Search of Higgs production with decay into the Z boson and a photon

Proklova Nadezda

MEPhI@ATLAS meeting, 14 April 2020

About me and my projects

Joint PhD student at MEPhI & Radboud University

Projects:

1) TRT software

- Qualification: Developments in the TRT dE/dx time-over-threshold algorithm for particle identification [link]
- Current project: MC Rt/T0 calibration studies

2) E/gamma performance

- photon isolation
- photon identification

3) Physics analysis: H→Zγ search

- thesis in progress

Extra personpower:

- Natalia Zubova (master student), involved in performance studies

Photon performance:
Photon ID

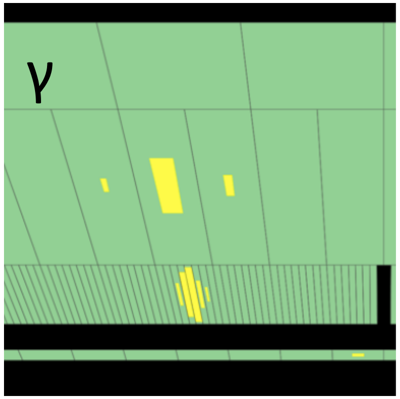
Photon reconstruction and identification

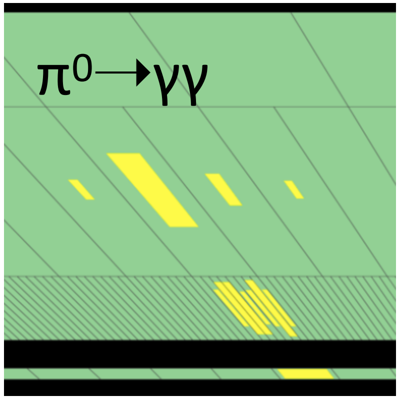
Prompt photons:

- Direct photon from the hard scattering process
- Fragmentation photon from a parton (less isolated)

Background:

- jets with large EM fraction (e.g. π0, η) that can fake photons
- Electron with similar interaction in calorimeter





ID: 9 discriminating variables (DVs) based on energy in cells of ECAL and leakage in hadronic calorimeter HCAL

Tight ID

- **Is optimised as function of E_T, depends on η and conversion status** (some details about optimisation are going to be provided in the analysis section)
- measurements are done in **data** :

1) MC shower shapes are shifted to data so that their means match the data means
2) residual differences in MC has to be corrected later to data

loose ID

used by triggers or as background control region

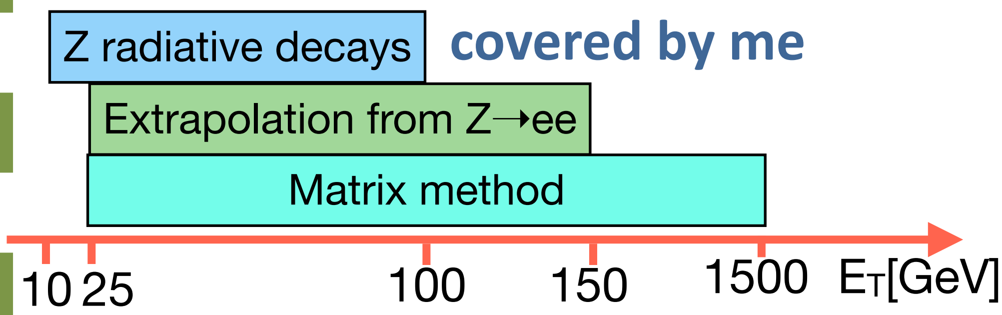
tight ID

- tighter cuts on DVs used by loose ID
- used for offline analysis

ATL-PHYS-PUB-2016-014

Photon reconstruction and identification

Tight ID efficiency is measured with **3 methods** with different ET ranges



RadZ: low E_T range, but pure photon sample (P=95-99%)

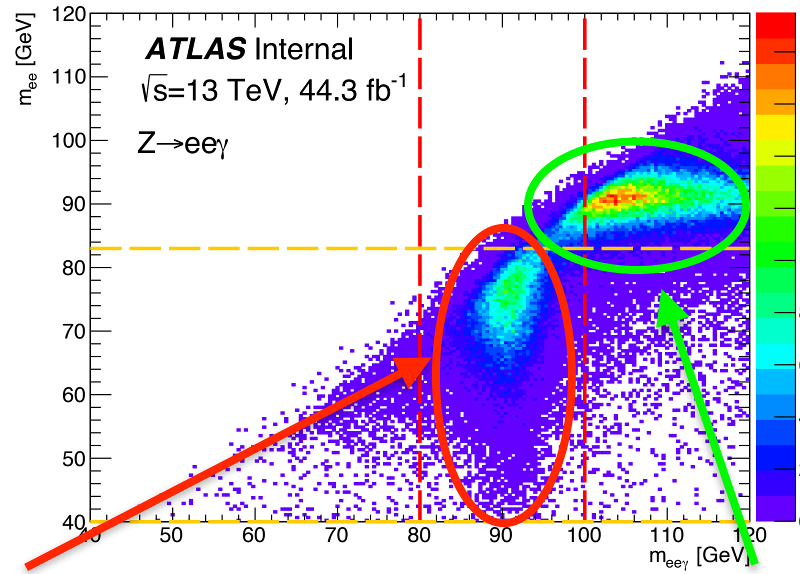
For FSR selection we use cuts on M_{ll} and $M_{ll\gamma}$

RadZ selection:

Leptons: Et(el,mu) > 10 GeV, $|\eta_{el}| < 2.47$, $|\eta_{mu}| < 2.7$, Loose isolation

Photons: Et > 10 GeV, Loose OR Tight iso, deltaR(el/mu, γ) > 0.4, $|\eta| < 2.37$

Event selection: 40 < Mll < 83 GeV; 80 < Mllγ < 100 GeV, trigger matching



Signal: FSR photons ($m_{ll\gamma} \approx m_Z$, $m_{ll} < m_Z$)

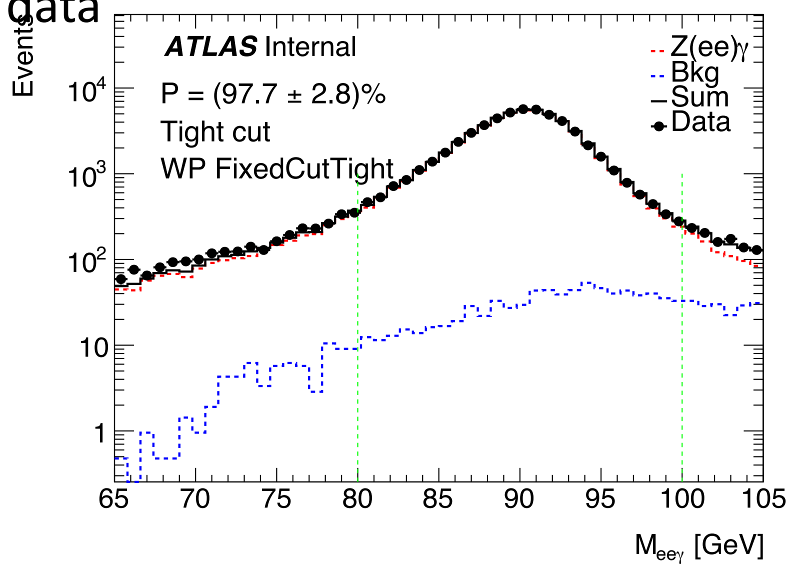
Bkg: ISR photons, jets ($m_{ll\gamma} > m_Z$, $m_{ll} \approx m_Z$)

