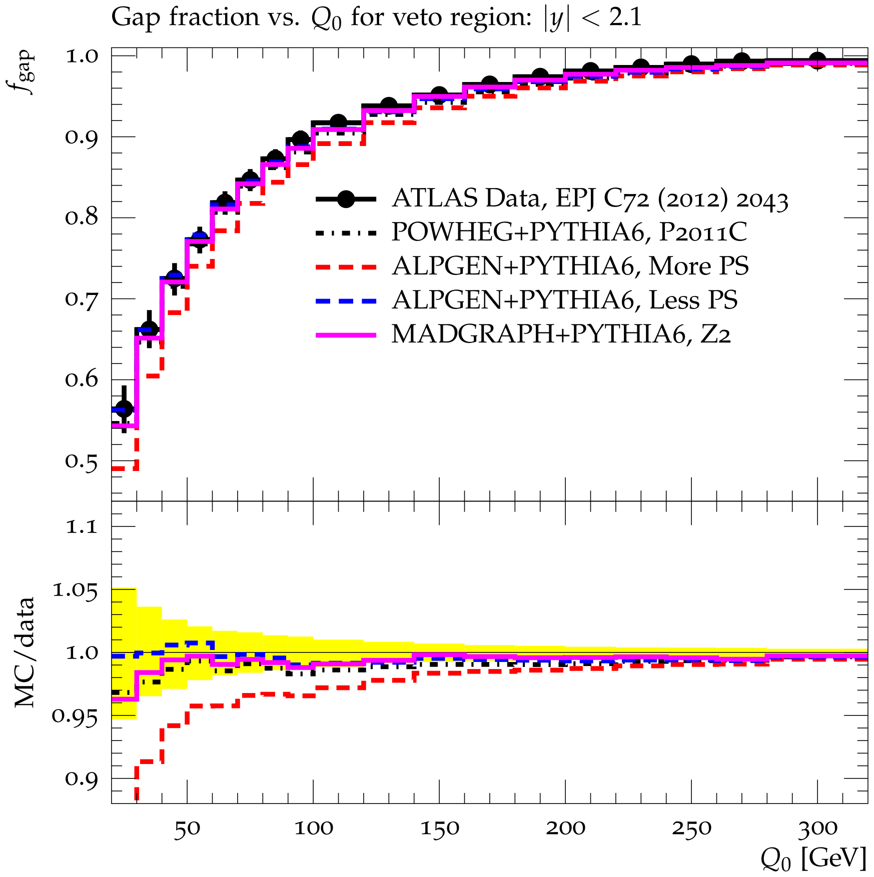
Table 9: For the $H \to WW^* \to \ell\nu\ell\nu$ analysis of the 8 TeV data, the numbers of events observed in the data and expected from signal ($m_H = 125.5$ GeV) and backgrounds inside the transverse mass regions $0.75\, m_H < m_{\mathrm{T}} < m_H$ for $N_{\mathrm{jet}} \leq 1$ and $m_{\mathrm{T}} < 1.2\, m_H$ for $N_{\mathrm{jet}} \geq 2$. All lepton flavours are combined. The total background as well as its main components are shown. The quoted uncertainties include the statistical and systematic contributions, and account for anticorrelations between the background predictions.

	$N_{\mathrm{jet}} = 0$	$N_{\mathrm{jet}} = 1$	$N_{\mathrm{jet}} \geq 2$
Observed	831	309	55
Signal	100 ± 21	41 ± 14	10.9 ± 1.4
Total background	739 ± 39	261 ± 28	36 ± 4
WW	551 ± 41	108 ± 40	4.1 ± 1.5
Other VV	58 ± 8	27 ± 6	1.9 ± 0.4
Top-quark	39 ± 5	95 ± 28	5.4 ± 2.1
Z+jets	30 ± 10	12 ± 6	22 ± 3
W+jets	61 ± 21	20 ± 5	0.7 ± 0.2

