EWK $(Z\to\nu\nu)\gamma$ process: combination of SM and EXOT analyses

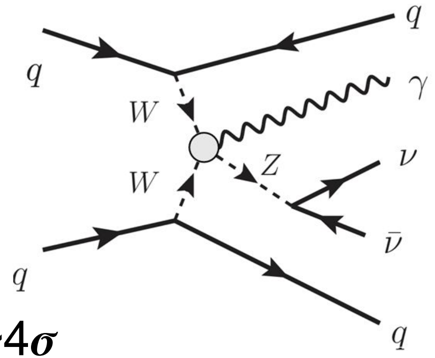
Douglas Schaefer,

Evgeny Soldatov

on behalf of SM Zγ and Exotics VBF H+γ teams

3 March 2021

Introduction and outline



- Our talks on this topic at SM EWK meetings are 1, 2.
- EXOT analysis closure talk was given last week, link.
  5.2(5.1)σ of observed(expected) significance obtained.
- SM Z($\nu\nu$)γ VBS analysis (close to SM approval) shows ~4σ of the expected significance.
- Since the signal extraction regions are independent, **we are working on the combination.**

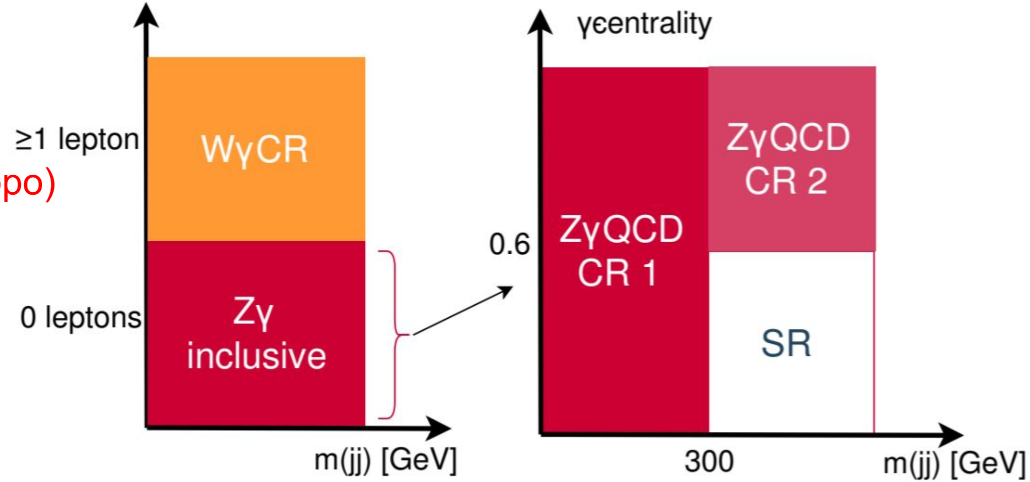
- So the aim of this talk is to discuss the preliminary combination results and possible issues.

Main selections and backgrounds (SM analysis)

Preselections:

Selections	Cut Value
E_T^{miss}	> 120 GeV
E_T^{γ}	> 150 GeV
Number of tight isolated photons	$N_{\gamma} = 1$
Number of jets	$N_{jets} \geq 2$ (AntiKt4Topo)
Lepton veto	$N_e = 0, N_{\mu} = 0$
E_T^{miss} significance	> 12
$\lvert\Delta\phi(\gamma, \vec{p}_T^{\,miss})\rvert$	> 0.4
$\lvert\Delta\phi(j_1, \vec{p}_T^{\,miss})\rvert$	> 0.3
$\lvert\Delta\phi(j_2, \vec{p}_T^{\,miss})\rvert$	> 0.3
$p_T^{SoftTerm}$	< 16 GeV

Additional selections to construct SR and 3 CRs:



Data-driven backgrounds:

- e→γ (mainly from W[eν]),
- jet→γ (mainly from Z[νν]+jet),
- fake E_T[miss] (mainly from γ+jet)

together constitute ~13% of the total yield (see next slide for the methodologies)

Backgrounds estimated from the fit in CRs:

- QCD Z(νν)γ (using Zγ CRs)
- W(lν)γ, ttγ (using Wγ CR)

Directly from MC:

- Z(ll)γ (<1% of the total event yield)

Main selections (EXOT analysis)

Period	Trigger
All 2015	HLT_xe70_mht
2016, Runs ≤ 302872	HLT_xe90_mht_L1XE50
2016, Runs > 302872	HLT_xe110_mht_L1XE50
2015–2016	HLT_noalg_J400
2017 Runs	HLT_xe110_pufit_L1XE55
2018 Runs	HLT_xe110_xe70_L1XE50

- MET triggered events
- PFlow jets p_T>25 GeV
- Tight photons with FCLoose Iso
  - 15 GeV < p_T<110 GeV
- Loose muons, LooseLH electron veto

γ -centrality rev. CR

$W\gamma$ - 1μ or 1e CR

SR - 0Lepton

Cut	SR
N_{photon}	=1
$N_{baseline\text{-}lepton}$	=0
N_{jet}	=2,3
$N_{b\text{-}jet}$ (77% WP)	<2
$p_T(j_1)$	> 60 GeV
$p_T(j_2)$	> 50 GeV
MET	> 150 GeV
MET_CST_jet	> 120 GeV
$\eta(j_1) \times \eta(j_2)$	<0
$\Delta\eta(j_1,j_2)$	> 3.0
$\Delta\varphi(j_1,j_2)$	< 2.5
$\Delta\varphi(j_{1,2,3},\text{MET})$	> 1
abs. $\Delta\varphi$(MET,photon)	> 1.8
Photon Centrality	> 0.4
Third jet centrality	< 0.7
$M(j_1,j_2)$	> 250 GeV
Photon p_T	< 110 GeV
abs(Photon Pointing)	< 250 mm

lead two jets fjvt>0.4

Table 17: Lepton selections

Cut	*inv.*	$W(\to e\nu)\gamma$+jets	$W(\to \mu\nu)\gamma$+jets
Lepton Flavours	0	e^-/e^+	μ^-/μ^+
"Veto" muons and electrons	0	1	1
"Signal" muons and electrons	0	1	1
$p_T(\ell_1)$	–	>30 GeV	>30 GeV
$p_T(\ell_2)$	–	–	–
$\|M(\ell\ell) - M_Z\|$	–	–	–
E_T^{miss}(with leptons)	> 80 GeV	–	–

Fit setup overview (SM analysis)

	Sample	Norm. coef	Systematic uncertainties	
			Designated	Common
MC estimated	Zγjj EWK	mu(ZγEWK), POI	Theory, interference	
	Zγjj QCD	mu(ZγQCD)	Theory	
	Wγ EWK		Theory	Experimental
	Wγ QCD	mu(Wγ)	Theory	
	ttγ		Theory	
	Z(ll)+γ		Theory	
Data driven	e -> γ		Data-driven flat	
	j+γ		Data-driven flat	
	j->γ		Data-driven flat	

Templates:

- mjj in the control regions
- BDT classifier response in the signal region

To account for limited MC statistics there is also an NP for every bin with MC stat error > 5%

	SR	ZγQCD CR 1	ZγQCD CR 2	Wγ
μ(ZγEWK)	v	v	v	
μ(ZγQCD)	v	v	v	
μ(Wγ)	v	v	v	v

The fit is performed in **3 steps**:

1. Fit MC to data in CRs with μ(ZγQCD) and μ(Wγ) as parameters of interest (POI).
2. Use fitted parameter values (norm.coef + NPs) to create Asimov pseudodata.
3. Fit MC to Asimov pseudodata in all regions with μ(ZγEWK) as POI and obtain the estimated median discovery significance.

Regions definition and fit setup (EXOT analysis)

- Control Regions:
  - 1 lepton regions binned in m_{jj}
    - Used to extract the normalization of the Wγ backgrounds
  - photon centrality reversed region binned in m_{jj}
    - Used to extract normalization of the Zγ strong background
- Signal Regions:
  - 0 lepton regions
  - Bins used in the maximum LH fit to extract signal in **4 m_{jj} bins**

	SR	$Z\gamma$ QCD rev. γ Centrality CR 2	$W\gamma$ CR
$\mu_{Z\gamma \mathrm{EWK}}$	✓	✓	
$\mu_{Z\gamma \mathrm{QCD}}$	✓	✓	
$\mu_{W\gamma}$	✓	✓	✓

Table 81: Table of regions where the normalization coefficients are used to calculate the likelihood function.

So the setups were well harmonized between analyses.

Theoretical systematic uncertainties

	Modelling		Scale	Alt. PDF set	NNPDF + αs
	Underlying Event + Parton showering	Sherpa vs MG			
Zγjj EWK	v		v	v	v
Zγjj QCD		v		v	v
Wγ EWK	v		v	v	v
Wγ QCD		v		v	v
ttγ	v		v		v
Z(ll)+γ			v	v	v

1 NP per combination of sample and systematic, e. g. α(ttγ UE+PS), α (WγQCD alt PDF set), except for **scale.**

Scale:
4 NPs per sample for:

- Wγ CR
- ZγQCD CR1
- ZγQCD CR2
- SR

Modelling

Zγ QCD - Difference between MadGraph and Sherpa samples was taken as modelling, Sherpa used as a central value.
Wγ QCD - no alternative sample available, made by with the ZgQCD modelling systematic relative uncertainties (see backup for more info)
ttγ - only fastsim alternative sample available, made by comparing fastsim nominal sample with fastsim alternative sample

Zγjj QCD and **Wγ QCD scale** uncertainties are omitted since they only affect the normalization and tend to double the designated normalization coefficients in the likelihood model.

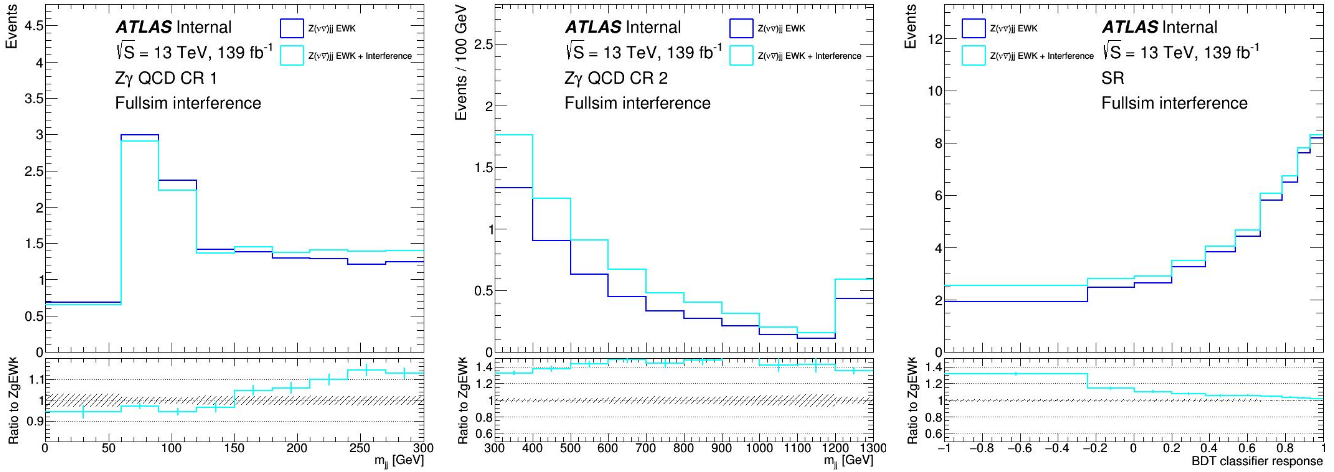
Zγjj EWK and QCD interference as systematic uncertainty

The interference between Zγjj EWK and QCD processes is treated as systematic uncertainty for the Zγjj EWK in 3 regions:

- ZγQCD CR 1
- ZγQCD CR 2
- SR

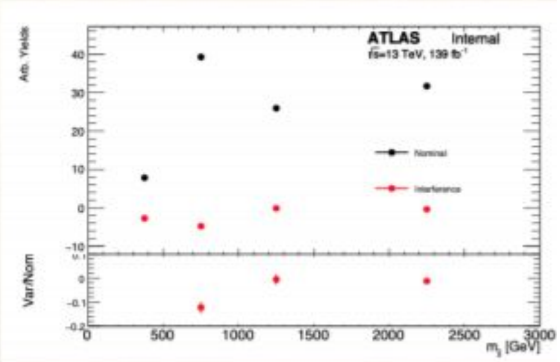
1 NP for all 3 regions

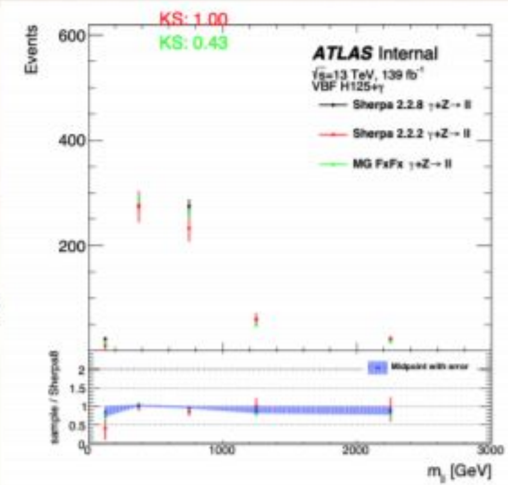
+σ	Zγjj EWK + interference
Nominal	Zγjj EWK
-σ	Zγjj EWK + interference

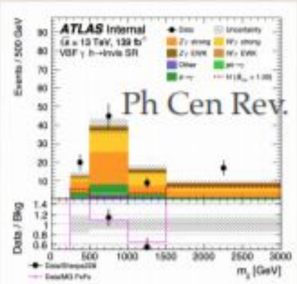


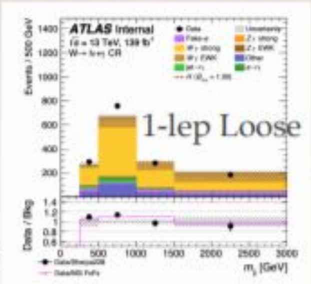
Theory uncertainties (EXOT analysis)

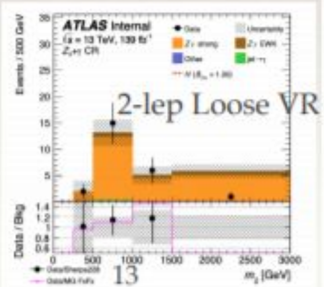
- Strong-EWK interference
  - Computed the LO strong-EWK Zg interference at truth level
  - Applied as an uncertainty on the EWK Zg as a function mjj
    - the impact is very small
- MG vs Sherpa Zg strong comparison
  - Compared for Z(ll)g at reco level
  - Full difference between MG & Sherpa taken as a systematic uncertainty with Sherpa as the central value
  - Same relative shape uncertainty is applied to Wg strong
  - Wg strong and Zg strong MG Comparison syst is not correlated because of differences in agreement with data in loosened CRs











Theory uncertainties correlation model (EXOT analysis)

Theory Sys.	SR (m_{jj} > 500 GeV)	SR (250 < m_{jj} < 500 GeV)	$Z\gamma$ QCD rev. γ Centrality CR 2	$W\gamma$ CR
pdf_Wg_strong	✓	✓	✓	✓
pdf_Zg_strong	✓	✓	✓	✓
pdf_Zg_EWK	✓	✓	✓	✓
pdf_Wg_EWK	✓	✓	✓	✓
renofact_Wg_strong	✓	✓	✓	✓
renofact_Zg_strong	✓	✓	✓	$\oplus$
renofact_Zg_EWK	✓	$\oplus$	$\otimes$	$\odot$
renofact_Wg_EWK	✓	$\oplus$	$\otimes$	$\odot$
ps_Zg_EWK	✓	✓	✓	$\odot$
ps_Wg_EWK	✓	✓	✓	✓
interfer_Zg_EWK	✓	✓	✓	✓
MGComparison_Wg_strong	✓	✓	✓	✓
MGComparison_Zg_strong	✓	✓	✓	✓

Table 118: List of correlations for the theory uncertainties. The same symbol indicates that the uncertainty is correlated across the SRs and CRs. Each line is uncorrelated with the other lines.