Methods of background estimation

Purity estimation with a template fit

- Number of background events could be estimated in data from the template fit

Signal (Zllg) PDF + background (Z+jets) PDF = fit to data



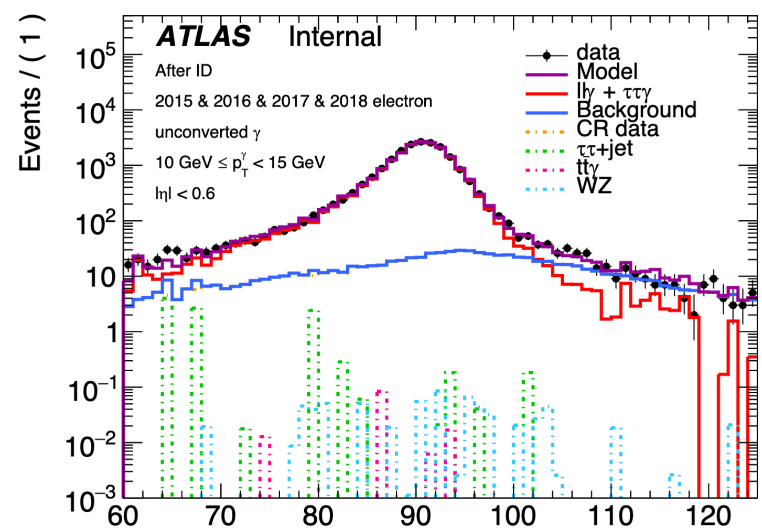
Purity: $P = \dfrac{N_{sig}}{N_{sig} + N_{bkg}}$

- Efficiency is corrected by doing background subtraction:

$$\varepsilon = \frac{N_{probes,pID} - N_{bkg,pID}}{N_{probes} - N_{bkg}} = \frac{N_{sig,pID}}{N_{sig}}$$

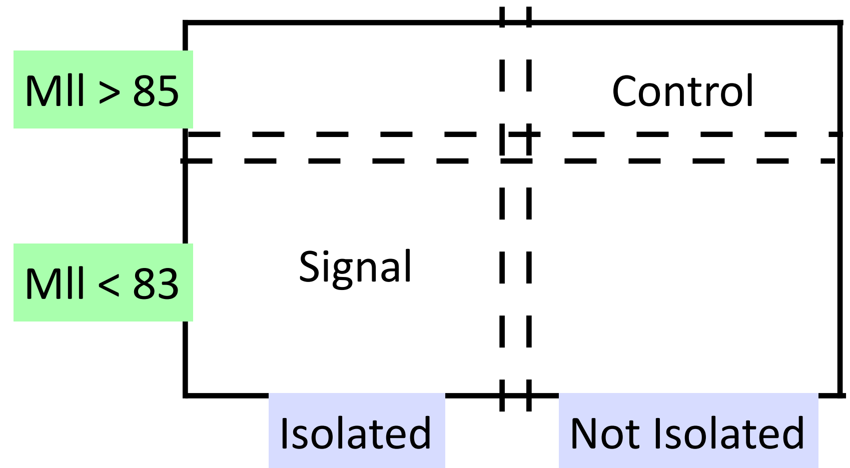
Method is not used anymore as a nominal method (only for a cross-check):

- allows to correct data only up to ~25 GeV (limited statistics)
- Z+jets only bkg doesn't describe tails of mass distribution
- Is replaced by template fit method with additional bkg sources

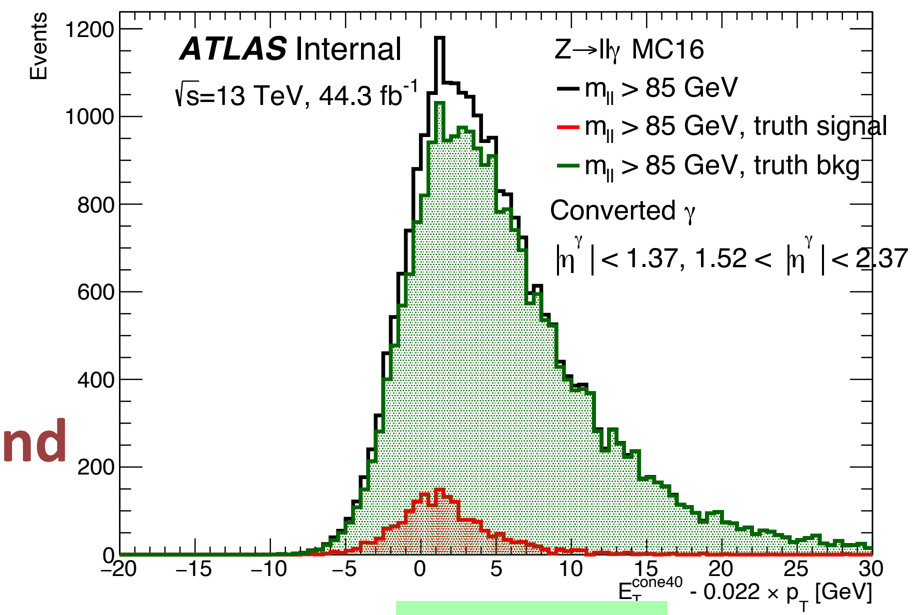


Methods of background estimation

2d sideband method



- Use of not-isolated photons: reversing one of the isolation variables
- Reversing Mll cut: for Mll > 85 GeV = almost all photons should be jets->γ
- N_{bkg} can be estimated by normalizing control bkg shape to the tail of signal (isolated) distribution
- Track isolation is pre-applied



Mll > 85 GeV

Method allows to estimate background contamination up to ~40 GeV

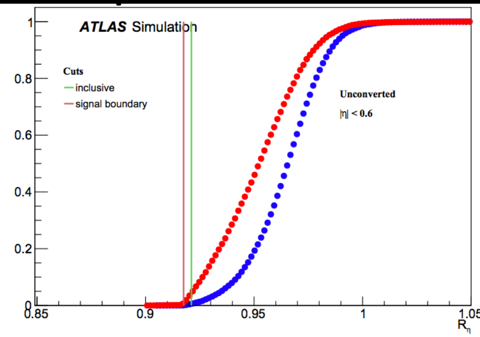
more on methods validation in backup slides