ATLAS-CONF-2013-030

Multi-leg vs. NLO vs. LO

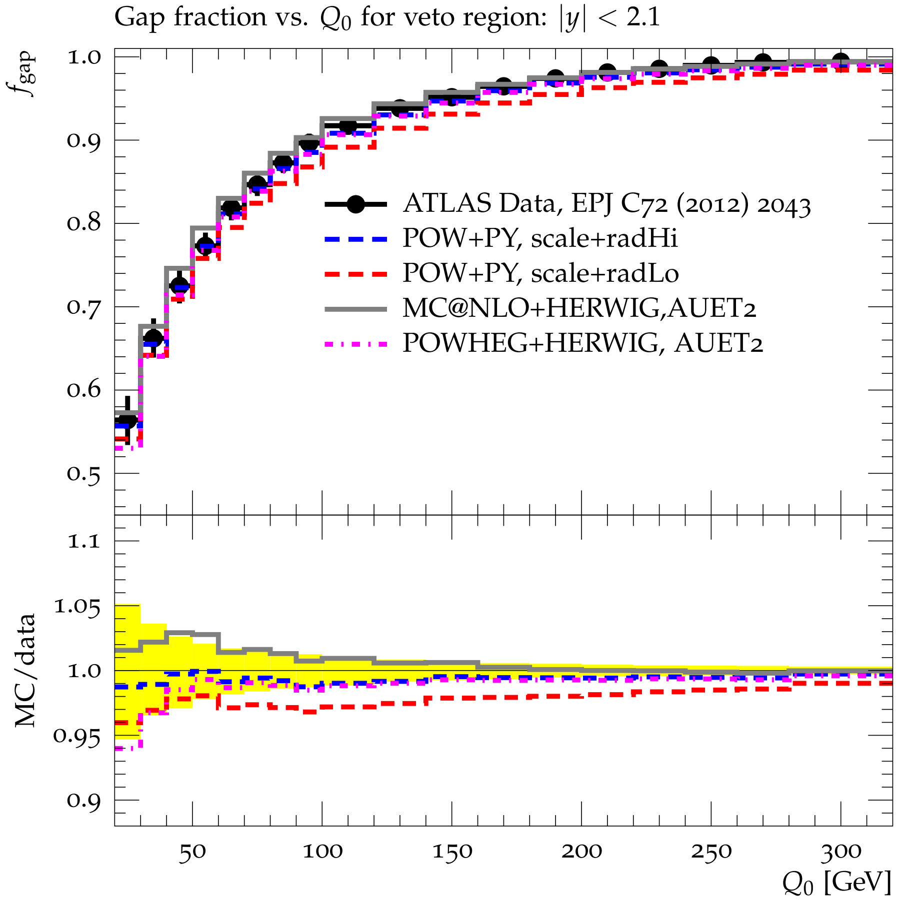
multi-leg



Alpgen+Pythia6 `as_down,radLO` and `as_up,radHi`(Perugia 2011 tune) sample simultaneously varying:

- Pythia6 parameters (`PARP(72)` and `PARP(64)`) for I/FSR
- `ktfac` parameters for renormalisation scale used at ME.