Decorrelated to avoid pulls in the 1-lepton CR. Has almost no impact because the contamination is ~1% in that CR

- Largest impact from correlation models comes from the choice to not correlate MG vs Sherpa differences in the Wg and Zg strong

Central value	Exp. Significance
Model config in Table 118	5.1σ
Correlate MGComparison between $W\gamma$+jets and $Z\gamma$+jets strong	4.9σ
Correlate systematics across all bins (no CR decorrelation)	5.1σ
Correlate ps_Zg_EWK across all bins	5.1σ

Experimental systematic uncertainties

Type	Set
JES	30 NPs, CategoryReduction
JER	8 NPs, SimpleJER
MET	3 NPs
e/γ	2 NP, "1NP set"
Muon	10 NPs
e efficiency	3 NPs
γ efficiency	2 NPs
Trigger efficiency	1 NP
JVT efficiency	1 NP
PRW	1 NP
Luminosity	1 NP
Pile-up bkg	1 NP

1 NP per systematic type for every background **estimated from the MC**, e.g. α(JET_EffectiveNP_Detector1)

The systematics are prunned using the following rules.

- The normalization part is dropped if the total effect on the event yield is < 1%
- The shape part is dropped if there's no bins with effect on the event yield is > 1%

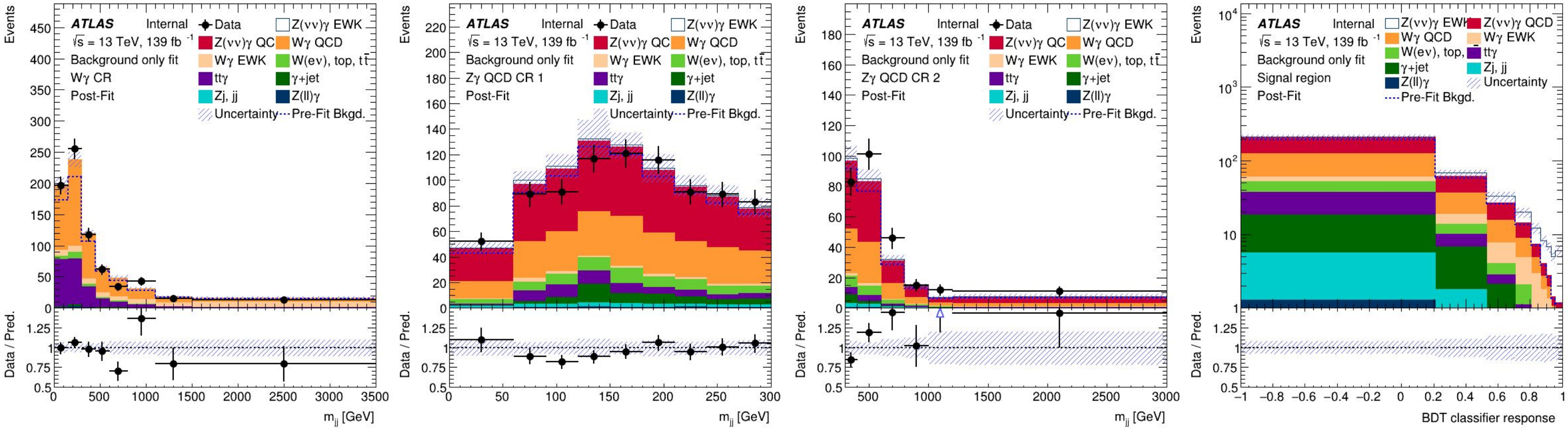
See backup for more info

Flat 1.9% systematic to account for pile-up background event yield

Every background with **data-driven estimations** also have an NP for their estimated systematic (see slide 5) which results in 3 more NPs

Fit results

Background only fit

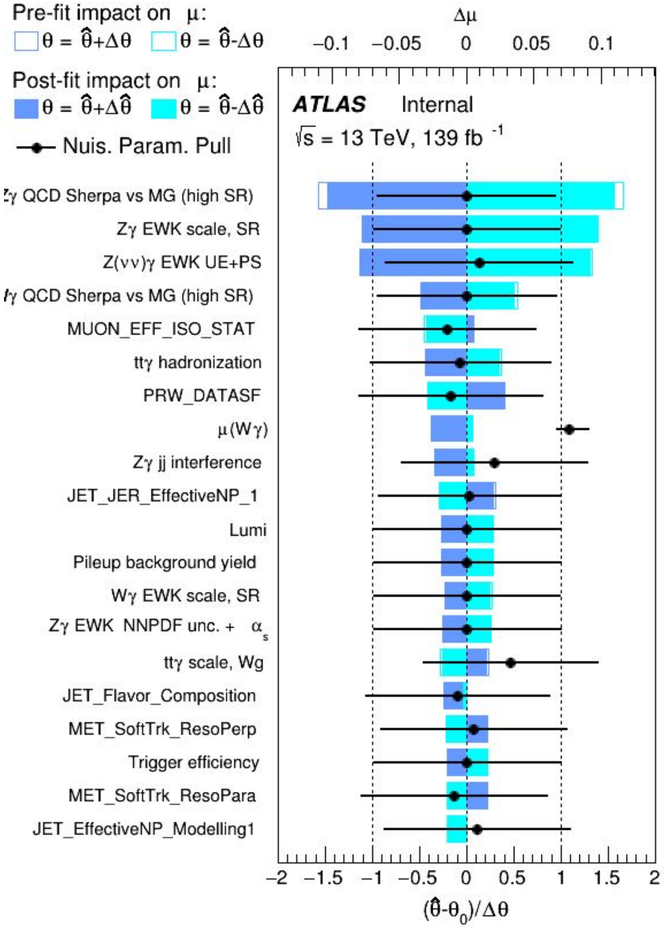


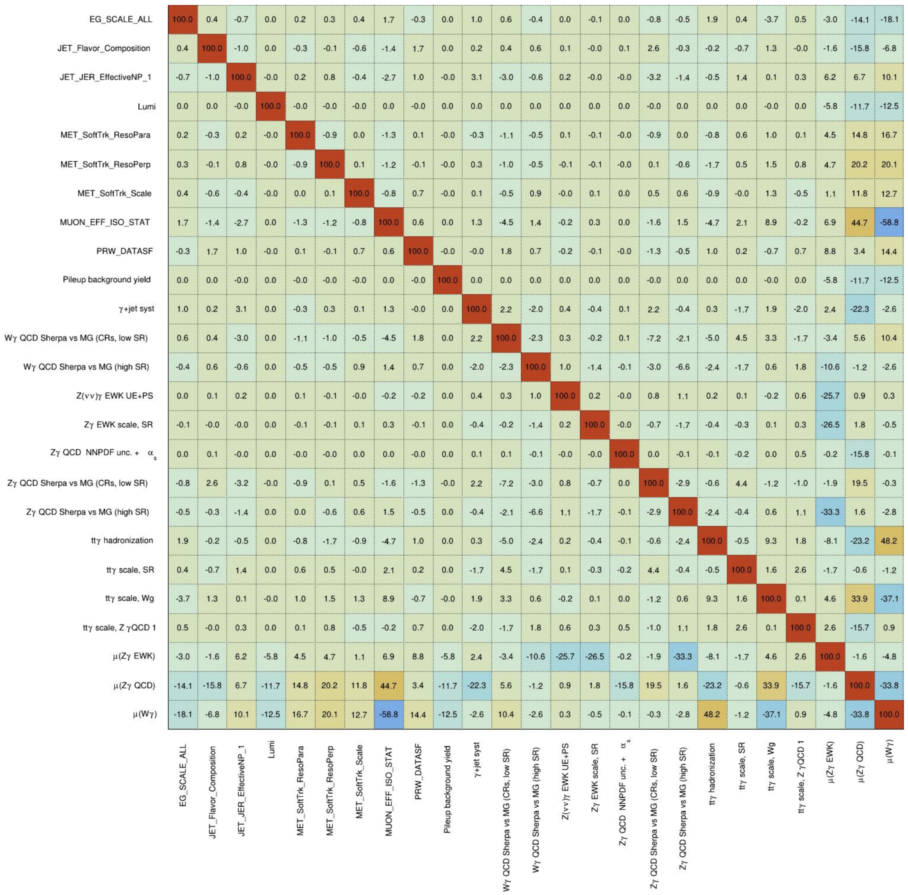
Asimov data fit

μ(ZγEWK)	$1.04^{+0.27}_{-0.25}$ (stat)$^{+0.24}_{-0.17}$ (syst)
μ(ZγQCD)	1.08 ± 0.08 (stat)$^{+0.16}_{-0.17}$ (syst)
μ(Wγ)	1.09 ± 0.04 (stat)$^{+0.20}_{-0.14}$ (syst)

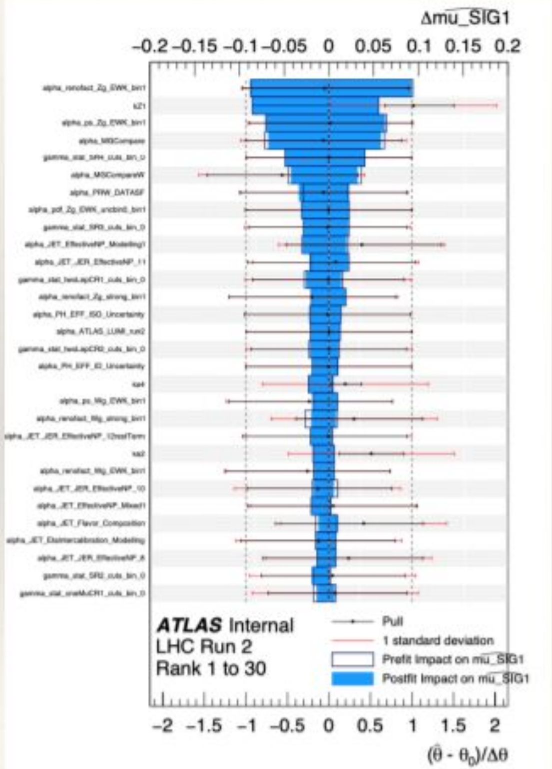
Expected median significance:
3.8 σ

Systematic uncertainties ranking and correlations





Fit results (EXOT analysis)



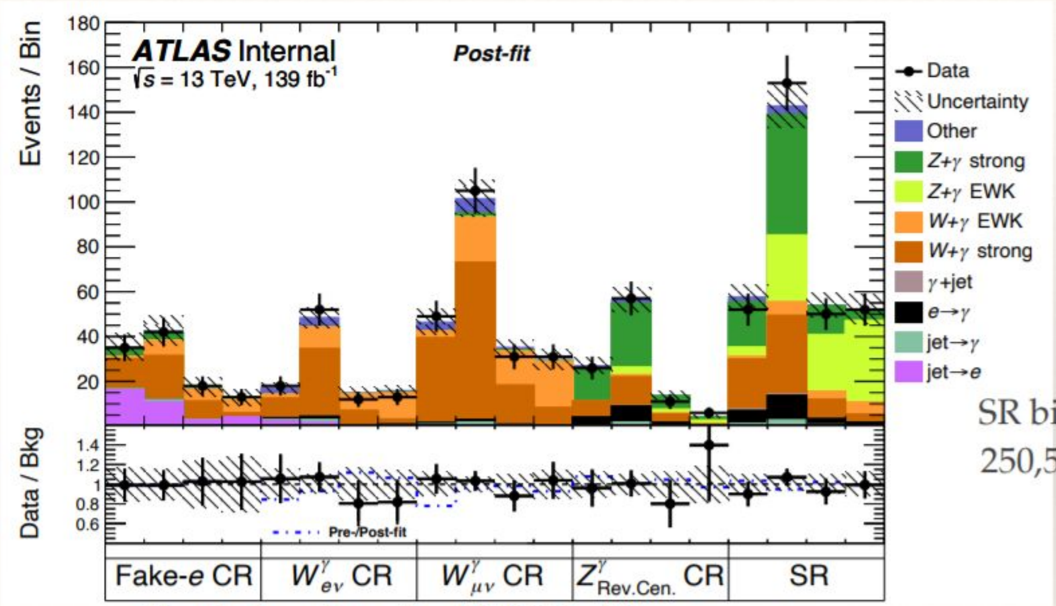
- W NF ($k_{W\gamma}$) 1.013±0.198
- Strong Z NF 1.024±0.412
- Fake e (k^{e}) 1.6±0.8
- EWK Z NF 1.03 ± 0.23

Source	1σ Uncertainty on $\mu_{Z\gamma_{EWK}}$ Unc.	% of Total
Jet Scale and Resolution	0.076	34
Photon	0.035	16
E_T^{miss}	0.023	10
Lepton	0.027	12
$V\gamma$+jets theory	0.067	30
Signal theory	0.054	24
Pileup	0.040	18
jet→ e, γ Bkg.	0.035	16
$W\gamma$+jets/$Z\gamma$+jets NF	0.073	32
MC stats.	0.063	28
Data stats.	0.157	70
Total	0.225	-

- Leading systematic uncertainties on EWK Zg
  - Zg EWK scale variations
  - Normalization of Zg strong
  - Zg EWK parton shower
  - MG vs Sherpa comparison for Zg strong
  - MC stats in SR bin 4

Observed(expected) significance: **5.2(5.1)σ**

PreFit vs PostFit



SR bins are in mjj: 250,500,1000,1500 GeV

Control regions for EWK Zg fit

- Pre- and Post-Fit results for the m_{jj} fit: showing the CR and SR bins
- Observed (expected) significance 5.2σ (5.1σ) for EWK Zgjj

Discussion

<u>Differences that we currently see:</u>

- **AntiKt4Topo vs PFlowJets jets.** Actually there are combinations of analyzes with different jet collections in ATLAS.
- **Simple JER vs Full JER.** The impact of JER systematics on the final result is small.
  - We are correlating alpha_JET_JER_EffectiveNP_{1..6}, alpha_JET_JER_DataVsMC_MC16 between analyses.
- **JET veto.** EXOT analysis has jet veto (only events with 2-3 jets are allowed), SM analysis doesn’t have jet veto.

If you see any other possible problems/issues - it would be very useful for us!

Combination

We are combining two workspaces using **workspaceCombiner** tool.

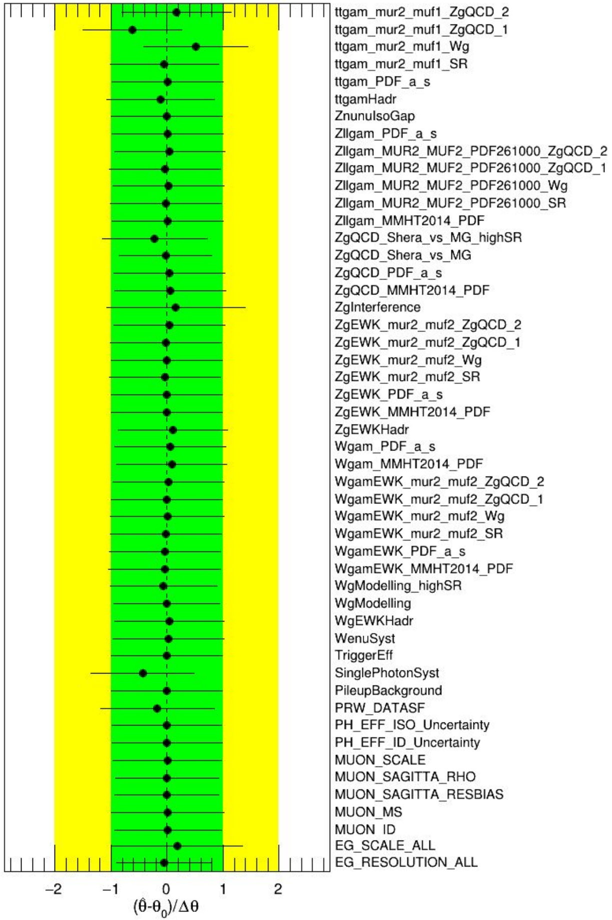
Same systematics in different workspaces are described by the same NPs.