Photon ID efficiency vs pT

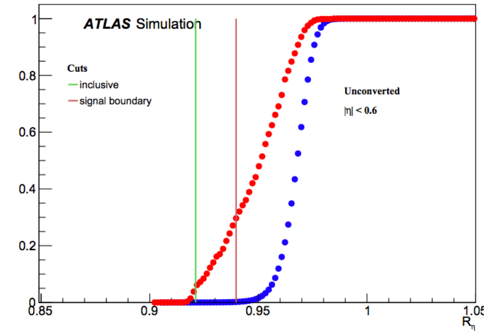
All shower shapes are pT -dependent

Tight ID menu was optimised as pT-dependent in 2018:

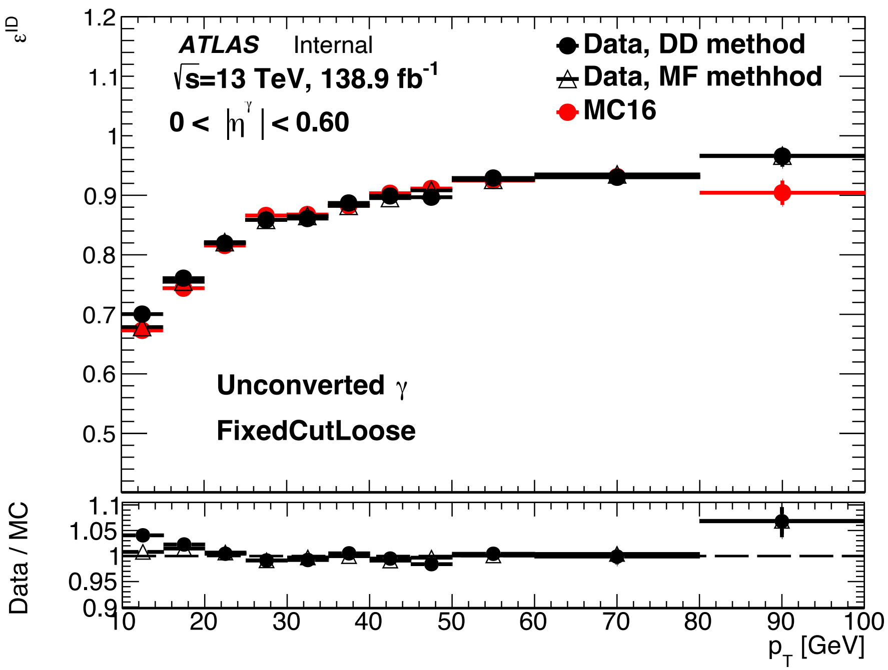
Benefits:
- x1.5-2 improvement in BKG rejection at pT > 100 GeV
- increased signal efficiency at 10 < pT < 25 GeV (+ ~10-20%)

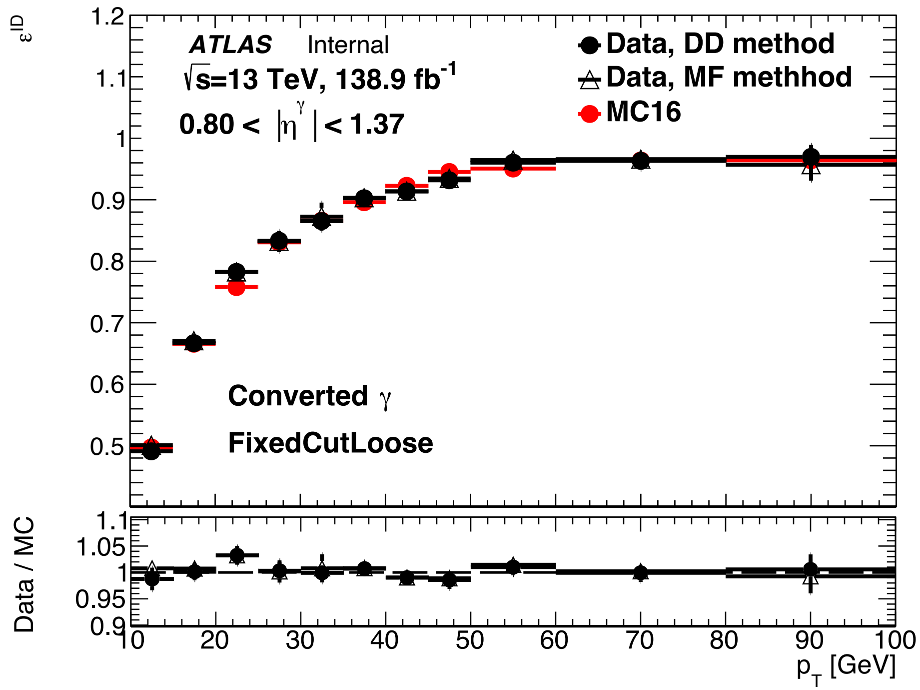


$25 < p_T < 30$



$80 < p_T < 100$





Monte-Carlo and Data efficiencies agrees well, SFs ~ 1 for different pT/eta bins

Systematic uncertainties

1) 2 methods of background estimation:
   - take the difference between isolation based method to mass-based method

3) uncertainty on FF variations - tight ID is varied for DD method

4) uncertainty of DD method: closure method

5) uncertainty of Mass fit method:
   - Mass fit range variation
   - Signal leakage

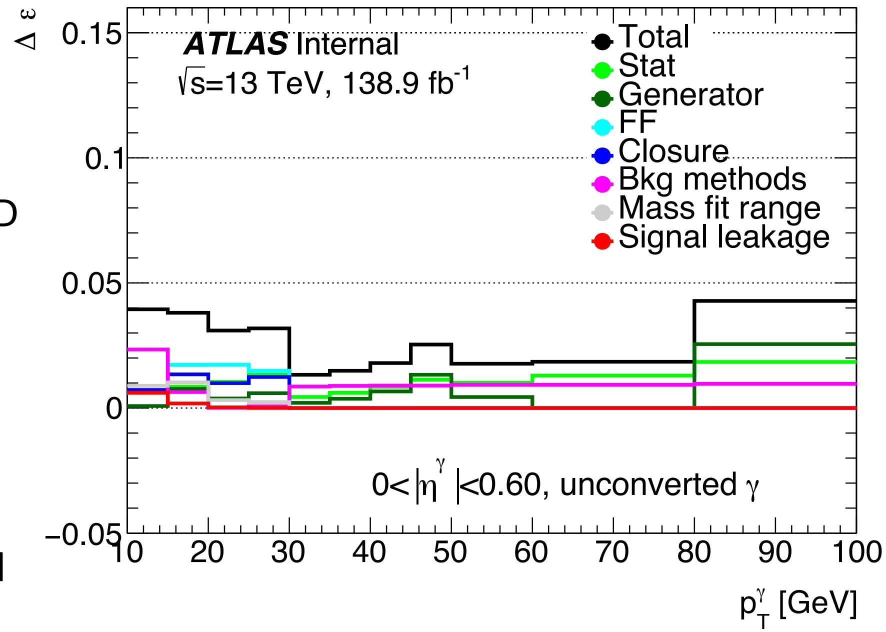
6) MC generator
   - difference in MC efficiency between Sherpa and PowhegPythia generators