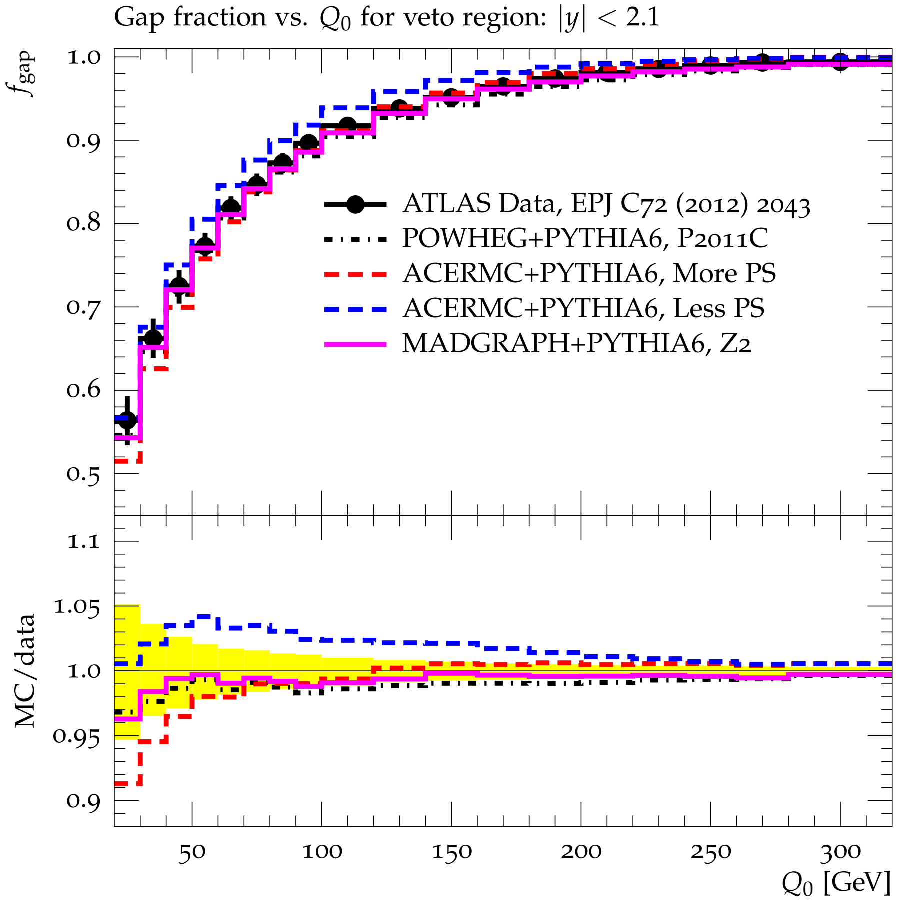
NLO



Powheg+Pythia6 `radLO` and `radHi` Perugia 2012 tune. Simultaneously varied

- renomalisation and factorisation scales in the presence of damping (of high p_T radiation in Powheg)

LO



AcerMC+Pythia6 (Perugia 2011) varied Pythia I/FSR parameters (`PARP(72)`, `PARP(61)` and `PARP(64)`) to get MorePS and LessPS samples.

Pythia parameters controlling ISR and FSR

- PARP(67): controls high-pt ISR branchings phase-space;
  ISR branchings with pTevol > m_dip/2 * PARP(67) are
  power suppressed by a factor (m_dip/(2pTevol))**2