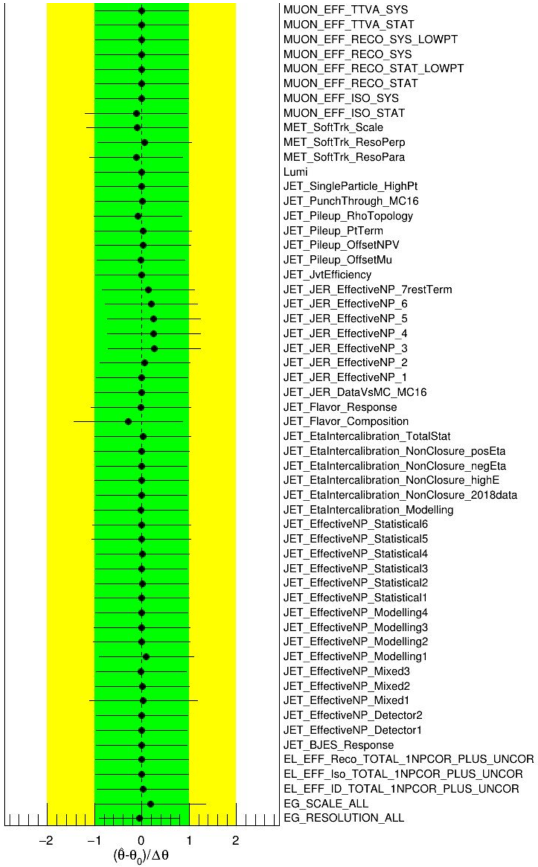
- Data-driven background NPs are not correlated (different methods):
  SinglePhotons, ZnunuIsoGap, WenuSyst (SM)
  JetFakePh*, GJet*, EFakePh* (EXOT)
- Also not correlated NPs:
  TriggerEff, PileUpBkg, ttgam_mur2_muf1_* (SM)
  xeSFTriggerWeight, renofact_* (EXOT)

Since the triggers are different, pile-up background is not significant in EXOT study.

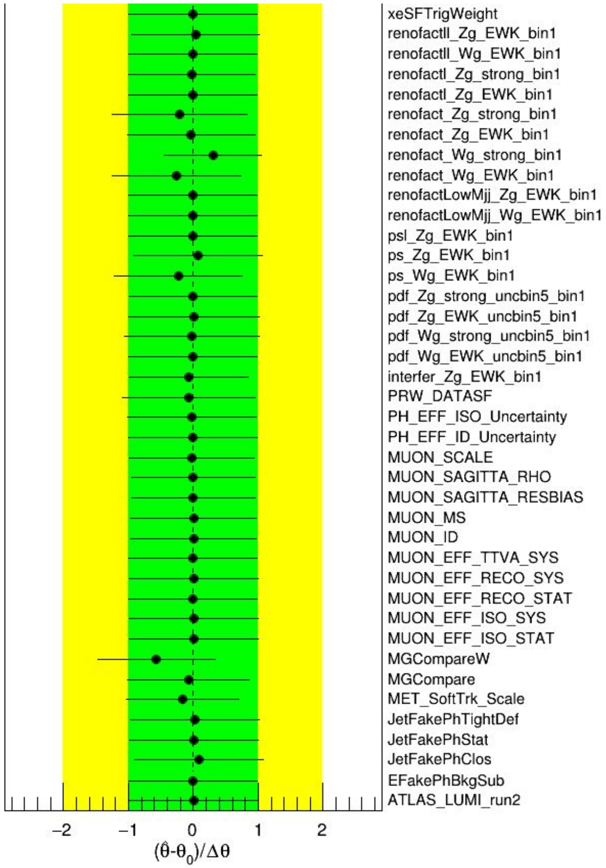
Scale uncertainties were removed in SM analysis, since they are flat and can be fully accounted by normalizations. More details in backup.

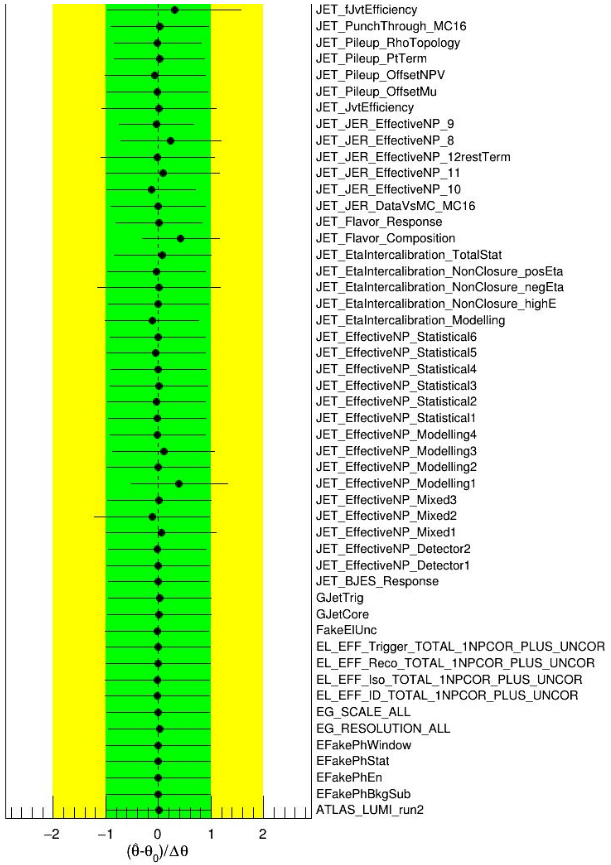
SM pulls plot for combination





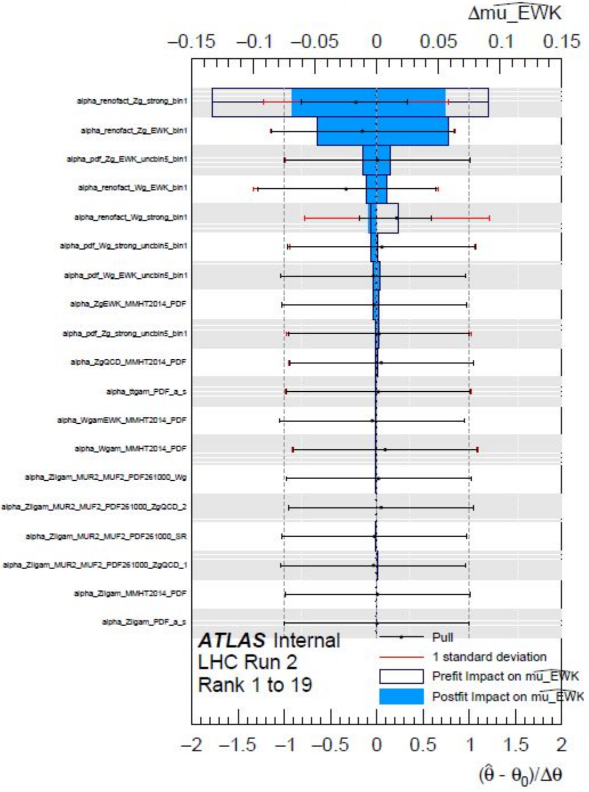
EXOT pulls plot for combination



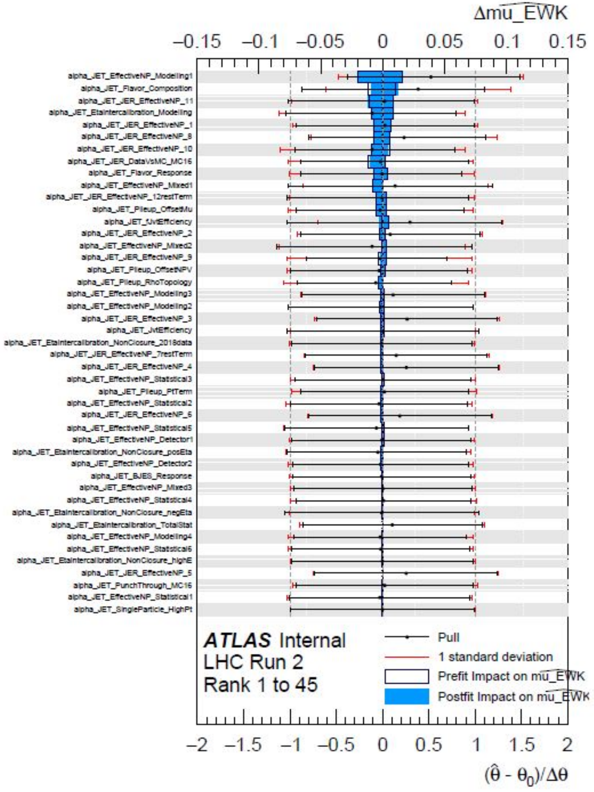


Combination

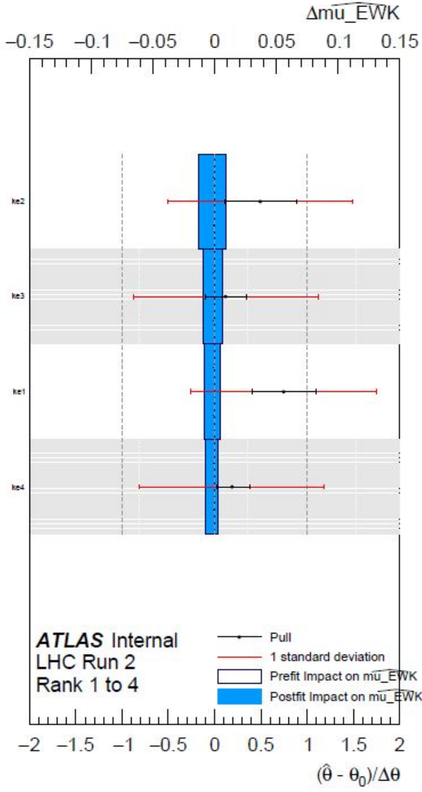
Theory



Jet

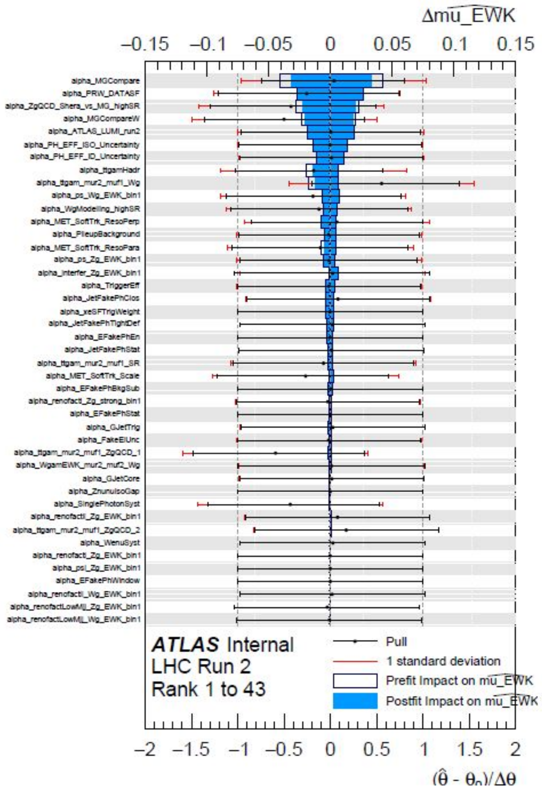


CR extrapolation

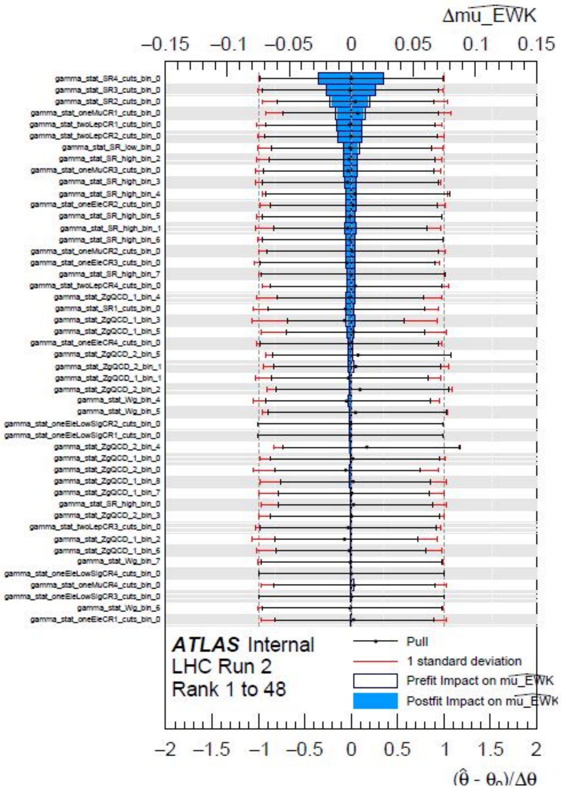


Combination

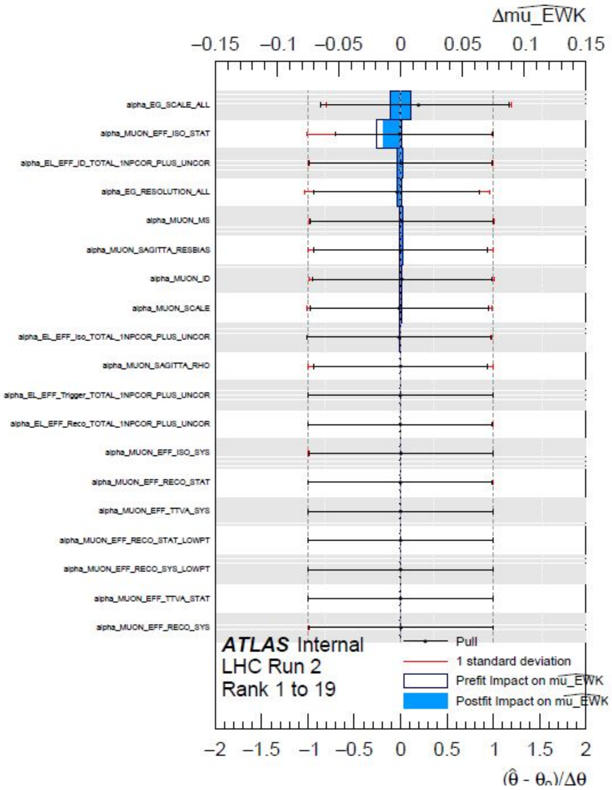
Other



MC stats



Lepton



Combination

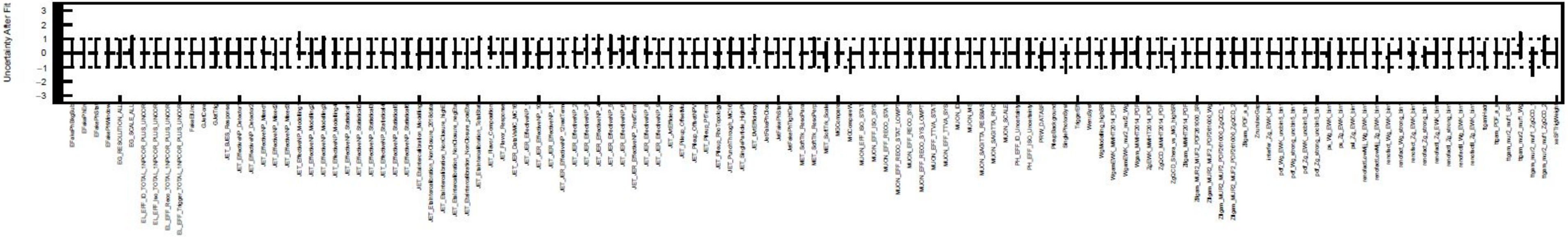
ke1 = 0.748487 +/- 0.348663 (limited)
ke2 = 0.49387 +/- 0.385764 (limited)
ke3 = 0.119505 +/- 0.221266 (limited)
ke4 = 0.187899 +/- 0.178368 (limited)
mu_EWK =1 +/- 0.180418 (limited)
mu_QCD = 1.07007 +/- 0.149665 (limited)
mu_Wg = 1.0381 +/- 0.116488 (limited)
Observed pValue: 1.48207e-10
Median test stat val: 44.9834
Median significance: 6.70696
Median pValue: 9.93583e-12

Expected significance=**6.7σ**

Large pull
JET_EffectiveNP_Modelling1 - 0.524852 +\- 0.935979 (mostly from EXOT WS)
ttgam_mur2_muf1_Wg - 0.549696 +\- 0.896381 (from SM WS)
ttgam_mur2_muf1_ZgQCD_1 - -0.593573 +\- 0.95735 (from SM WS)

Over constrained
renofact_Wg_strong_bin1 - 0.208557 +\- 0.553312 (from EXOT WS)
renofact_Zg_strong_bin1 - -0.222028 +\- 0.620581 (from EXOT WS)