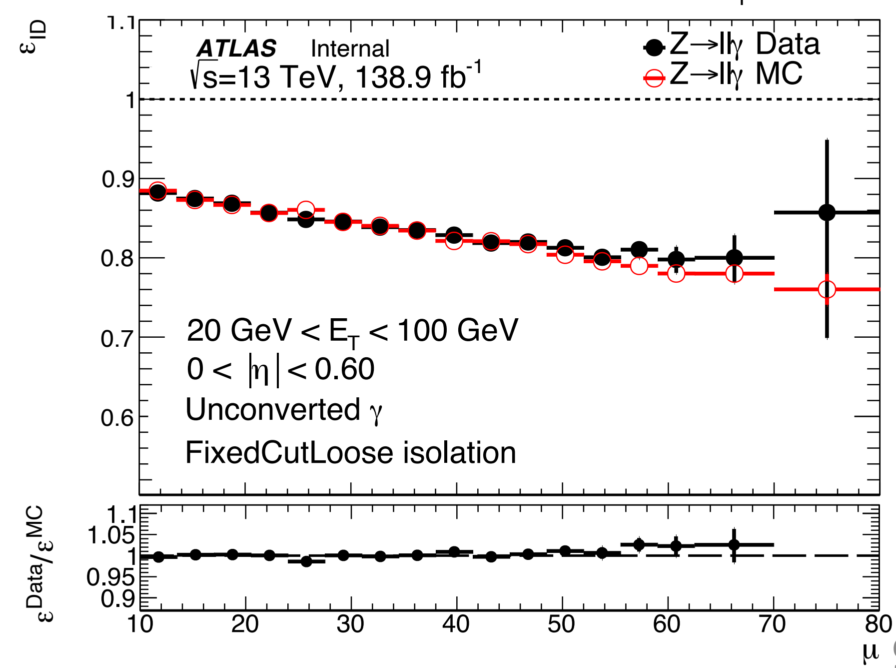
Extra checks:

1) compatibility between different isolation WPs
2) pileup dependence

MC describes Data well except for very high pileup > 60

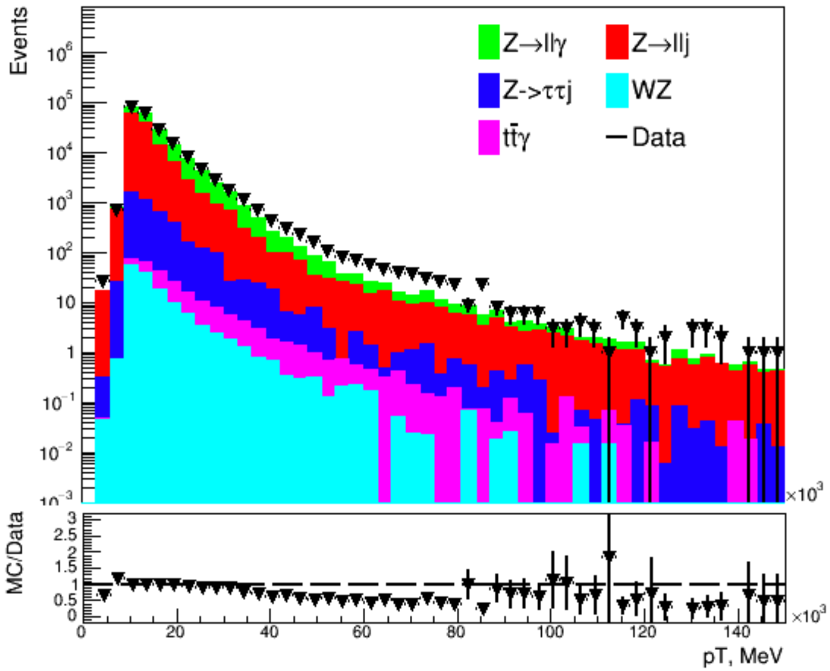
Visible decrease of phID with pileup ~10-20%

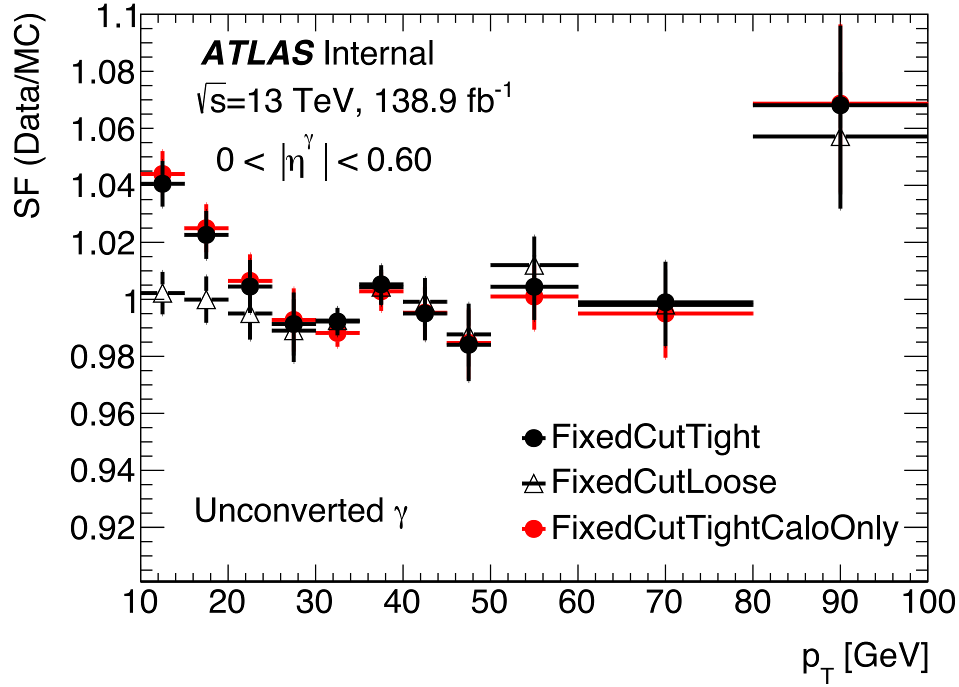




Known problems

N. Zubova





Known MC backgrounds don't describe data at high pT > 40 GeV (difference is ~10-50%) for anti-tight ID and anti or no isolation