- PARP(64): multiplicative factor of the mom. scale^2 in running
  alpha_s used in ISR

- PARP(72): multiplicative factor of the lam_QCD in
  running alpha_s used in FSR central param. setting is
  motivated by ATLAS FSR QCD jet shapes,
  variations correspond to *1/2 and *1.5 central value

- PARJ(82): FSR low-pt cutoff

ISR/FSR - a truth level study

- For this study MC@NLO(v4.10 link) + Herwig (v6.521, Fortran link) were used.

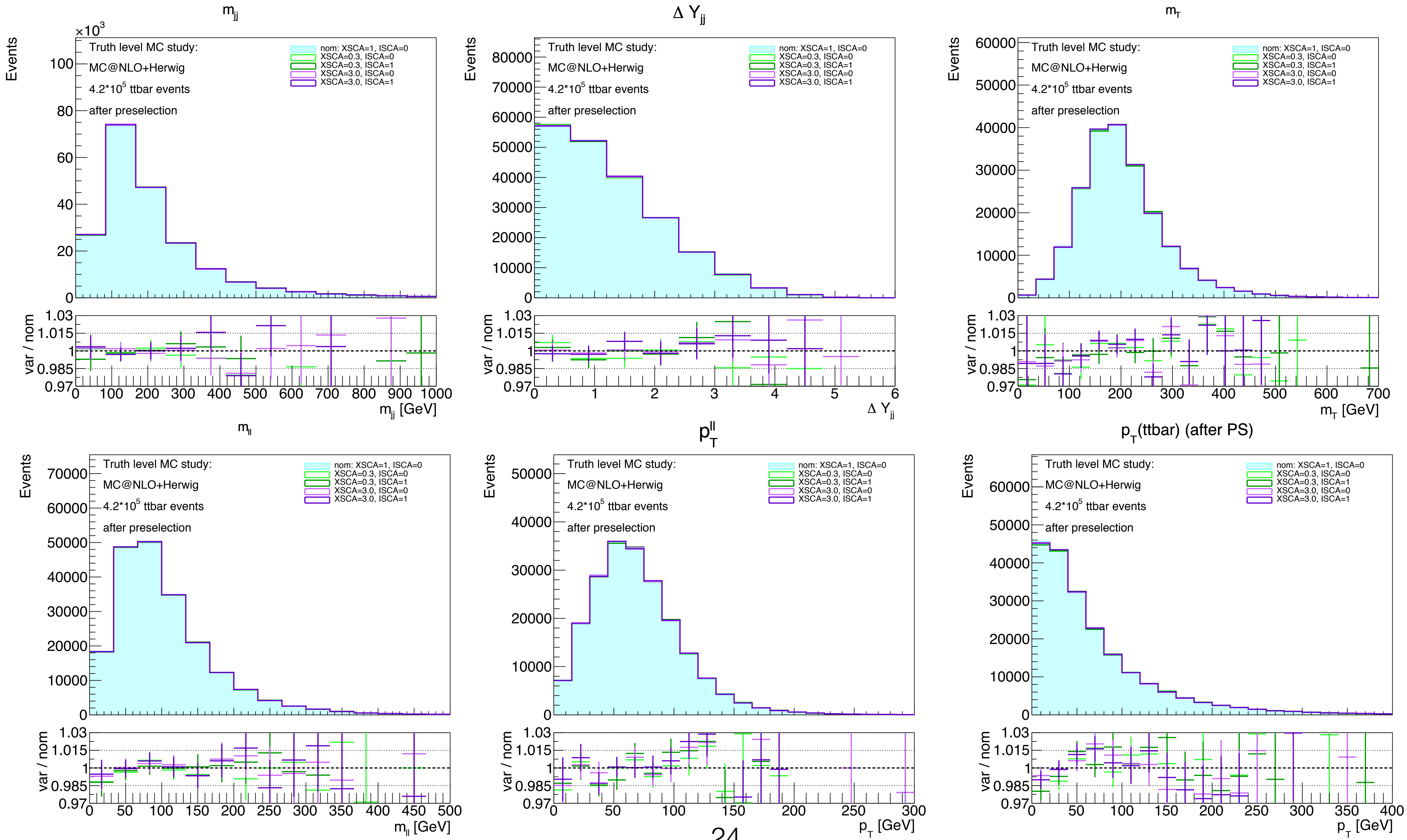
- Amount of ISR/FSR in Herwig is varied by changing the veto on emissions in the parton showers with $p_T >$ SCALUP.