Combination



Current issues

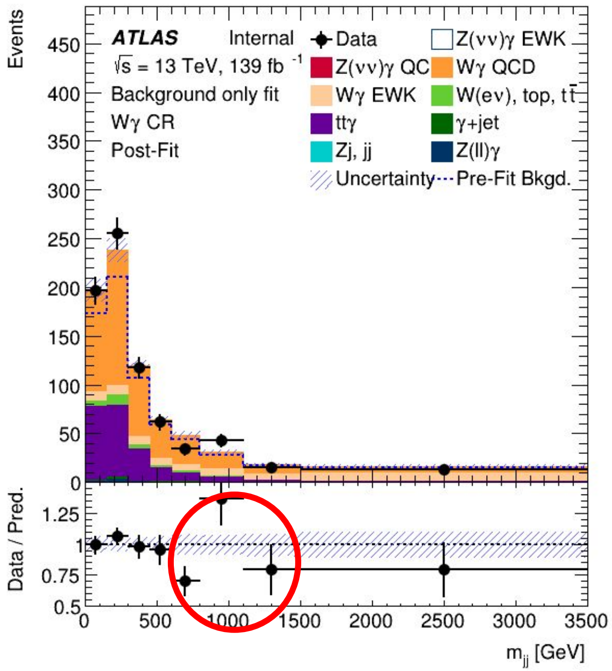
1) Decide on the scale systematics model: no systematics/decorrelated between regions systematics/correlated among the regions systematics

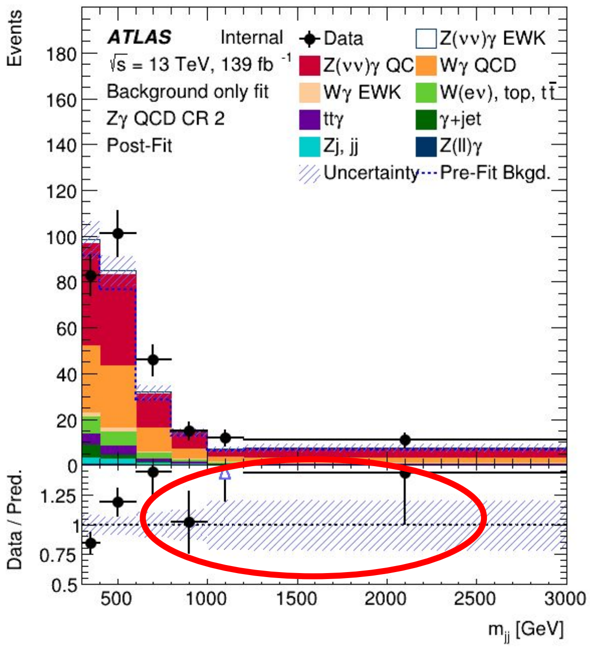
2) TReXFitter somehow gives a bit different results in case it was used for all steps or its Workspace was fitted independently
    - We've temporarily solved the problems with NPs (pulls/underconstraints) by smoothing them.
    - However mu_EWK is not 1, if SM workspace is used outside.

Feedback

- To check different schemes of JER NPs correlations
- To check impact of custom calculation of jet flavor uncertainty in EXOT

Current issues (SM)

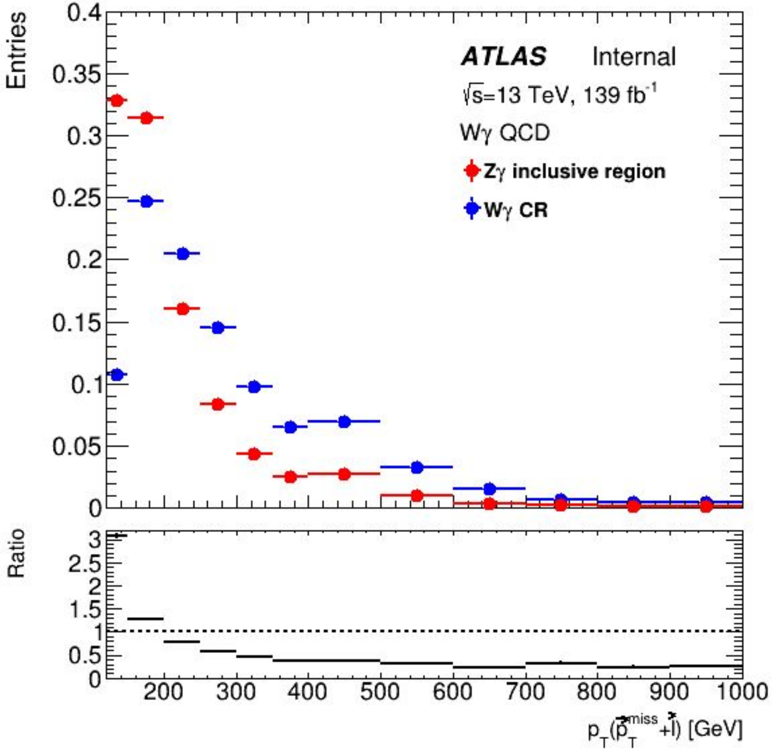


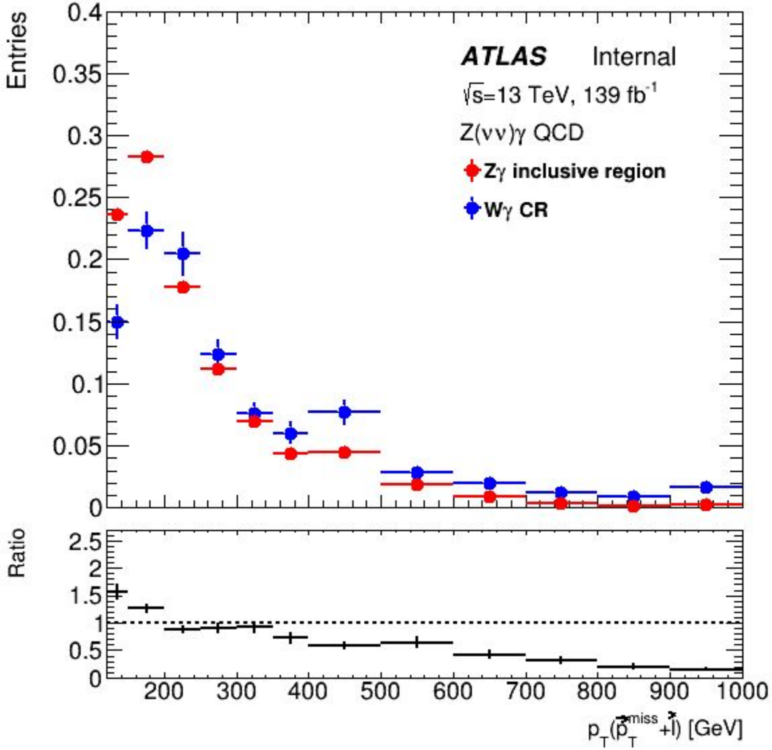


Mismodelling?
Of Z**γ** QCD only or both Z**γ** QCD and W**γ** QCD?

Current issues (SM)

Andy Pilkington: By topologically similar, I mean that the jet pt cuts are the same and that the pt of the Wy system is forced to be similar to the pt of the Zy system (i.e. you have a cut on photon pt and MET>XGeV, which can be mimicked in the Wy CR by requiring pt(lep+MET)>XGeV). If it is topologically similar then you expect the same mismodelling in both regions.



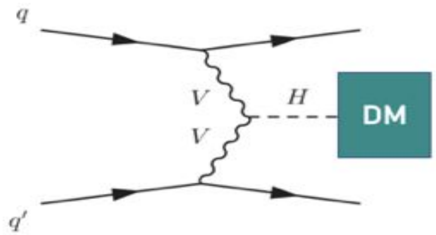


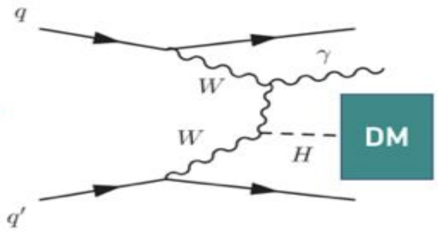
Back-up slides

Intro

Why + Photon?

VBF + MET → **VBF + MET + photon**





- The analysis will be the first search for the VBF+photon+H-> Invisible, which deserve its own paper
- Full Run-2 dataset
- The VBF+MET dominant backgrounds (QCD V+jets) are greatly suppressed requiring one ISR photon
- Production mode has been proposed for some time
  - Nucl. Phys. B781 p64-84 (2007) , JHEP 1607 (2016) 003
- Production mode already utilized in an ATLAS analysis
  - Utilized H→bb: Phys. Rev. D 98, 052003 (2018)

3

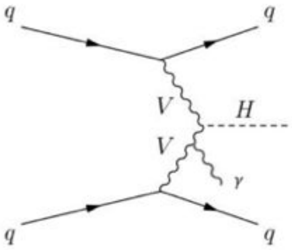
x

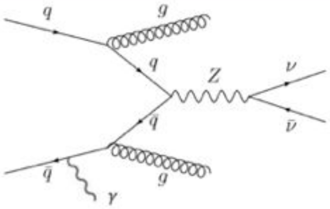
Signal and Backgrounds

Signal and Backgrounds

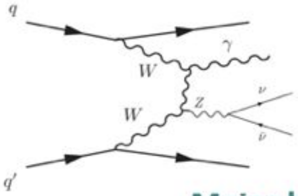
Signature:

- Two jets
  - High m_{jj}
  - Highly separated in η
  - Not back to back
- One photon
- High MET





Strong production



EWK production

Major backgrounds:

- $Z(\rightarrow \nu\nu)\gamma + jets$
- $W(\rightarrow \nu l_{lost})\gamma + jets$

6

X

Background Treatment

Table 70: Yields of signal and major backgrounds 139 fb^{-1} selection using the 2015–2018 dataset. The yields prefit.

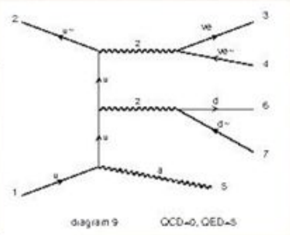
Samples	SR	$Z(\to \ell\ell) + \gamma$ VR	$W(\to \mu\nu) + \gamma$ CR	$W(\to e\nu) + \gamma$ CR	Fake-e CR
VBFγH125	86.853	0.000	0.000	0.028	0.000
ggFH125	7.296	0.000	0.000	0.000	0.000
$Z\gamma$ QCD	82.845	15.975	1.964	0.000	6.071
$Z\gamma$ EWK	66.492	14.660	0.195	0.236	0.455
$W\gamma$ QCD	47.566	0.000	97.960	35.968	34.948
$W\gamma$ EWK	12.288	0.000	45.380	24.195	13.831
Top/VV/VVV/VBFWW	6.459	0.076	10.765	6.724	2.617
γ+j	0.980	0.000	0.000	0.000	0.000
$e \to \gamma$	17.132	0.000	2.564	2.397	0.000
$j \to \gamma$	3.916	0.110	2.341	1.593	1.310
eleFakes	0.000	0.000	0.000	3.000	21.300
data	251	34	170	86	84
total bkg	237.678	30.821	161.170	74.113	80.533
data/bkg	1.056	1.103	1.055	1.160	1.043

- **Gamma+jet** enters the signal region through reconstruction of fake MET
  - Use a jet smearing approach with normalization to a low MET validation region
- **Electron faking photons**
  - Measured the fake rate of electrons being reconstructed as photons in the Z mass peak. Then apply this eta and pT dependent correction to Medium LH electrons
- **Jets faking photons:** measured using an ABCD method in isolation and PID. Checked dependence in photon pT, mjj and MET. Large uncertainties but it is a very small background
- **Top/VV/VVV**: normalized to their cross-section. No special theoretical treatment is made
- Lepton veto uncertainty for Wg in the 0-lepton selection

Lepton inefficiency	
eleANTISFEL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID inefficiency uncertainty
γ + jet Background	
GJetCore	Core+tail+truth E_T^{miss} variations
GJetTrig	MET Trigger variations and normalisation
e-fake-γ Background	
EFakePhWindow	Vary mass window
EFakePhStat	Fake rate stat unc.
EFakePhBkgSub	Bkg subtraction unc.
EFakePhEn	Energy differences between γ and e

X

EWK Triboson



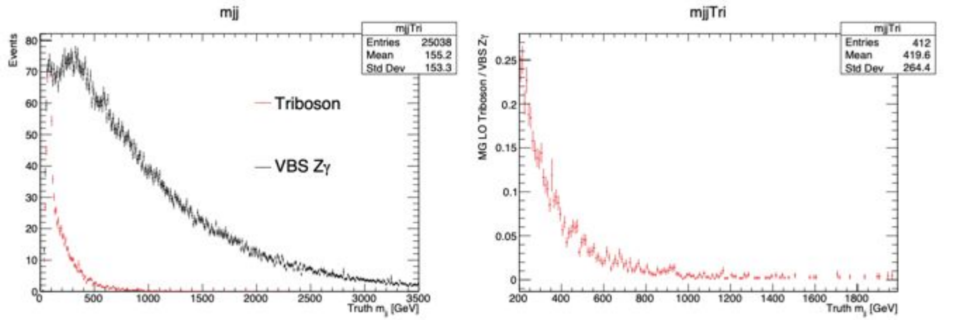


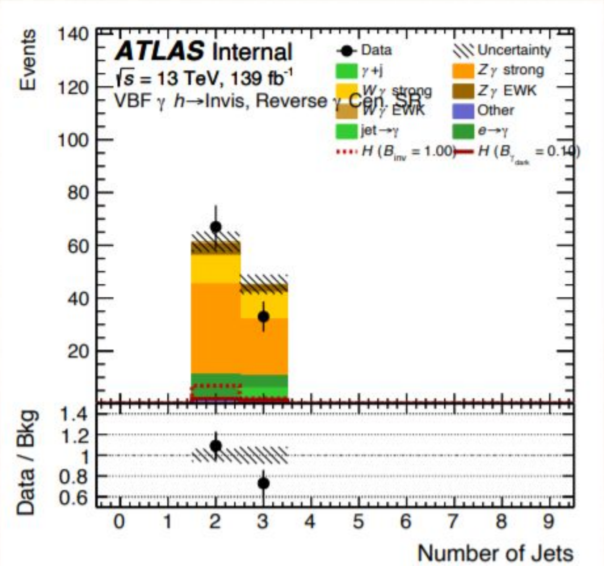
Figure 76: The fraction of triboson versus VBS Zgjj in the LO EWK MG samples at truth level using the same selections as Section 7.1.2.