- General correlation between photon ID and isolation (both use calorimeter information)
- isolation cone dependency at low pT (ie cone40/cone20), noticed the effect only with full Run2 data, rel.21

Photon performance: Photon isolation

Photon isolation

Photon isolation suppress further backgrounds after photonID

- Energy flow around fakes is larger than for prompt photons

Calorimeter isolation:

$\Sigma E_T/E_T < 0.065/0.022$, the sum is over all calo clusters in a cone $\Delta R < 0.2$ / 0.4

photon energy

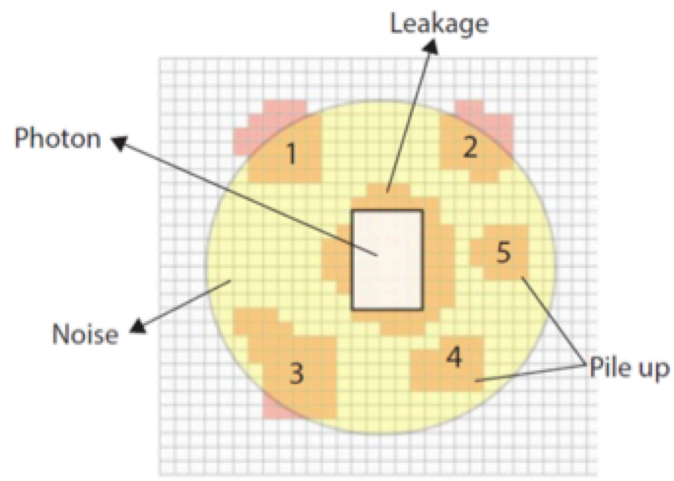
$$E^T_{iso,corr} = \sum_{i,\Delta R<0.4}^{clusters} E^T_{i,raw} - \left[E^T_{core} - E^T_{leakage}(p_T,\eta)\right] - \left[E^T_{pileup}(\eta)\right]$$

Pileup correction

Track isolation:

$\Sigma p_T/E_T < 0.05$, loose vertex association, sum is over all tracks (with $p_T > 1$ GeV) in a cone $\Delta R = 0.2$

Calorimeter isolation



Fakes: neutral hadrons in jets decaying into two photons

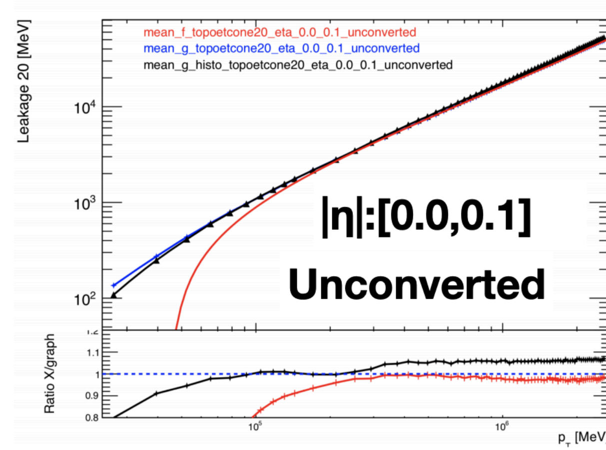
- Clusters around prompt (isolated) photons are coming from pileup.
- Clusters around fakes (non isolated photon candidates) are coming from other objects in the jet.

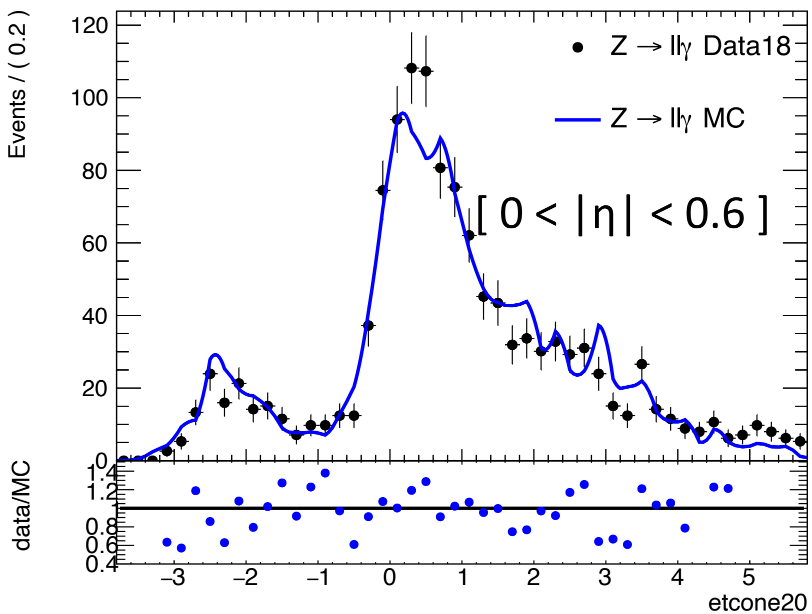
Photon calorimeter isolation

Photon energy component: fixed size window + energy leaking outside the window

Photon calo leakage corrections: corrects the energy leaking outside the fixed mask

- parametrization of the isolation energy distribution with ET for each of the |η| bins using Crystal Ball functions.
- fit the model with a pol-2 to obtain an ET-dependent correction in |η| bins and conversion
- Converted photons divided in two categories: second peak observed in the isolation distribution for the category called "trouble".





Converted trouble category:

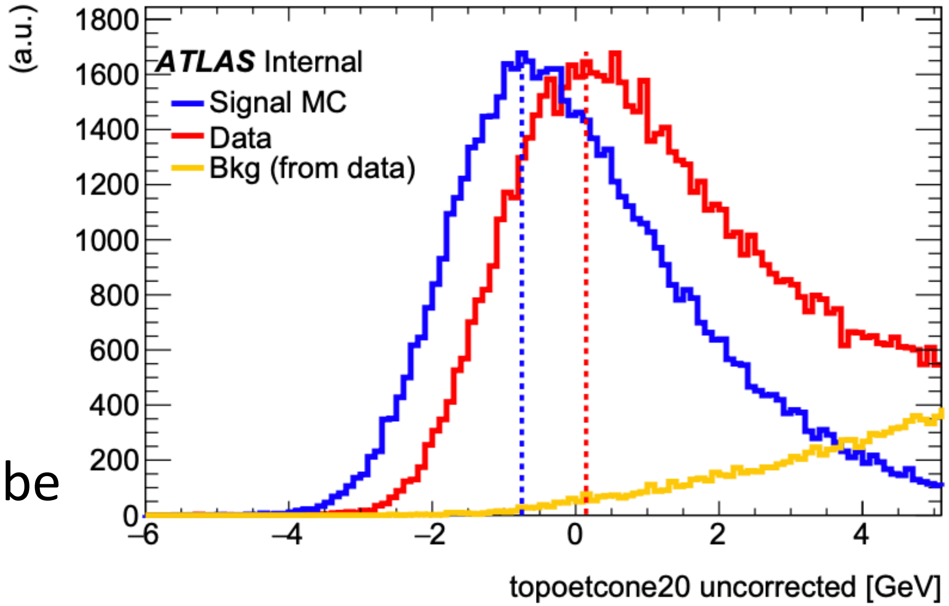
- Conversion type 3: two tracks, both with Si hits
- Asymmetric conversion: $E_{T,e_{subl}}/E_{T,e_{lead}} < 0.3$
- 30% of converted photons are trouble with increasing ET (~ 10 < ET < 25 GeV)
- Structures observed as a function of |η| (barrel is affected)
- Around ~5-10% of the photons are classified in the "trouble" category.