- SCALUP - a parameter used by Herwig in its showering stage affecting both ISR and FSR. It is varied by changing 2 additional parameters in Herwig:
  - ➡ XSCA $\in [0.3, 3.0]$ → SCALUP = SCALUP * XSCA
  - ➡ ISCA = 0 → SCALUP = ECM - 2PTR,
ISCA = 1 → SCALUP = ECM

  ECM - subprocess center of mass energy
  PTR - p_T of hard emission in the collider frame

- Variations on SCALUP parameter only possible for Herwig Fortran v6.521 and ttbar pair production.

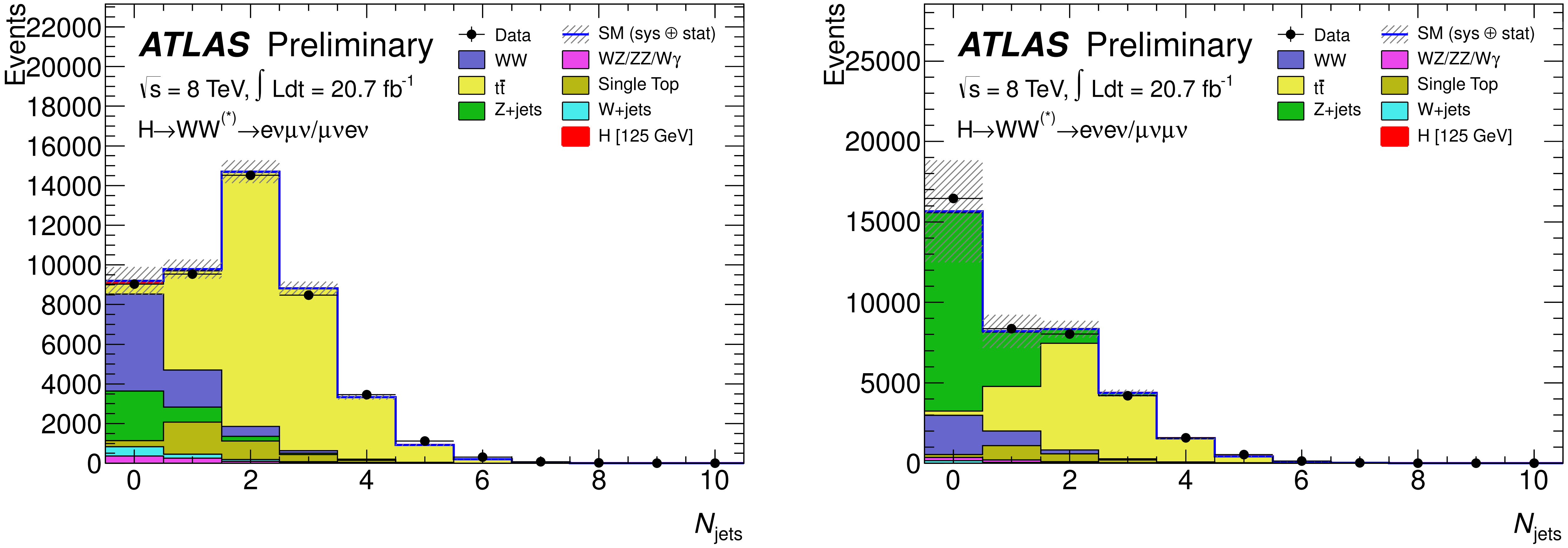
More I/FSR Results



Additional plots

Njet

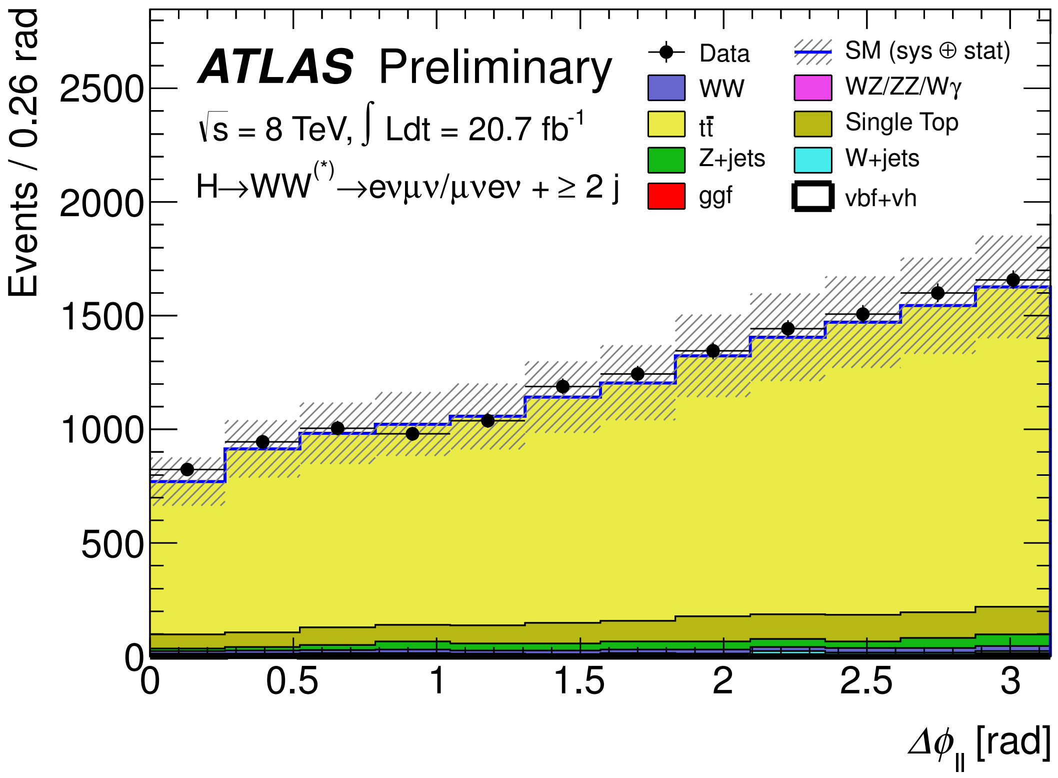
after preselection

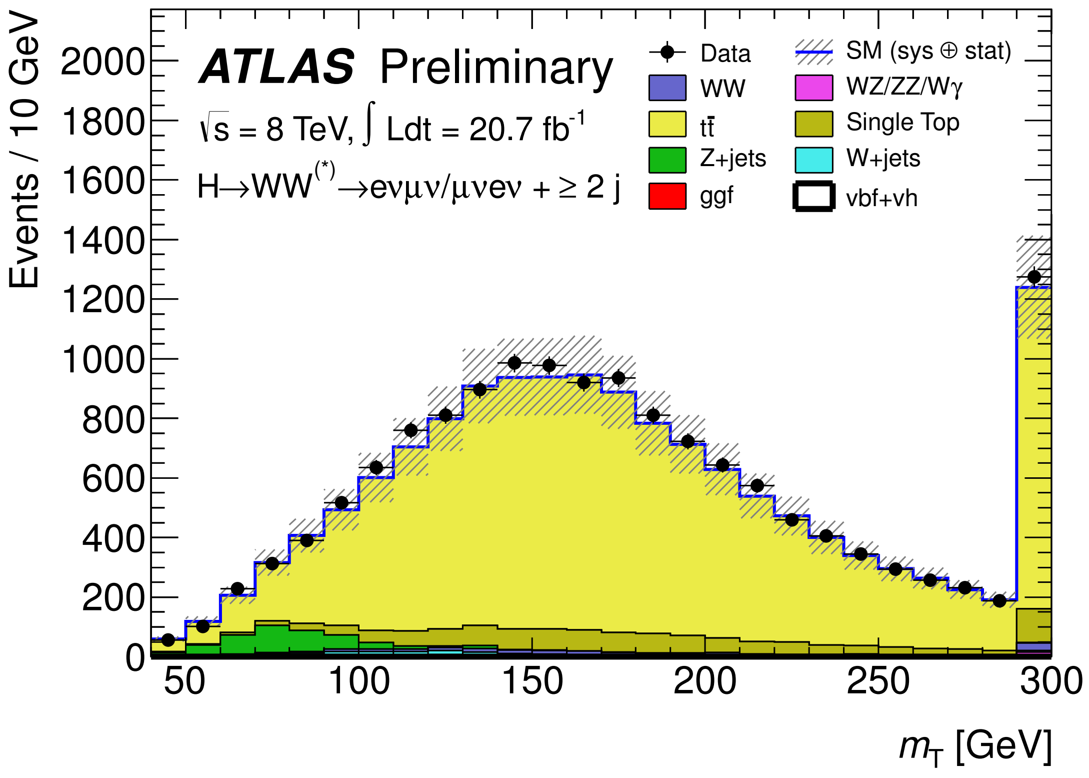


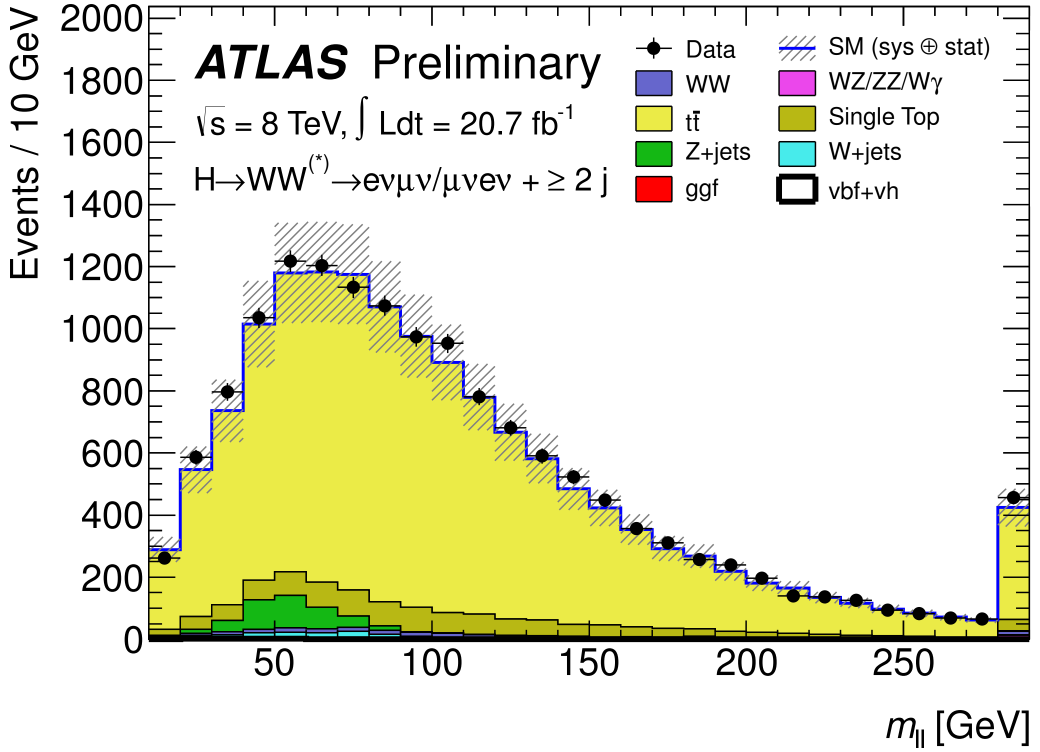
(a) N_{jet} distribution for $e\mu + \mu e$