- Electroweak Triboson is included in the madgraph samples
- We have not separated triboson and EWK Zg at this point, so it would be part of the EWK Zg signal
- Does the SM group separate the triboson processes out from the signal sample? Perhaps we should?
  - No major impact is expected given that the contamination for Mjj>500 GeV is less than 2% at LO. Most of the signal has mjj>500 GeV

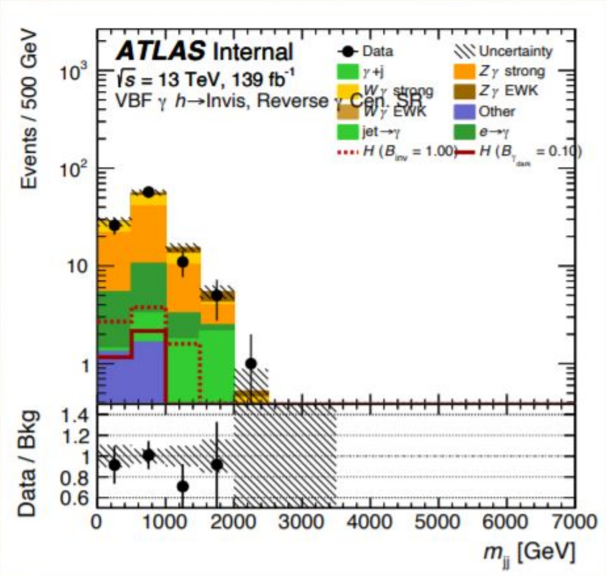
X

Reversing Photon Centrality

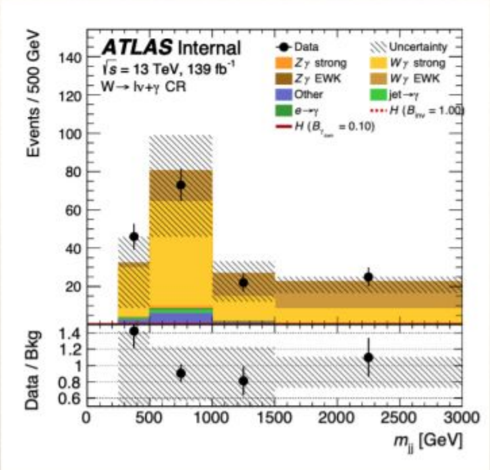
- Reversing the photon centrality cut to control the Zg strong background
  - Used to normalize the Zg strong background
  - 50% purity of Zg strong

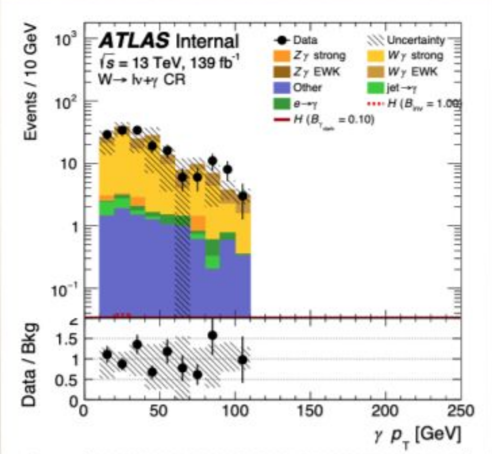


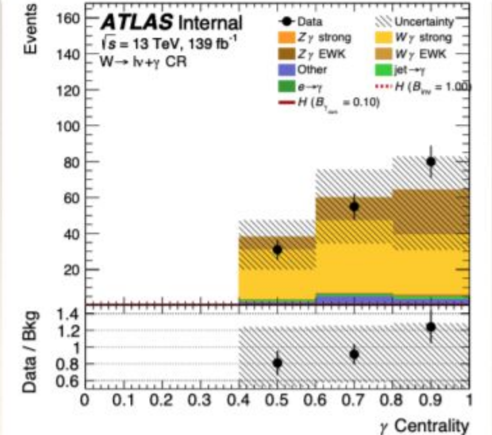
11



One Lepton CR







- One lepton CR
  - 87% pure in Wg
  - Data/MC differences are covered by systematic+statistical uncertainties

PreFit vs PostFit [Note this H->inv signal]

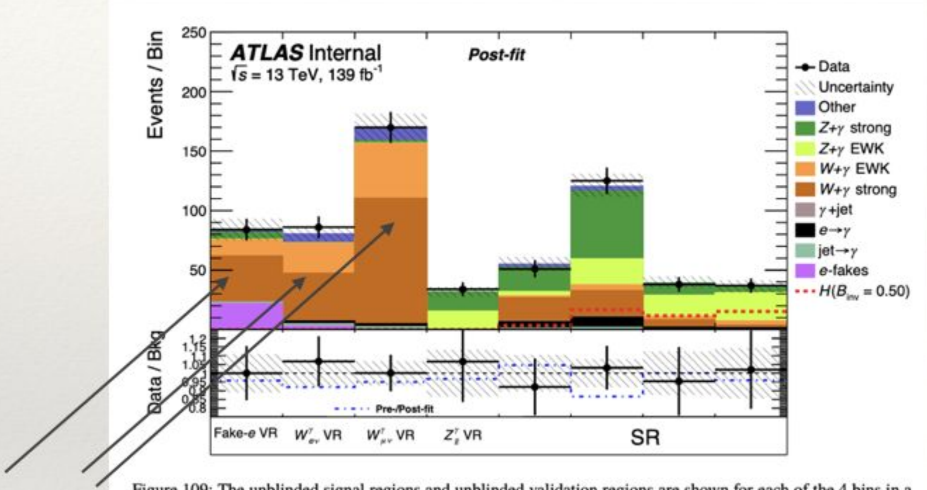


Figure 109: The unblinded signal regions and unblinded validation regions are shown for each of the 4 bins in a signal floated fit. The post-fit errors are statistical plus reconstruction plus theory systematics, and the normalisations are post fit.

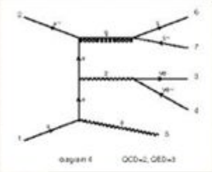
Control regions for EWK Zg fit

These are the fit regions except the Z(ll) one.

SR bins are in mjj: 250,500,1000,1500 GeV

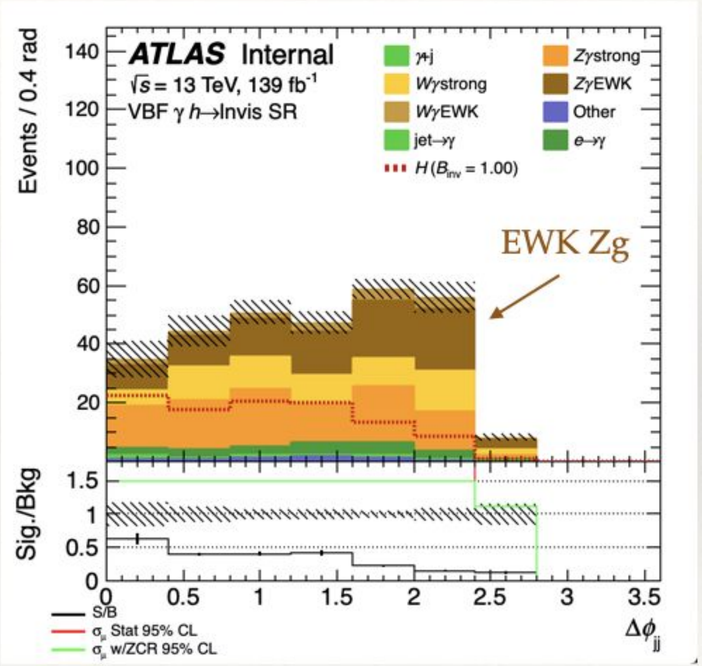
- Plot above is the post-fit SR with dphijj<2 fitting for the Higgs to invisible signature. This is just an example as we'd will prepare a plot for the EWK Zg fit

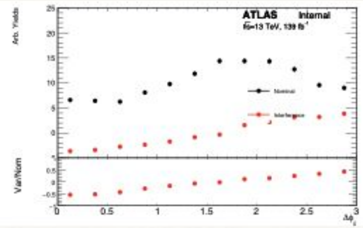
x

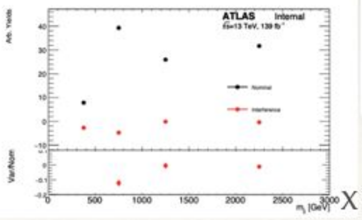


Δϕjj Selection & Keras NN

- Nominally we cut $\Delta\phi_{jj}<2$ for H->invisible
  - If we extend to $\Delta\phi_{jj}<2.5$, we reduce the interference uncertainties
  - Expected significance goes from 4.1->4.6σ for EWK Zγ
  - Would suggest we make this change for the Higgs->invisible analysis.
- keras NN:
  - Expected limit improves slightly from 0.36->0.35 for H->invisible & it reduces the interference uncertainty that we missed during approval
  - Expected significance for EWK Zγ is closer to 3.3σ because it is tuning for lower $\Delta\phi_{jj}$.
  - Prefer the m_{jj} fit







Systematics Treatment

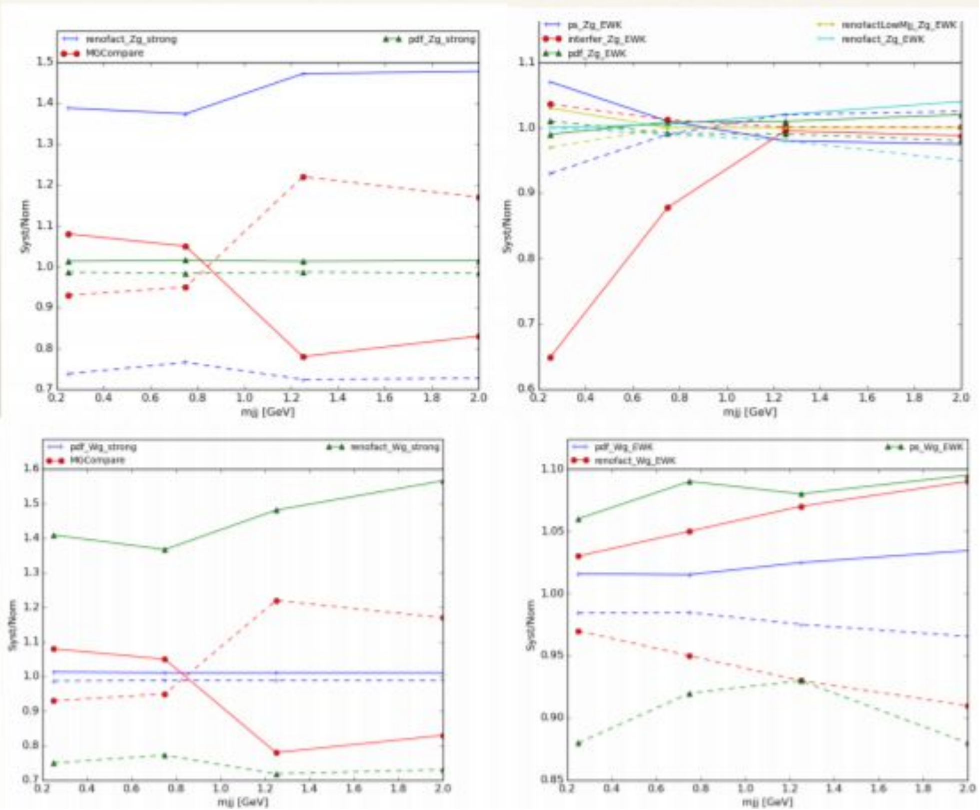
- MET Trigger uncertainty is applied to account for the MET trigger turn on curve
- Full JER is used. typically small so maybe this is ok for combination?
  - This analysis uses p-flow jets
- Theory uncertainties:
  - **Zg strong**: scale variations (these are dropped in the SM analysis), PDF (one set), comparison to MG
  - **Wg strong**: scale variations (these are dropped in the SM analysis), PDF (one set), (**comparison to MG to be added using the Zg strong comparison**)

  - **Zg EWK**: scale variations from VBFNLO, PDF(one set), comparison to H7 parton shower model, electroweak/strong interference
  - To match the SM analysis, these uncertainties are fully correlated across all SR analysis bins. Both the shape and normalization are considered as uncertainties.
  - One difference is that we correlated the Wg strong uncertainties also in the one-lepton CRs, whereas the SM analysis has these uncorrelated. We correlated them with the SR and CR to cancel the normalization component of these uncertainties

E_T^{miss}-Trigger and E_T^{miss}-Terms	
xeSFTrigWeight	trigger efficiency uncertainty
MET_SoftTrk_ResoPerp	track-based soft term related to transversal resolution uncertainty
MET_SoftTrk_ResoPara	track-based soft term related to longitudinal resolution uncertainty
MET_SoftTrk_ScaleUp	track-based soft term related to longitudinal scale uncertainty
MET_SoftTrk_ScaleDown	track-based soft term related to longitudinal scale uncertainty

Theory uncertainties (EXOT analysis)

- Theory systematic uncertainties in the SR
- Zg EWK theory uncertainties were normalized to the fiducial volume
  - Only acceptance+shape uncertainties should be included for the signal



Theory uncertainties (EXOT analysis)

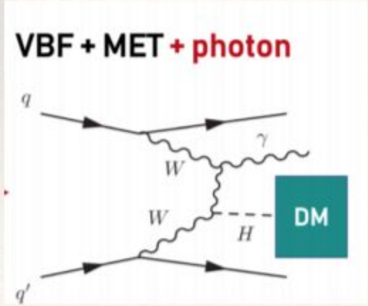
PostFit Yields

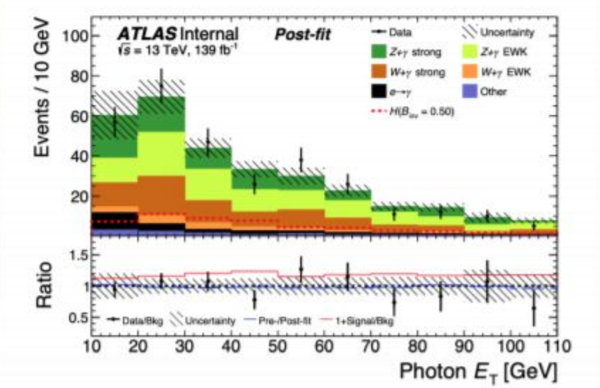
Table 84: Yields of data and background predictions, after the fit to Data, in 139 fb^{-1} for the four bins of the SR based on the m_{jj}. The uncertainties on the backgrounds are derived by the fit including the effects of nuisance parameter constraints and the correlation of systematic uncertainties. A dash "-" indicates less than 0.01 events. The uncertainties on the "Data/Bkg" treat the Data and Bkg components as uncorrelated.

Process	Fake−e CR	$W^{\gamma}_{e\nu}$ CR	$W^{\gamma}_{\mu\nu}$ CR	$Z^{\gamma}_{\text{Rev.Cen.}}$ CR	SR - m_{jj} [TeV] 0.25-0.5	0.5-1.0	1.0-1.5	≥ 1.5
Strong $Z\gamma$+jets	8± 8	0± 1	3± 2	50± 12	20± 6	54± 12	13± 5	5± 2
EWK $Z\gamma$+jets	0.6± 0.2	0.3± 0.2	0.4± 0.2	7± 2	4± 1	30± 7	25± 5	36± 7
Strong $W\gamma$+jets	43± 9	47± 9	133± 21	24± 6	22± 6	35± 10	9± 3	3± 1
EWK $W\gamma$+jets	19± 6	31± 7	59± 13	1.4± 0.5	2± 1	6± 1	4± 1	5± 1
jet→ γ	1± 1	2± 2	3± 2	2± 2	1± 1	2± 2	1± 1	0.4± 0.3
jet→ e	34± 17	5± 3	-	-	-	-	-	-
$e \to \gamma$	-	2.7± 0.4	2.9± 0.4	13± 1	6± 1	11± 1	2.6± 0.4	1.4± 0.3
γ + jet	-	-	-	0.7± 0.5	0.7± 0.5	0.4± 0.3	0.1± 0.1	0.1± 0.1
$t\bar{t}\gamma/V\gamma\gamma$	3± 1	9± 2	13± 2	3± 1	2± 1	4± 1	0.4± 0.2	0.1± 0.1
Total Bkg	108± 10	96± 8	213± 14	102± 9	58± 6	143± 12	54± 5	52± 6
Data	108	95	216	100	52	153	50	52
Data/Bkg	1.00± 0.14	0.99± 0.12	1.01± 0.09	0.98± 0.13	0.90± 0.15	1.07± 0.11	0.93± 0.16	0.99± 0.18

Plans (EXOT analysis)

- Publication plans for the exotics results:
  - Exotics results should go out for a CONF note for Moriond EWK: dark photon search + H->invisible
  - SM EWK Zγ (exotics) result will be included in an EPJC paper along with the exotics results shortly after Moriond
    - After this approval, the SM results should be added to the CONF note within 1-2 weeks in order to start the paper approval process
    - Plan to include the SM EWK Zg team in the author list given the extensive discussions
- In the SM paper, planned combination with the SM analysis and the combination is on-going with ~6.5σ expected
  - Initial discussion tomorrow https://indico.cern.ch/event/1010000/



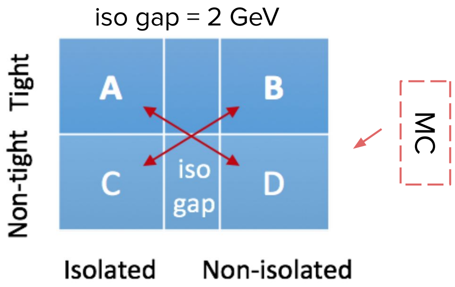


(a) Photon p_T distribution in SR

21

jet→γ misID background I: correlation factor

Source: Z(νν)+jets and multi-jet processes.
Background is estimated from data using **2D-sideband method**:
Photon isolation and identification variables are used to construct the sidebands.



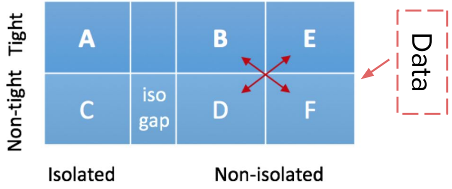
$p_T^{cone20}/p_T^{\gamma} < 0.05$

FixedCutTight:

A: tight, $E_T^{cone40} - 0.022\, p_T^{\gamma} < 2.45$ [GeV]
B: tight, $2.45 + \text{gap} < E_T^{cone40} - 0.022\, p_T^{\gamma} < 29.45$ [GeV]
C: non-tight, $E_T^{cone40} - 0.022\, p_T^{\gamma} < 2.45$ [GeV]
D: non-tight, $2.45 + \text{gap} < E_T^{cone40} - 0.022\, p_T^{\gamma} < 29.45$ [GeV]

Non-tight: at least one of the cuts on the following variables should fail in these:

- *loose'2*: w_{s3}, F_{side}
- *loose'3*: $w_{s3}, F_{side}, \Delta E$
- *loose'4*: $w_{s3}, F_{side}, \Delta E, E_{ratio}$
- *loose'5*: $w_{s3}, F_{side}, \Delta E, E_{ratio}, w_{tot}$



Isolation should not correlate with non-tight ID!

$$\frac{N_A^{jet\to\gamma}}{N_B} = \frac{N_C}{N_D}$$

Correlation is measured in data and MC by $R = \frac{N_A N_D}{N_B N_C}$.