Photon calorimeter isolation

Isolation Working Points efficiency: measurement are done in data, and MC is corrected according to it

1) MC shower shapes are shifted to data so that their means match the data means (DD shifts)

2) residual differences in MC has to be corrected later to data with SFs



Both measurements are obtained in different |η|, pT and photon conversion status.

covered by me

$10 < E_T < 100$ GeV

SFs are measured for 3 isolation WPs (track+calo isolation)

$25 < E_T < 1000$ GeV

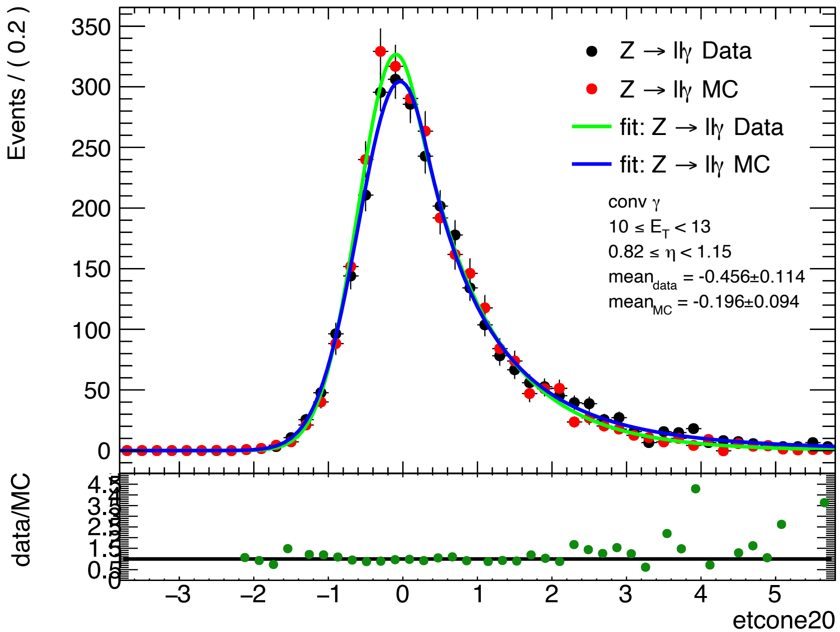
SF are measured with Single Photons, separately for track and calo isolations

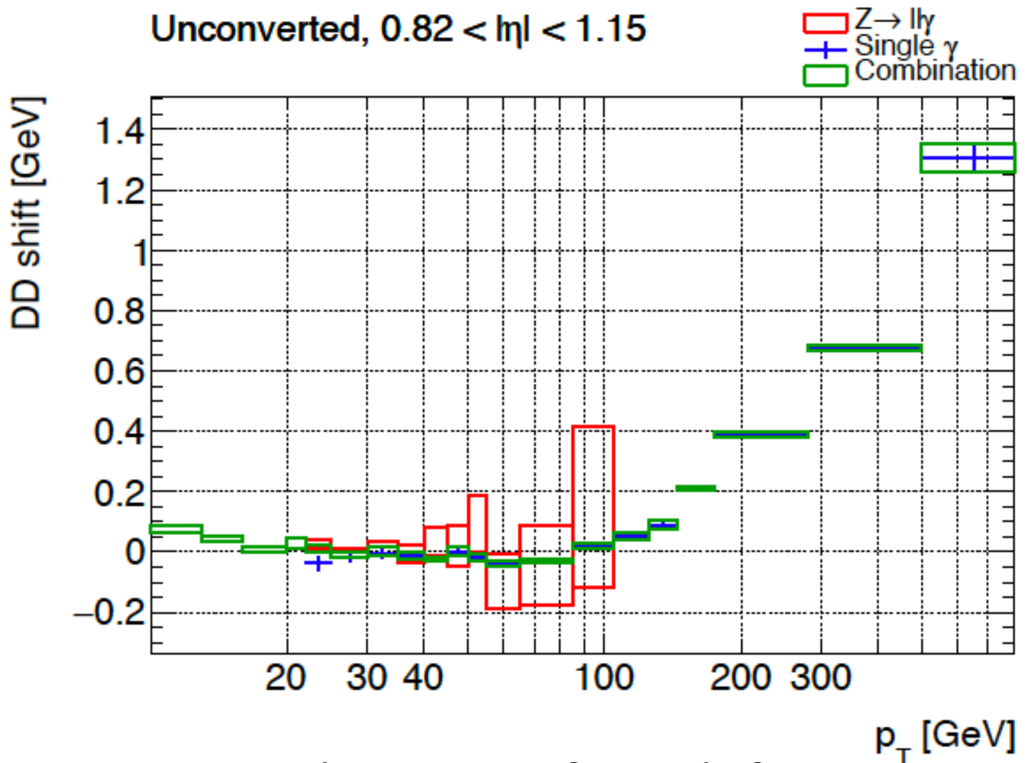
Photon calorimeter isolation

RadZ measurements: selection is described in phID

DD shifts are measured as $\mu_{Data} - \mu_{MC}$ of isolation distribution

N. Zubova





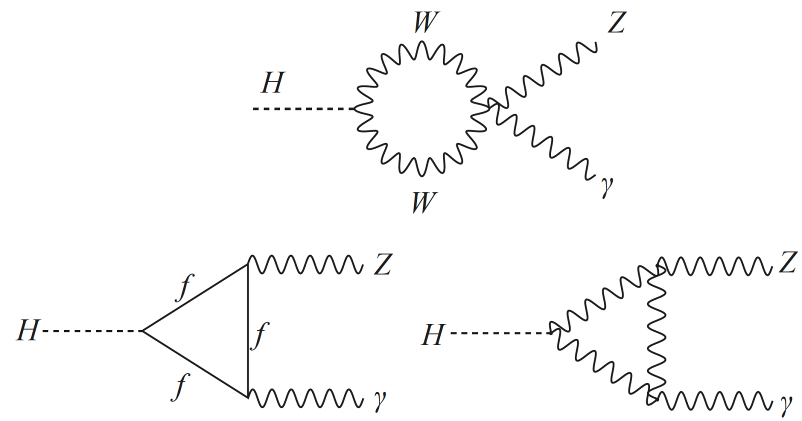
Fitting is done with asymmetric Crystal Ball function