(b) N_{jet} distribution for $ee + \mu\mu$

H-decay topological variables in topCR





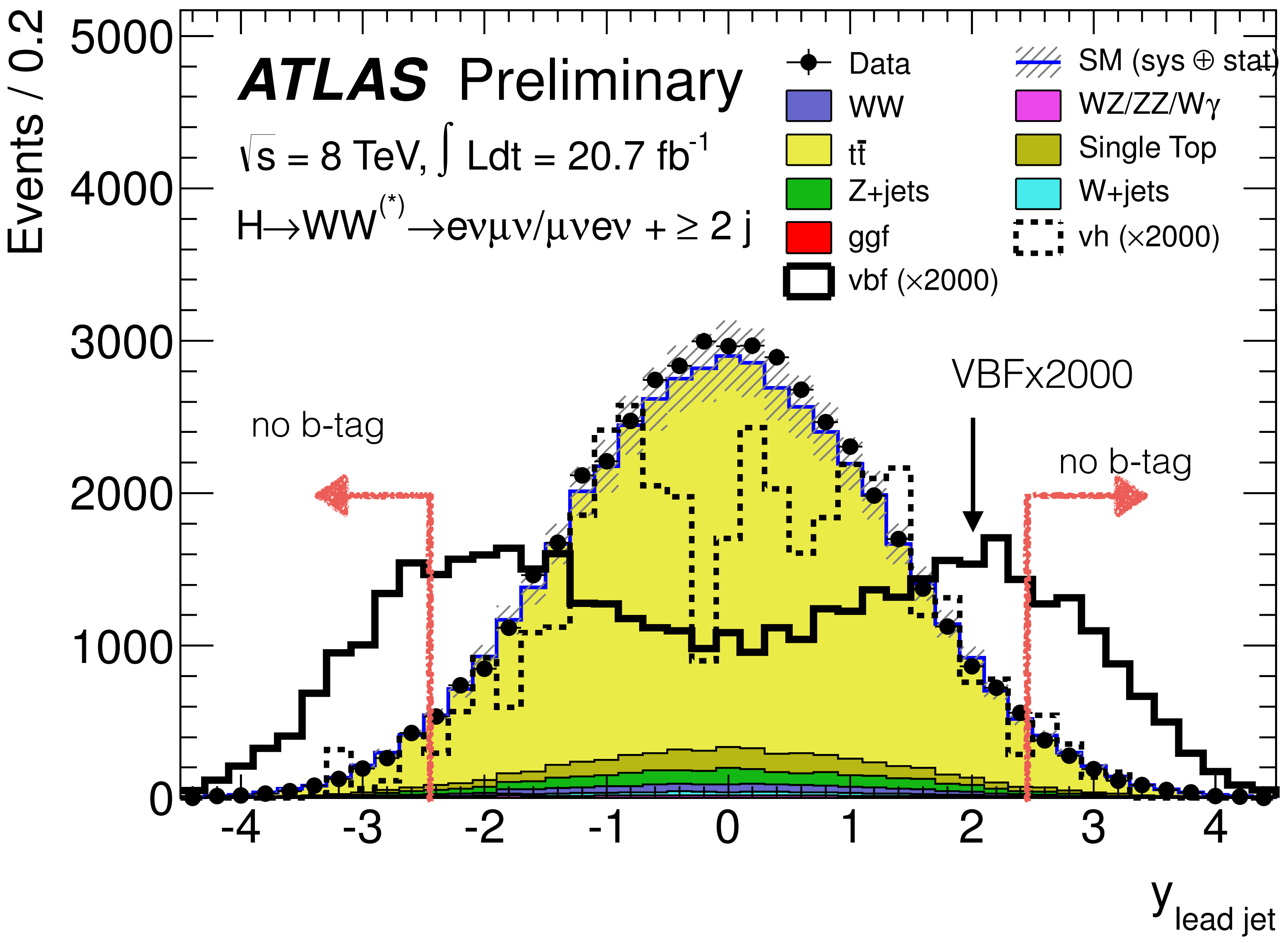


The $|\Delta\phi_{ll}|$, m_T and m_{ll} distributions after the $p^{tot} < 45$ GeV cut in the top CR, defined by the requirement of one and only one b tagged jet . p^{tot} is defined as the total transverse momentum of all leptons, jets and missing E_T passing the selection. The shaded area represents the uncertainty on the signal and background yields from statistical, experimental, and theoretical sources