(relaxed selections)

with upper cut

R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.02 ± 0.12	1.12 ± 0.12	1.21 ± 0.12	1.41 ± 0.14
Data-driven	0.92 ± 0.07	0.93 ± 0.07	0.88 ± 0.06	0.90 ± 0.06

Best agreement
Smallest correlation

without upper cut

R factor	*loose'2*	*loose'3*	*loose'4*	*loose'5*
MC	1.01 ± 0.12	1.11 ± 0.12	1.19 ± 0.11	1.39 ± 0.13
Data-driven	0.71 ± 0.04	0.73 ± 0.04	0.67 ± 0.03	0.66 ± 0.03

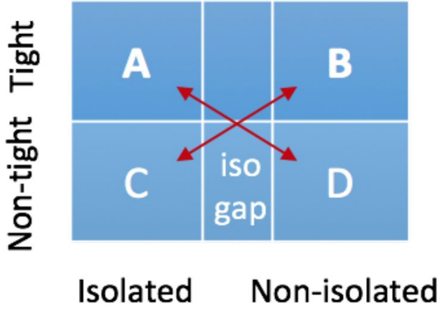
Unstable

Used as nominal

Purity	*loose'2*	*loose'3*	*loose'4*	*loose'5*
Region C	0.16 ± 0.06	0.16 ± 0.04	0.15 ± 0.03	0.14 ± 0.03
Region D	0.37 ± 0.10	0.37 ± 0.10	0.29 ± 0.05	0.30 ± 0.05

jet→γ misID background II: estimation technique

iso gap = 2 GeV



The number of events arising in each of the regions:

$$N_{\rm A} = N_{\rm A}^{Z(\nu\bar{\nu})\gamma} + N_{\rm A}^{\rm bkg} + N_{\rm A}^{\rm jet\to\gamma};$$
$$N_{\rm B} = c_{\rm B} N_{\rm A}^{Z(\nu\bar{\nu})\gamma} + N_{\rm B}^{\rm bkg} + N_{\rm B}^{\rm jet\to\gamma};$$
$$N_{\rm C} = c_{\rm C} N_{\rm A}^{Z(\nu\bar{\nu})\gamma} + N_{\rm C}^{\rm bkg} + N_{\rm C}^{\rm jet\to\gamma};$$
$$N_{\rm D} = c_{\rm D} N_{\rm A}^{Z(\nu\bar{\nu})\gamma} + N_{\rm D}^{\rm bkg} + N_{\rm D}^{\rm jet\to\gamma};$$

The signal leakage parameters:

$$c_{\rm B} = \frac{N_{\rm B}^{Z(\nu\bar{\nu})\gamma}}{N_{\rm A}^{Z(\nu\bar{\nu})\gamma}}, \quad c_{\rm C} = \frac{N_{\rm C}^{Z(\nu\bar{\nu})\gamma}}{N_{\rm A}^{Z(\nu\bar{\nu})\gamma}}, \quad c_{\rm D} = \frac{N_{\rm D}^{Z(\nu\bar{\nu})\gamma}}{N_{\rm A}^{Z(\nu\bar{\nu})\gamma}}$$

MC →

$$N_{\rm A}^{Z(\nu\bar{\nu})\gamma} = \tilde{N}_{\rm A} - (\tilde{N}_{\rm B} - c_{\rm B} N_{\rm A}^{Z(\nu\bar{\nu})\gamma}) \frac{\tilde{N}_{\rm C} - c_{\rm C} N_{\rm A}^{Z(\nu\bar{\nu})\gamma}}{\tilde{N}_{\rm D} - c_{\rm D} N_{\rm A}^{Z(\nu\bar{\nu})\gamma}}$$

($\tilde{N}_i$ - data N_i with subtracted $N_i^{\rm bkg}$)

$$a = c_{\rm D} - c_{\rm B} c_{\rm C};$$
$$b = \tilde{N}_{\rm D} + c_{\rm D}\tilde{N}_{\rm A} - (c_{\rm B}\tilde{N}_{\rm C} + c_{\rm C}\tilde{N}_{\rm B});$$
$$c = \tilde{N}_{\rm D}\tilde{N}_{\rm A} - \tilde{N}_{\rm C}\tilde{N}_{\rm B}.$$

$$N_{\rm A}^{Z(\nu\bar{\nu})\gamma} = \frac{b - \sqrt{b^2 - 4ac}}{2a}$$

→ $N_{\rm A}^{\rm jet\to\gamma}$

	Data	$Z(\nu\nu)\gamma$ QCD	$W\gamma$ QCD	$W\gamma$ EWK	$W(e\nu), top, tt$	$tt\gamma$	γ+jet	$Z(ll)\gamma$	$W(\tau\nu)$
A	blinded	569.5 ± 1.7	396 ± 6	41.6 ± 0.4	97 ± 3	87 ± 2	112 ± 10	11.0 ± 0.8	7 ± 3
B	101 ± 10	29.4 ± 0.4	16.1 ± 1.9	1.51 ± 0.08	3.7 ± 0.2	4.5 ± 0.4	2.1 ± 1.1	0.5 ± 0.2	5.2 ± 1.2
C	38 ± 6	5.08 ± 0.16	3.9 ± 0.6	0.42 ± 0.04	1.30 ± 0.14	1.1 ± 0.3	0.8 ± 0.7	0.11 ± 0.05	5.1 ± 1.9
D	27 ± 5	0.22 ± 0.03	0.14 ± 0.08	0.015 ± 0.007	0.4 ± 0.4	0.14 ± 0.08	0.3 ± 0.3	0 ± 0	6.0 ± 1.4

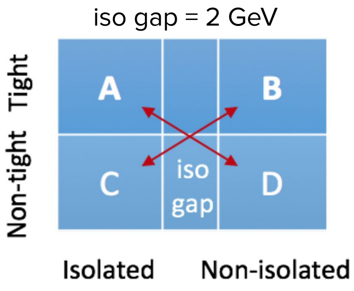
jet→γ misID background III: uncertainties

➢ **Statistical uncertainty:**

- The event yields of four regions in data and non jet ➜ γ background are varied by ±1σ independently.
- The statistical uncertainty on the signal leakage parameters is **3%**.

Total statistics: 60%.

Central value	35^{+19}_{-21}
Loose'3	−6
Loose'4	−2
Loose'5	−5
Isolation gap +0.9 GeV	+8
Isolation gap −0.6 GeV	±0



➢ **Systematic uncertainty:**

- Anti-tight definition and isolation gap choice – variations of ABCD region determination for ±1σ changes in data yield (**23%**).
- Uncertainty coming from the signal leakage parameters is obtained via using two different parton shower models (**0%**).

Signal leakage parameters	MadGraph+Pythia8	MadGraph+Herwig7
c_B	0.0313 ± 0.0006	0.0315 ± 0.0006
c_C	0.0085 ± 0.0003	0.0089 ± 0.0003
c_D	0.00031 ± 0.00005	0.00043 ± 0.00006
$jet \to \gamma$ estimation	35^{+19}_{-21}	35^{+19}_{-21}

- The iso/ID uncertainty on reconstruction photon efficiency $\delta_{eff}^{iso/ID}$ (**9%**):

  - $\sigma_{iso}^{c_B}(\text{relative}) = \delta_{iso}^{eff} * (c_B + 1)/c_B$
  - $\sigma_{ID}^{c_C}(\text{relative}) = \delta_{ID}^{eff} * (c_C + 1)/c_C$
  - $\sigma_{iso}^{c_D}(\text{relative}) = \delta_{iso}^{eff} * (c_B + 1)/c_B$
  - $\sigma_{ID}^{c_D}(\text{relative}) = \delta_{ID}^{eff} * (c_C + 1)/c_C$

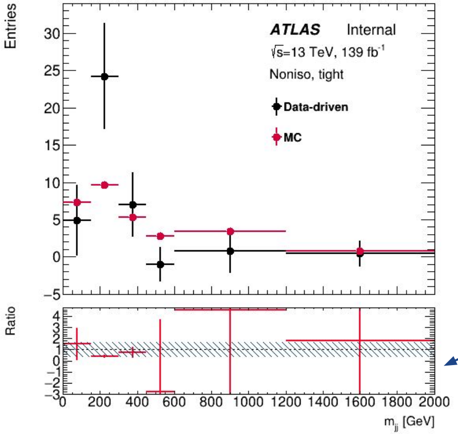
$\delta^{eff}_{iso} = 0.023$
$\delta^{eff}_{iso/ID} = 0.019$

Total systematics: 25%.

★ **Resulting number** of jet➜γ events in <u>Zγ inclusive region</u> is $35^{+19}_{-21} \pm 9$. Z(νν)+jets and multi-jet MC predict 7±2 events.

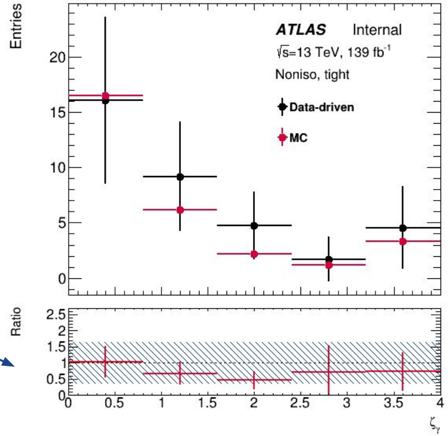
jet→γ misID background IV: Zγ inclusive and signal regions

The extrapolation of jet→γ background estimation from Zγ inclusive region to <u>the signal region</u>:

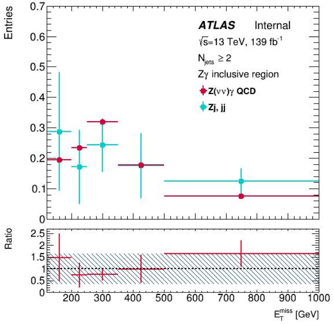


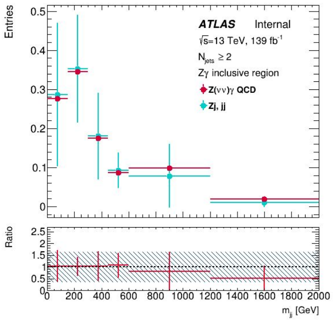
$Z\gamma jj$ EWK signal region	
N_{leptons}	$= 0$
m_{jj}	> 400 GeV
γ-centrality	< 0.6

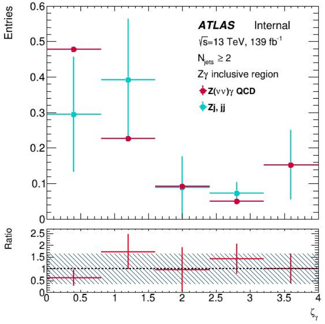
Statistical and systematic uncertainties of jet→γ background estimation



The current uncertainty for this estimation covers all differences and should not be increased.







Good agreement of the shapes.

⇩

The shape of jet → γ background for normalization is taken from Zγ QCD.

Electron misidentification as photon (e⟶γ)

Background estimation method:

1. estimating e→γ fake-rate as $rate_{e\to\gamma} = \frac{(N_{e\gamma} - N_{bkg})}{(N_{ee} - N_{bkg})}$,

   where $N_{e\gamma}$, N_{ee} − number of ee and eγ events in Z-peak mass window (M_Z−10 GeV, M_Z+10 GeV),

   N^{bkg} − background in Z-peak mass window extrapolated from sideband with exponential pol1 or pol2 fit (more details below)

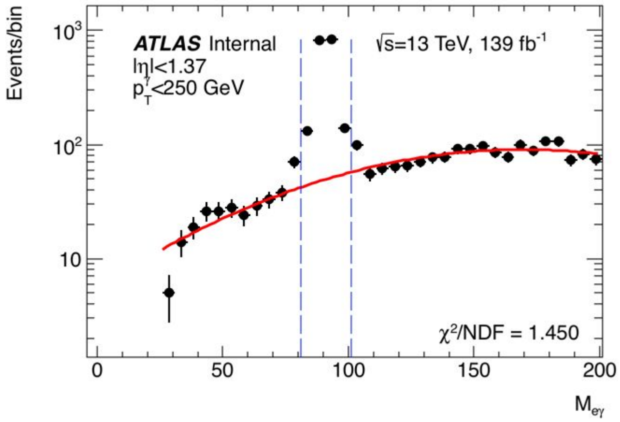
   Wγ background rejection: E_T^{miss} < 40 GeV

   eγ pair selection:

   signal region photon with p_T>150 GeV (probe), selected Tight electron with p_T>25 GeV (tag)

   ee pair selection:

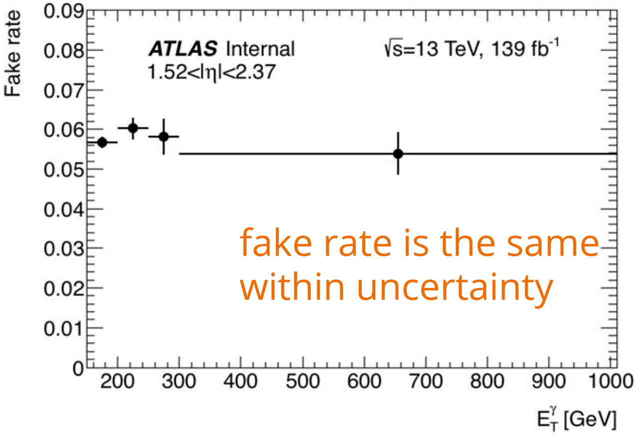
   selected electron with p_T>150 GeV (probe), selected opposite sign Tight electron with p_T>25 GeV (tag)

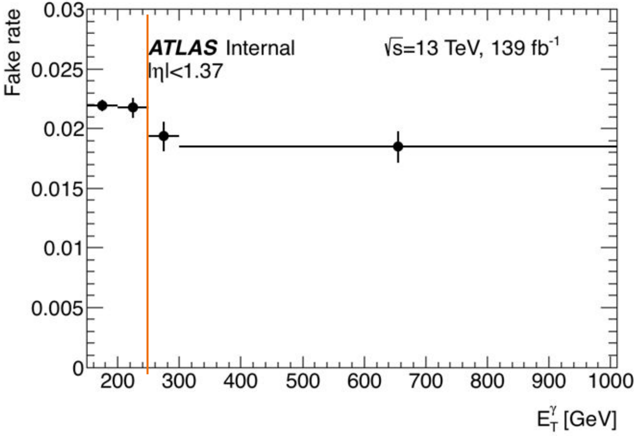


2. building e-probe/incl. e-probe control region (CR): signal region/Zγ incl. region with selected Tight electron with p_T>150 GeV instead of photon