Combination of DD shifts for the whole pT range

Higgs->Zγ search

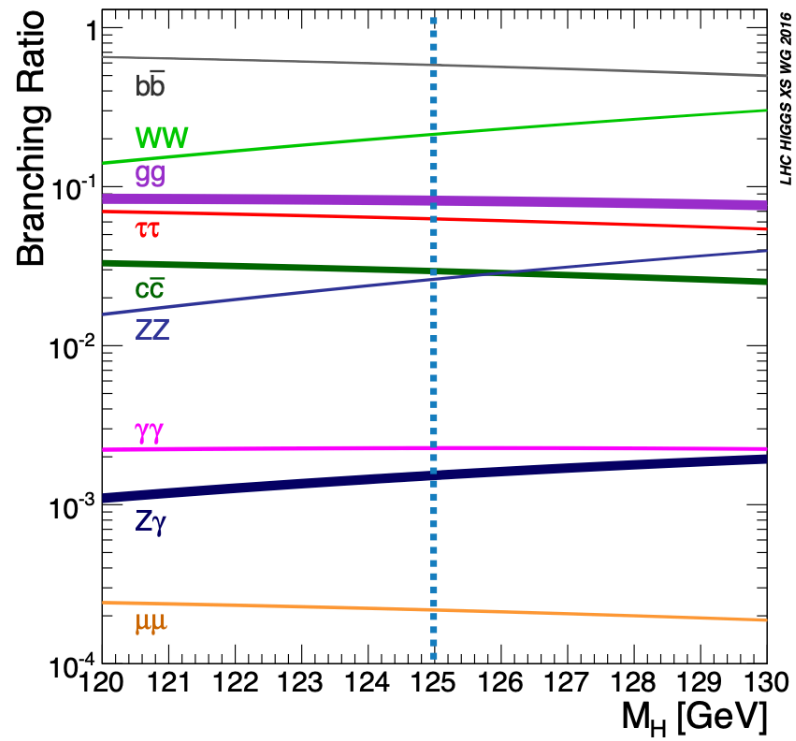
Motivation

- more difficult than Higgs→γγ
  - -(B[H→Zγ]×B[Z→ee/μμ]= ~10^{-4})
- -**BUT**: Small background → great sensitivity (large QCD component for γγ backgrounds)



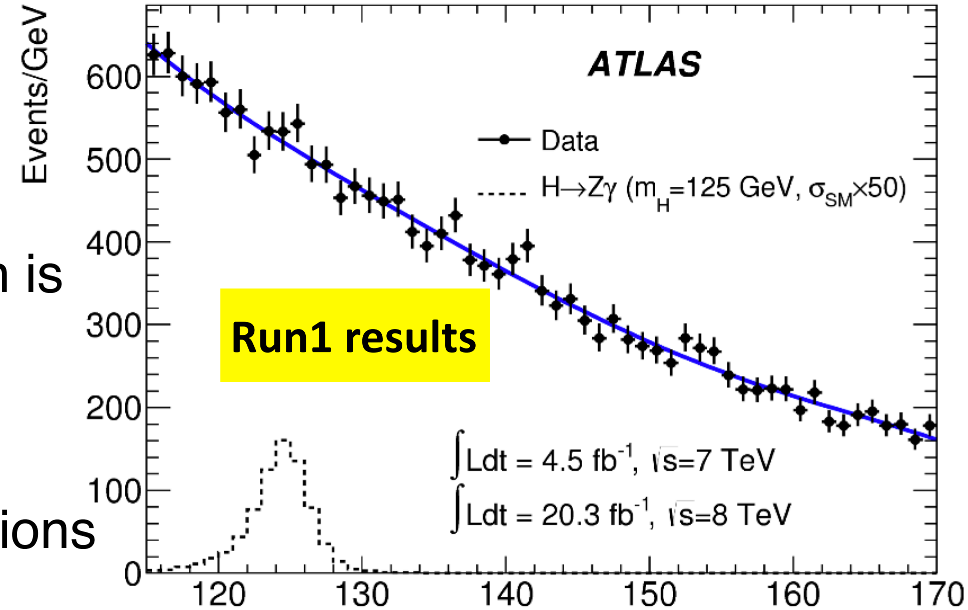
Something what can be checked:

- **B(H→γγ)/B(H→Zγ) may differentiate**
  - -In extended Higgs sector models [arxiv 1207.1065v2]
  - -With additional light charged electroweak particles [arxiv 1206.1082v3]

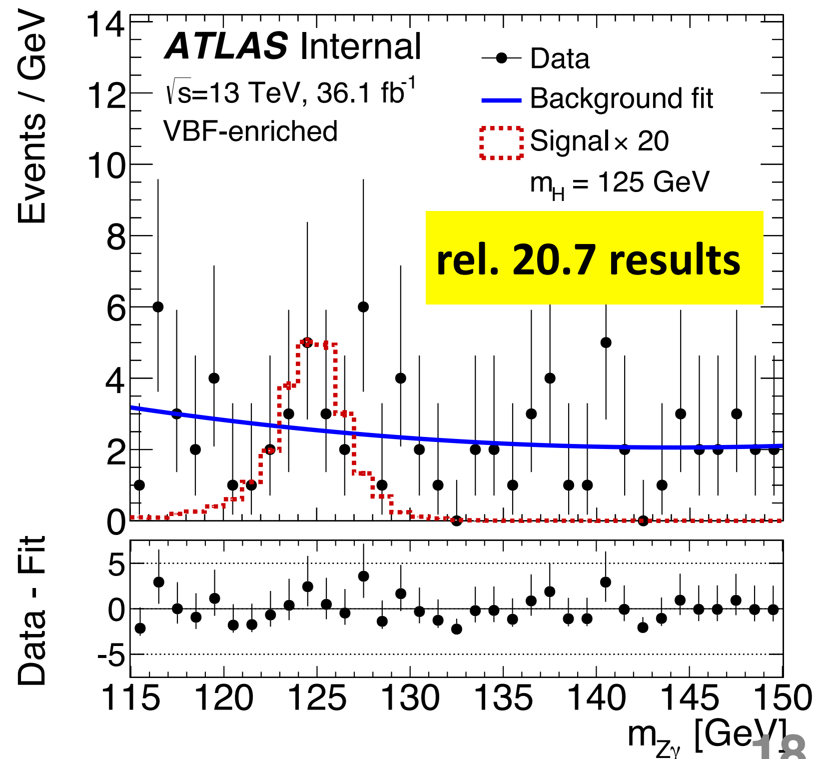


Previous results

Run1: [Phys. Lett. B 732 (2014) 8–27]
24.8 fb^{-1} of 7 + 8 TeV data
eeγ and μμγ final states, categorisation is based on kinematic properties
No significant excesses found
Expected (Observed) limits at mH = 125 GeV are: 9(11) x SM predictions