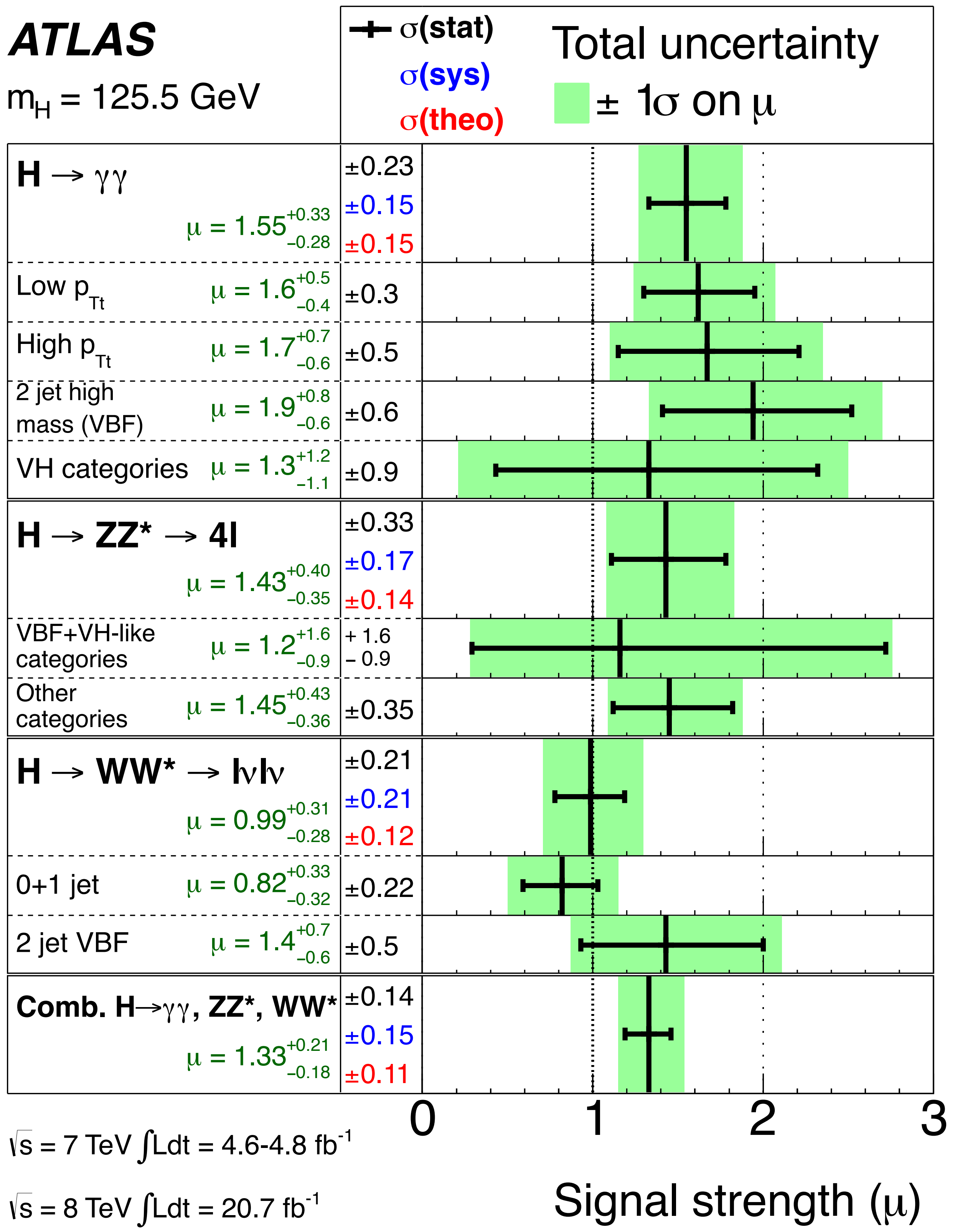
ATLAS-CONF-2013-030

Rapidity of the leading jet

after Njets cut



Signal Strength combination



ATLAS Higgs Coupling

Figure 6: The measured production strengths for a Higgs boson of mass *mH* =125.5 GeV, normalised to the SM expectations, for diboson final states and their combination. Results are also given for the main categories of each analysis (described in Sections 4.2, 5.2 and 6.2). The best-fit values are shown by the solid vertical lines, with the total ±1σ uncertainty indicated by the shaded band, and the statistical un- certainty by the superimposed horizontal error bars. The numbers in the second column specify the contributions of the (symmetrised) sta- tistical uncertainty (top), the total (experimental and theoretical) sys- tematic uncertainty (middle), and the theory uncertainty (bottom) on the signal cross section (from QCD scale, PDF, and branching ratios) alone; for the individual categories only the statistical uncertainty is given.