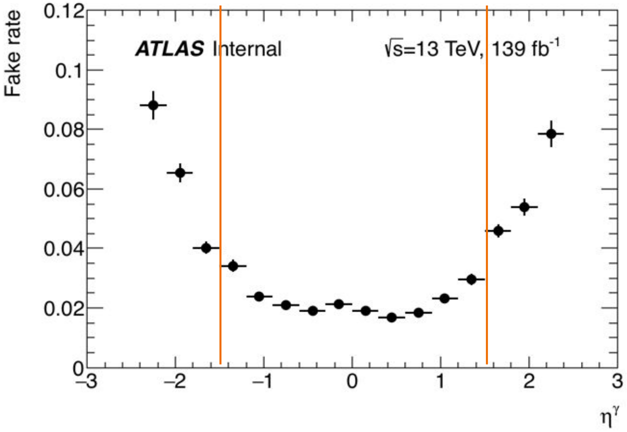
3. scaling data distributions from e-probe CR/incl. e-probe CR on fake rate

e→γ fake rate dependencies on eta and pT





There are no background subtraction under Z peak on the plots

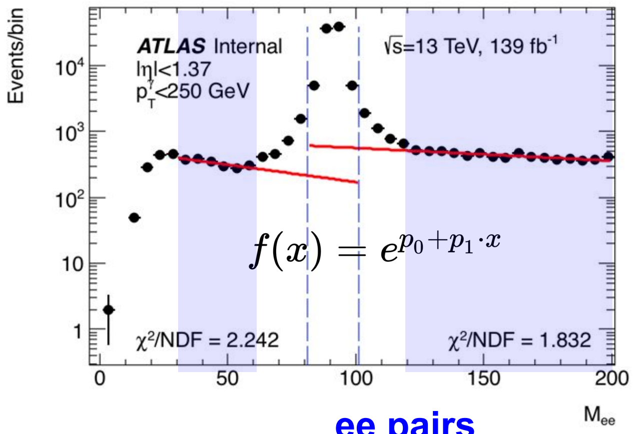


Since e→γ fake rate depends on η и p_T it is estimated in three regions

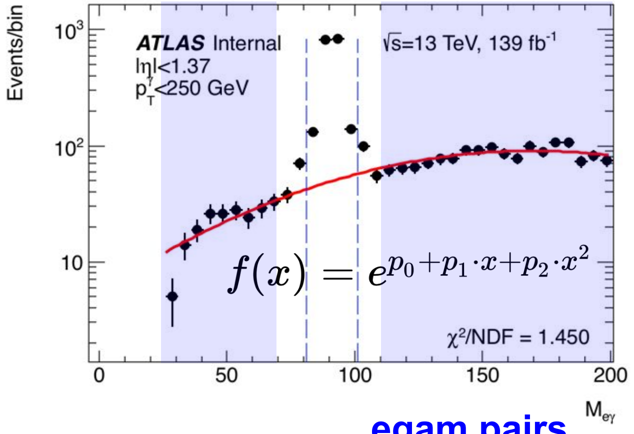
|η|<1.37, p_T<250 GeV
|η|<1.37, p_T>250 GeV
1.52<|η|<2.37

finner division isn’t possible due to insufficient statistics for under Z-peak bkg subtraction

e⟶γ: background under Z-peak



ee pairs



egam pairs

fit range

Background fit is extrapolated to Z peak mass window from both sides. Integrals under the fit function is this region give N_{min} and N_{max}
Average used as N_{ee}^{bkg} in fake-rate calculation: $N^{bkg}_{average} = \frac{N^{bkg}_{min}+N^{bkg}_{min}}{2}$

Background fit is extrapolated to Z peak mass window after the fit. Integral of extrapolated function in Z peak mass window is used as $N_{e\gamma}^{bkg}$

Systematics on bkg estimation under the Z peak is evaluated by variation of N^{bkg} values in ee and eγ pairs.

N_{min} and N_{max} values are used as variations of N_{ee}^{bkg}. In eγ pairs extrapolation function parameters were varied by their statistical uncertainties one by one. Resulting integral of the function is used for variation of $N_{e\gamma}^{bkg}$. Sum in quadrature of the largest variations of $N^{bkg}_{e\gamma}$ and N^{bkg}_{ee} is taken as systematics.

e→γ: systematics and result

Systematics on fake-rate estimation (ascending contribution):

- Z peak mass window size variation by σ
- Background under Z peak evaluation
- Difference between "real fake rate" in Z(ee) MC and tag-and-probe method performed on *Z*(*ee*) MC

fake rate	$150 < E_T^\gamma < 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$E_T^\gamma > 250$ GeV $0 < \lvert\eta\rvert < 1.37$	$1.52 < \lvert\eta\rvert < 2.37$
syst. from mass window var.:	0.4%	0.3%	0.6%
syst. from tag-n-probe and real f.r.:	3.1%	13.4%	8.7%
Background fit variation	3.7%	10.8%	3.3%
Total syst.:	4.8%	17.2%	9.3%

$e \to \gamma$ fake rates		
	$150 < E_T^\gamma < 250$ GeV	$E_T^\gamma > 250$ GeV
$0 < \lvert\eta\rvert < 1.37$	$0.0205 \pm 0.0005 \pm 0.0010$	$0.0183 \pm 0.0012 \pm 0.0032$
$1.52 < \lvert\eta\rvert < 2.37$	$0.0571 \pm 0.0016 \pm 0.0053$	

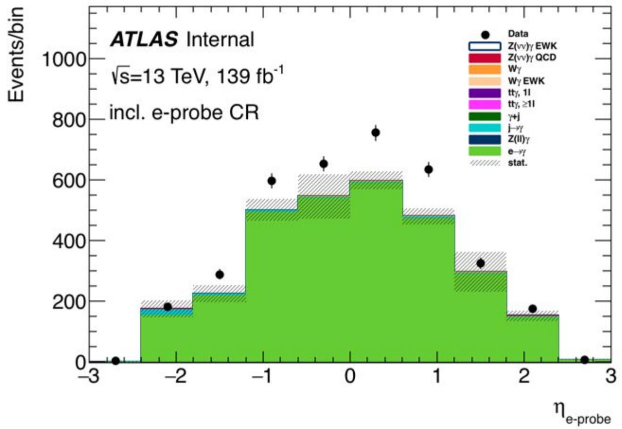
First uncertainty is stat., second is syst.

Total background (e-probe region scaled by fake-rate):

Zγ inclusive region 97 ± 2 ± 4

Signal region 19.5 ± 0.8 ± 0.7

Total systematics does not exceed *4.12%*



Percent of e-probe CR contamination is taken as additional systematic uncertainty: 1-2%
Contamination is determined as

$$1 - \frac{all\ other\ MC}{W(e\nu) + t\bar{t} + t\ MC}$$

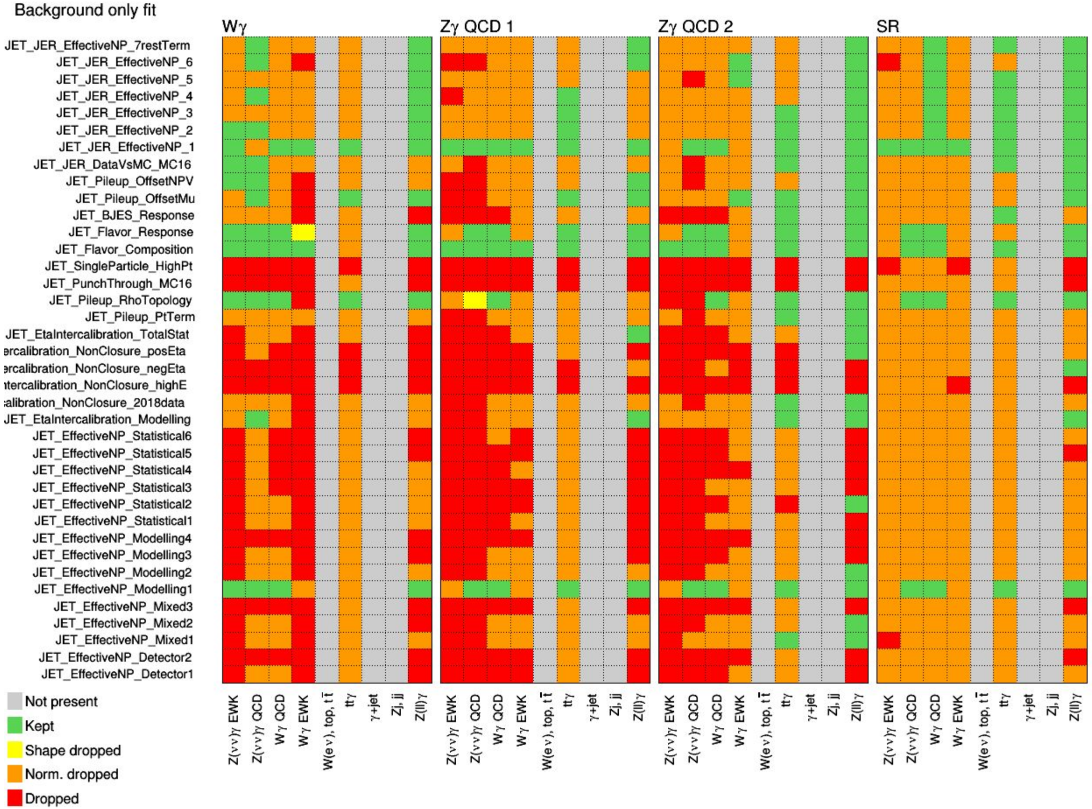
Total systematics in every region of fake-rate estimation combines fake-rate statistical unc., fake-rate systematics unc., systematics from contamination

Total event yields (pre-fit)

	$W\gamma$ CR	$Z\gamma$ inclusive	$Z\gamma$ QCD CR 1	$Z\gamma$ QCD CR 2	Signal region
$Z(\nu\nu)\gamma$ EWK	$0.108 \pm 0.008 \pm 0.042$	$63.73 \pm 0.2 \pm 24.77$	$13.71 \pm 0.09 \pm 5.96$	$4.72 \pm 0.05 \pm 2.23$	$46.16 \pm 0.17 \pm 17.92$
$Z(\nu\nu)\gamma$ QCD	$0.98 \pm 0.07 \pm 0.58$	$569.4 \pm 1.7 \pm 267.7$	$353.4 \pm 1.4 \pm 164.6$	$101.5 \pm 0.7 \pm 47.8$	$114.5 \pm 0.8 \pm 55.8$
$W\gamma$ QCD	$390 \pm 6 \pm 183$	$396 \pm 6 \pm 187$	$237 \pm 5 \pm 109$	$68 \pm 2 \pm 34$	$91 \pm 3 \pm 45$
$W\gamma$ EWK	$61.3 \pm 0.5 \pm 21.3$	$41.2 \pm 0.4 \pm 15.3$	$13.7 \pm 0.2 \pm 5.2$	$4.34 \pm 0.13 \pm 1.71$	$23.5 \pm 0.3 \pm 8.8$
$W(e\nu)$, top, $t\bar{t}$	$19.1 \pm 0.8 \pm 1.3$	$97.0 \pm 1.8 \pm 6.8$	$61.8 \pm 1.5 \pm 4.3$	$15.6 \pm 0.8 \pm 1.1$	$19.7 \pm 0.8 \pm 1.4$
$tt\gamma$	$168 \pm 3 \pm 63$	$86.8 \pm 2.0 \pm 35.7$	$57.8 \pm 1.6 \pm 29.5$	$8.7 \pm 0.6 \pm 8.3$	$20.3 \pm 1.0 \pm 9.2$
γ+jet	$6.0 \pm 5.3 \pm 1.9$	$90 \pm 20 \pm 30$	$63 \pm 21 \pm 19$	$9 \pm 7 \pm 3$	$23 \pm 11 \pm 7$
Zj, jj	$0.071 \pm 0.703 \pm 0.014$	$35 \pm 19 \pm 9$	$25 \pm 14 \pm 5$	$7.3 \pm 7.2 \pm 1.4$	$8.2 \pm 8.1 \pm 1.6$
$Z(\ell\ell)\gamma$	$8.9 \pm 0.6 \pm 4.3$	$11.0 \pm 0.8 \pm 4.4$	$6.8 \pm 0.6 \pm 3.0$	$2.0 \pm 0.3 \pm 0.9$	$2.1 \pm 0.4 \pm 1.4$
Total	$654 \pm 8 \pm 202$	$1400 \pm 30 \pm 500$	$830 \pm 30 \pm 290$	$222 \pm 10 \pm 88$	$348 \pm 14 \pm 118$
Data	720 ± 30		830 ± 30	257 ± 16	

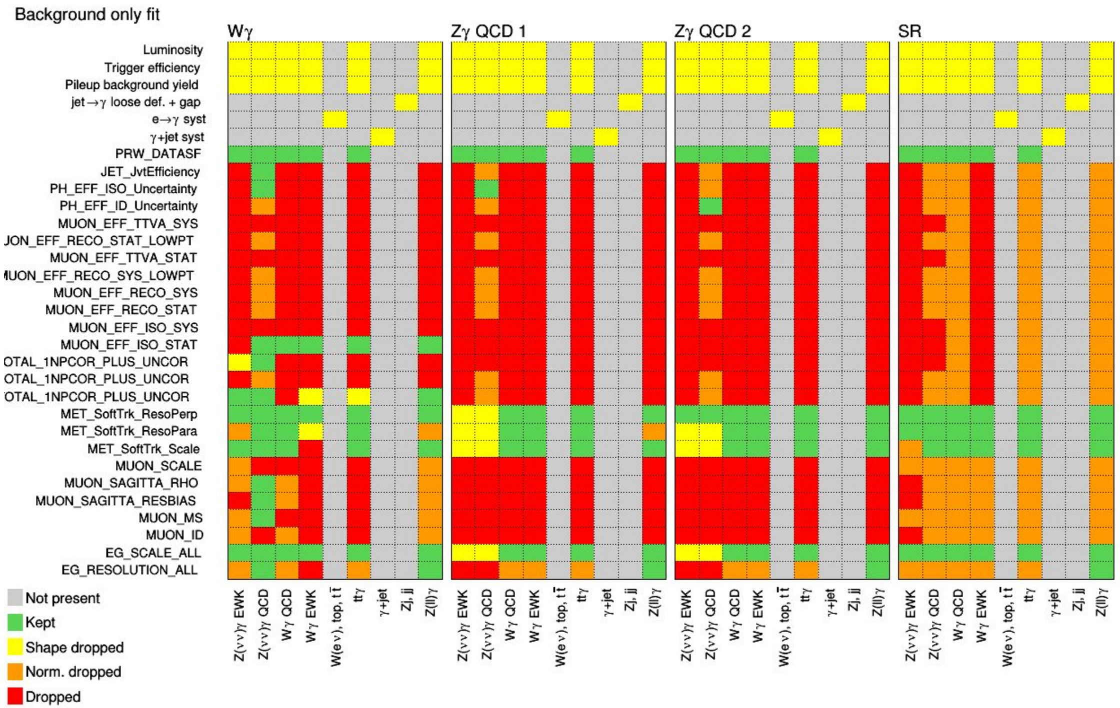
Table 16: Event yields for the signal and all of the background processes considered in this analysis before the normalization of the $Z\gamma$, $W\gamma$ and $tt\gamma$ background in the background only fit described in Section 12. The yields are presented in the regions described in Section 6.7. The uncertainty of the expected consists of statistic and systematic uncertainties. For the moment data in the signal region (and thus the $Z\gamma$ inclusive region) is blinded.

Systematic uncertainties pruning. Experimental 1/2



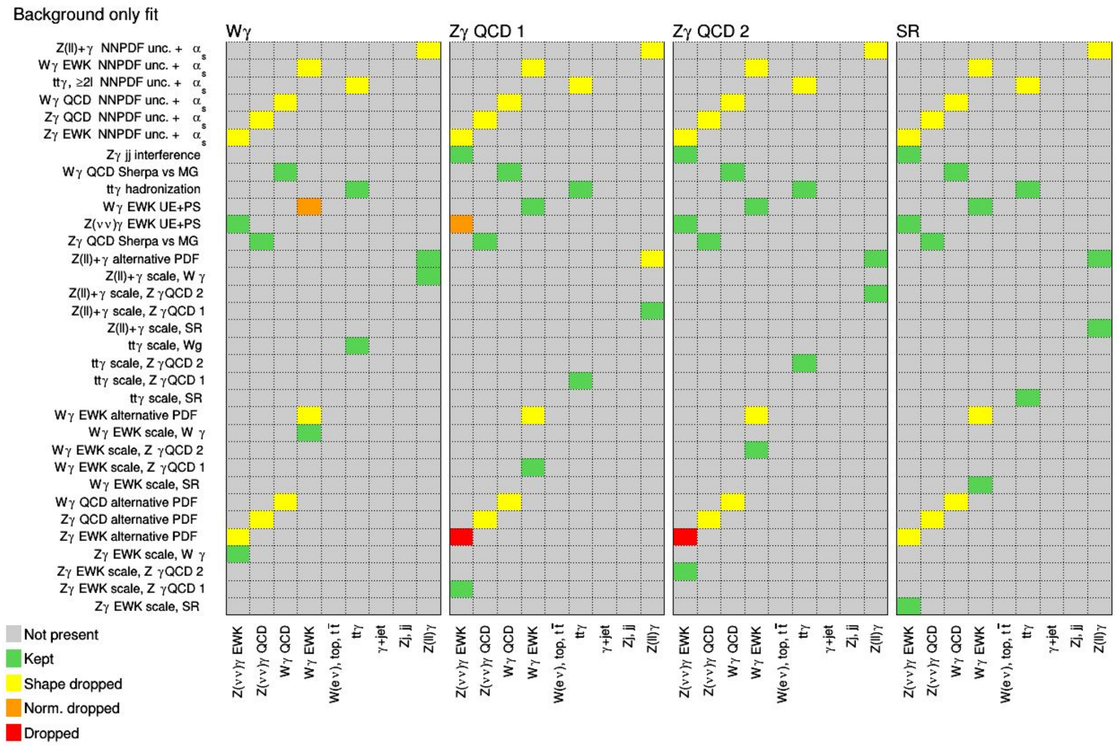
- The normalization part is dropped if the total effect on the event yield is < 1%
- The shape part is dropped if there’s no bins with effect on the event yield is > 1%

Systematic uncertainties pruning. Experimental 2/2



- The normalization part is dropped if the total effect on the event yield is < 1%
- The shape part is dropped if there’s no bins with effect on the event yield is > 1%

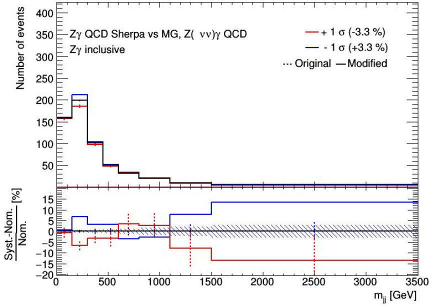
Systematic uncertainties pruning. Theoretical

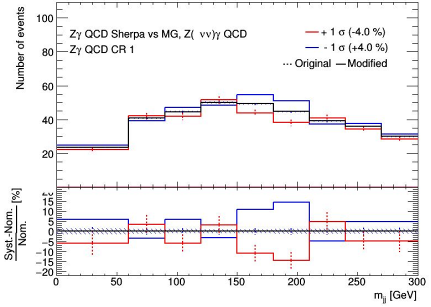


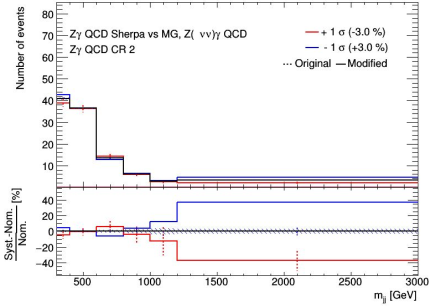
- The normalization part is dropped if the total effect on the event yield is < 1%
- The shape part is dropped if there's no bins with effect on the event yield is > 1%

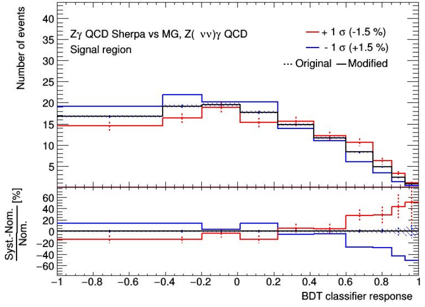
Wg QCD modelling uncertainty

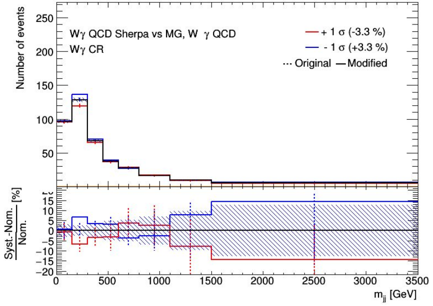
- Since there are no alternative samples for WγQCD, the modelling uncertainty was created by taking relative uncertainties of Z(vv)γ QCD Sherpa/Madgraph comparison in every region.
- Zγ inclusive region is used to model the uncertainty in the Wγregion, since Z(vv)γ QCD has low statistics and fluctuations as high as 1000% in Wγregion

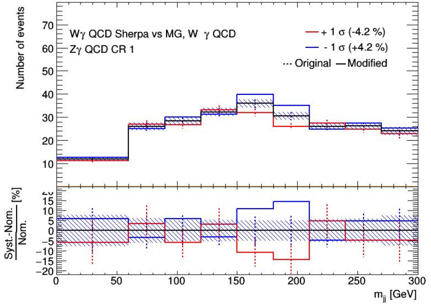


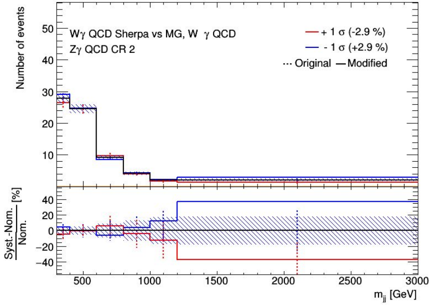


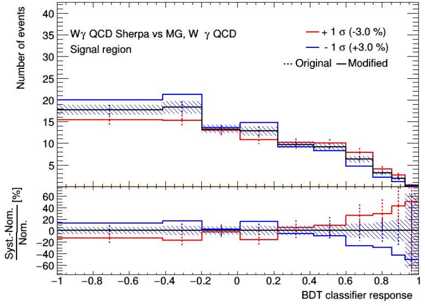








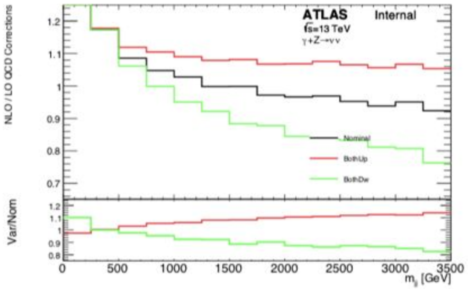




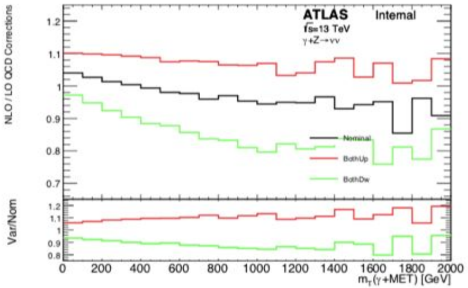
Electroweak $V\gamma$ + jets theory uncertainties

- NLO QCD corrections are about 2–3%, with 3–10% scale variations
- No strong dependence on $m_{\mathrm{T}}(\gamma, E_{\mathrm{T}}^{\mathrm{miss}})$

vs. m_{jj}



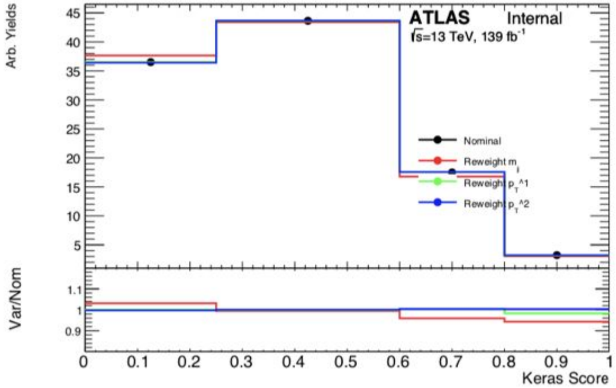
vs. $m_{\mathrm{T}}(\gamma, E_{\mathrm{T}}^{\mathrm{miss}})$



September 11, 2020 Lacey Rainbolt, VBF + MET + γ 11 / 19

Electroweak $V\gamma$ + jets theory uncertainties

- Applied NLO QCD corrections to m_{jj}, leading jet p_{T}, and subleading jet p_{T}
- Only the reweighting in m_{jj} affects the shape of the DNN score distribution



Smoothing Added

scale is different!

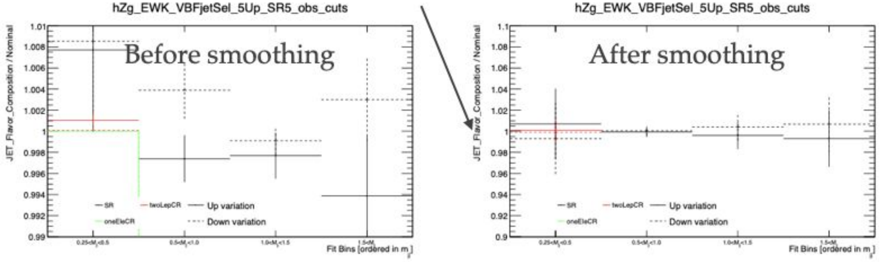


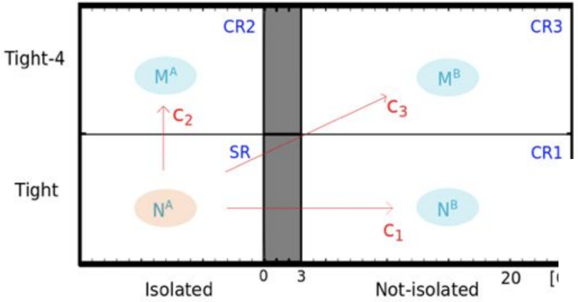
Figure 98: An example of the shape of the uncertainties before (left) and after (right) smoothing is shown for JET_Flavor_Composition. For the after smoothing, the two electron and two muon are the same shape and the same for the one electron and one muon.

- Smoothed using parabolic smoothing with 1 peak. Done separately for the mjj, mT and DNN inputs
- Smoothed 12 systematics uncertainties, which were problematic in the fit
  - Set a minimal list because smoothing can remove real information of the systematic uncertainties

X

Jets faking Photons

Data driven ABCD method to estimate how often a jets gets mis-reconstructed as a photon



Isolated = TopoEtCone40 – 0.022 pT – 2.45 GeV < 0 and ptcone20/pγT <

Model assumptions:

- Correlation of Tightness ID and Isolation negligible for the fake photons background
- Negligible number of photon candidates in the three CR

Both hypotheses relaxed to account for possible correlation and leakage from SR (respectively R_{MC} and $c_{1,2,3}$, extracted from MC)

SR photon Purity →

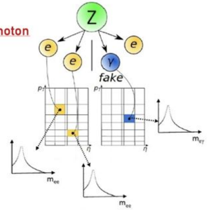
$$P = \frac{(M^B + N^A c_3 - N^B c_2 R_{MC} - M^A c_1 R_{MC})}{2N^A(c_1 c_2 R_{MC} - c_3)} \cdot \left(-1 + \sqrt{1 + \frac{4(c_1 c_2 R_{MC} - c_3)(N^A M^B - N^B M^A R_{MC})}{(M^B + N^A c_3 - N^B c_2 R_{MC} - M^A c_1 R_{MC})^2}}\right)$$

11

Electrons faking Photons – 1/2

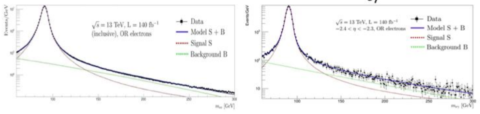
Fake signal given e.g. from $W(\to e\nu) + jets$ Where the e is reconstructed as a photon

Probability of an electron to be reconstructed as a photon → $F_{e\to\gamma} \equiv \frac{\epsilon(e^{true} \to \gamma^{reco})\epsilon_\gamma}{\epsilon(e^{true} \to e^{reco})\epsilon_e}$



- $Z(\to ee)$ data sample: all e and γ from the ee/eγ/γγ pairs are classified in η and pT bins and invariant mass spectra are built

$F_{e\to\gamma} = \frac{N_{e\gamma}}{2N_{ee}}$, → Numbers estimated by integrating between $\pm 3\sigma$ and subtracting background



← DSCB + e^{-ax-bx^2}

Milano group study

12

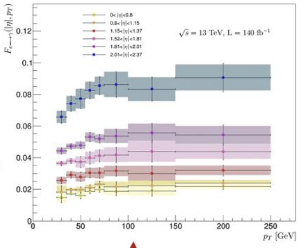
Electrons faking Photons – 2/2

Three major uncertainty sources:

- Mass Range: estimated as max absolute difference between nominal and integrating between $\pm 2\sigma$ / $\pm 4\sigma$ (~1 – 4%)
- Background subtraction: estimated as max absolute difference between nominal and integrating w/o subtracting background (~3 – 8%)
- Bias in photon reco energy: the photon energy is varied up and down by 1.5% (estimated from the difference between the ee and eγ m_{ex} peak positions) (~8 – 10%)

	Electron fakes	Stat. for number of events in probe-e CR	Stat. for fake rate	Syst. for fake rate	Total unc.
SR	27.4	1.05	2.0	5.0	5.6
1ℓ CR	0.92	0.16	0.13	0.1	0.09
2ℓ CR	0.0	0.0	0.0	0.0	0.0

Table 26: Number of electrons faking photons estimated in the SR and its CRs for full Run-2 data. The uncertainty is expressed in three terms: the first term is the statistical uncertainty related to the number of events found in the probe-e CR; the second and third terms are the statistical and systematic uncertainties related to the electron fake rate, respectively. The total uncertainty, calculated as the combination of the three uncertainties listed above, is also shown.



Inclusive in m_{jj}

13

Updating the Central Value for Vg Strong

- The NLO EWK+QCD corrections are available as on-the-fly weights in the Vg ntuples.
- Frank says that it is the most precise estimate of Vg strong that is currently available, so we should be using it.
- Practically it makes no change. If we are updating the dphijj cut, then we might as well do the too.

	Bin 1	Bin 2	Bin 3	Bin 4
SR Total	-1.5%	-4.1%	-3.0%	-4.2%
SR W	-1.4%	-4.8%	-3.1%	-3.1%
$W \to e\nu$ VR	-1.4%	-7.7%	-2.8%	-2.7%
$W \to \mu\nu$ VR	-1.5%	-2.1%	-2.5%	-7.4%
$Z \to ee$ VR	-1.3%	-3.3%	-3.6%	-2.7%
$Z \to \mu\mu$ VR	-1.3%	-1.9%	-0.8%	-4.6%

Changes relative to the NLO QCD normalization

X

Sherpa vs MG Vg strong

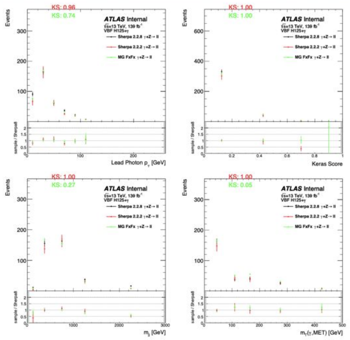


Figure 60: The $Z(\to \ell\ell)\gamma$+jets strong distributions of photon p_T, m_{jj}, Keras score, and $m_T(\gamma, E_T^{miss})$ with 139 fb^{-1} of simulation in the Loose $Z\gamma$+jets validation region, defined in Section 4.5.7. The predictions from Sherpa 2.2.8, 2.2.2, and Madgraph are compared at reconstruction level. The simulation is normalised to cross-sections times luminosity, and the uncertainties are statistical only.

Zγjj/Wγjj QCD scale uncertainties

	Scale
Zγjj EWK	v
Zγjj QCD	
Wγ EWK	v
Wγ QCD	
ttγ	v
Z(ll)+γ	v

Scale:
4 NPs per sample for:
- Wγ CR
- ZγQCD CR1
- ZγQCD CR2
- SR

Zγjj QCD and **Wγ QCD scale** uncertainties are omitted since they only affect the normalization and tend to double the designated normalization coefficients in the likelihood model.

However the EXOT analysis does have those uncertainties.

When we try to include **Zγjj QCD** and **Wγ QCD scale** uncertainties with the same 4 NPs scheme we got unnatural NP pulls, μ(ZγQCD) shift and μ(ZγQCD) error increase.

Trying to combat those issues we've also tested the 1 NP scheme for **all** of the scale uncertainties (i.e. all of the regions are now "correlated").

Zγjj/Wγjj QCD scale uncertainties

No Zγ/Wγ scale unc.
Asimov data fit

μ(ZγEWK)	$1.04^{+0.27}_{-0.25}$ (stat)$^{+0.24}_{-0.17}$ (syst)
μ(ZγQCD)	1.08 ± 0.08 (stat)$^{+0.16}_{-0.17}$ (syst)
μ(Wγ)	1.09 ± 0.04 (stat)$^{+0.20}_{-0.14}$ (syst)

Expected median significance:
3.8 σ

With Zγ/Wγ scale unc., 4 NP
Asimov data fit

μ(ZγEWK)	$1.04^{+0.28}_{-0.25}$ (stat)$^{+0.26}_{-0.20}$ (syst)
μ(ZγQCD)	1.30 ± 0.08 (stat)$^{+0.49}_{-0.33}$ (syst)
μ(Wγ)	1.09 ± 0.04 (stat)$^{+0.19}_{-0.13}$ (syst)

Expected median significance:
3.6 σ