2015-2016 **36.1 fb^{-1}** analysis:
JHEP 10 (2017) 112
eeγ and μμγ final states, categorisation is based on kinematic properties + **MVA**
No significant excess over the expected background
(6.6(5.2) x SM @95% CL)



Current Analysis:
Re-optimized full Run2 H→Zγ with ~3.9 times data as last publication;

Higgs->Zγ samples

Full run2 data: 139 fb^{-1}

Higgs MC:

Process	Technique	PYTHIA8 version & tune	PDF set	QCD (gen.)	Normalisation
ggF	MiNLO [90–93] & NNLOPS [94, 95]	8.186, AZNLO [87]	NNPDF30 [73]	NNLO	NNNLO (QCD), NLO (EW) [43–52]
VBF	Powheg	8.186, AZNLO	NNPDF30	NLO	NNLO (QCD), NLO (EW) [53–55]
$q\bar{q} \rightarrow VH$	MiNLO [92, 96]	8.186, AZNLO	NNPDF30	NLO	NNLO (QCD), NLO (EW) [56–63]
$t\bar{t}H$	Powheg	8.230, A14 [85]	NNPDF23 [38]	NLO	NNLO (QCD), NLO (EW) [64–67]

New: H→μμ signal to test its contamination (up to 3% among all categories)

Background:

- **Sherpa Zγ** - main background (~80%): Full-sim in pT slices (selection/ categorization optimization), Fast-sim in mass slices and in 2 parton region (background templates)
- **Madgraph Electroweak Zγjj (VBS)**: non-negligible after VBF selection
- **Z+jet** (~20%, reducible): Data-driven method (non-tight photons)

Higgs->Zγ triggers, object selection

Triggers:

Single-lepton, dilepton triggers

candidates	channel	single/di-lepton	trigger name
2015 data	$Z(\to ee)\gamma$	single electron	HLT_e24_lhmedium_L1EM20VH
			HLT_e60_lhmedium, HLT_e120_lhloose
2015 data	$Z(\to ee)\gamma$	di-electron	HLT_2e12_lhloose_L12EM10VH
2016 data	$Z(\to ee)\gamma$	single electron	HLT_e26_lhtight_nod0_ivarloose
			HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
2016 data	$Z(\to ee)\gamma$	di-electron	HLT_2e17_lhvloose_nod0
2017-2018 data	$Z(\to ee)\gamma$	single electron	HLT_e26_lhtight_nod0_ivarloose
			HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
2017-2018 data	$Z(\to ee)\gamma$	di-electron	HLT_2e24_lhvloose_nod0
2015 data	$Z(\to \mu\mu)\gamma$	single muon	HLT_mu26_imedium, HLT_mu50
2015 data	$Z(\to \mu\mu)\gamma$	di-muon	HLT_mu22_mu8noL1
2016 data	$Z(\to \mu\mu)\gamma$	single muon	HLT_mu26_imedium
			HLT_mu26_ivarmedium, HLT_mu50
2016 data	$Z(\to \mu\mu)\gamma$	di-muon	HLT_mu22_mu8noL1
2017-2018 data	$Z(\to \mu\mu)\gamma$	single muon	HLT_mu26_ivarmedium, HLT_mu50
2017-2018 data	$Z(\to \mu\mu)\gamma$	di-muon	HLT_mu22_mu8noL1

Object selection:

~5% higher sig. eff. than medium in Z→ee

Lepton and Photon Preselection (before overlap removal)			
Cut	Electrons	Muons	Photons
p_T	> 10 GeV	> 10 GeV	> 10 GeV
$\|\eta\|$	$\|\eta\| < 2.47$	$\|\eta\| < 2.7$	$\|\eta\| < 2.37$
	exclude $1.37 < \|\eta\| < 1.52$	-	exclude $1.37 < \|\eta\| < 1.52$
$\|d_0\|/\sigma_{d_0}$	< 5	< 3	-
$z_0 \sin\theta$	< 0.5 mm	< 0.5 mm	-
Identification	Loose	Medium	Loose
Isolation	FCLoose	FCLoose	-

Jet object for VBF: Anti-kt, pT > 25 GeV, |eta| < 4.4, removed in a cone size of 0.2 in the 2 final leptons or in any photon

Higgs->Zγ event selection

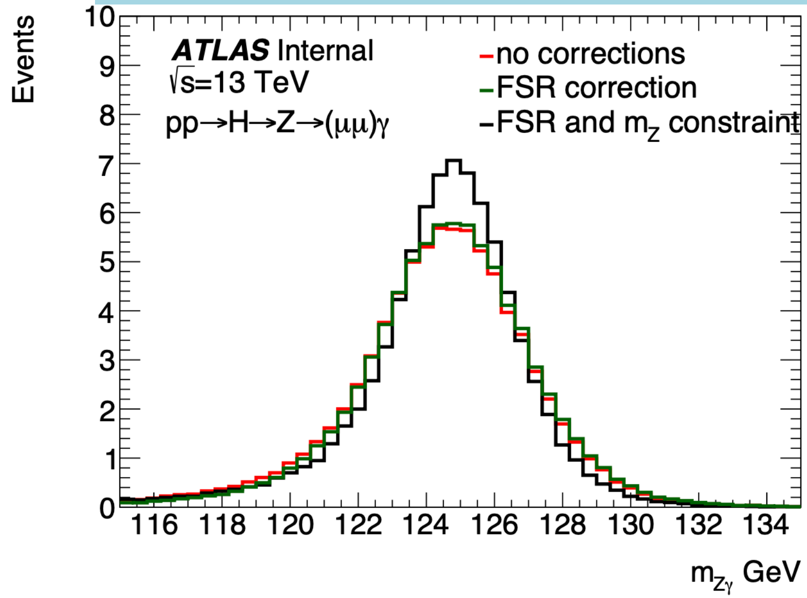
Overlap removal: remove photons in a cone size < 0.3 of each one in the di-lepton pair to build Z candidate; remove jets in a cone size < 0.2 of other objects

- **~2% higher sig. eff. than removal with all leptons, helps to account for increased number of photon fakes**

Z construction:

- Opposite charge lepton pair whose mass (m_{ll}) is closest to Z mass
- m_{ll} corrected by FSR correction and mass constraint
- 81.2-corrected mll-101.2 GeV
  - **~3% higher signal sensitivity than 15GeV Z mass window**

FSR reduce Res. by 3%
Mass constraint reduce Res. by 14%



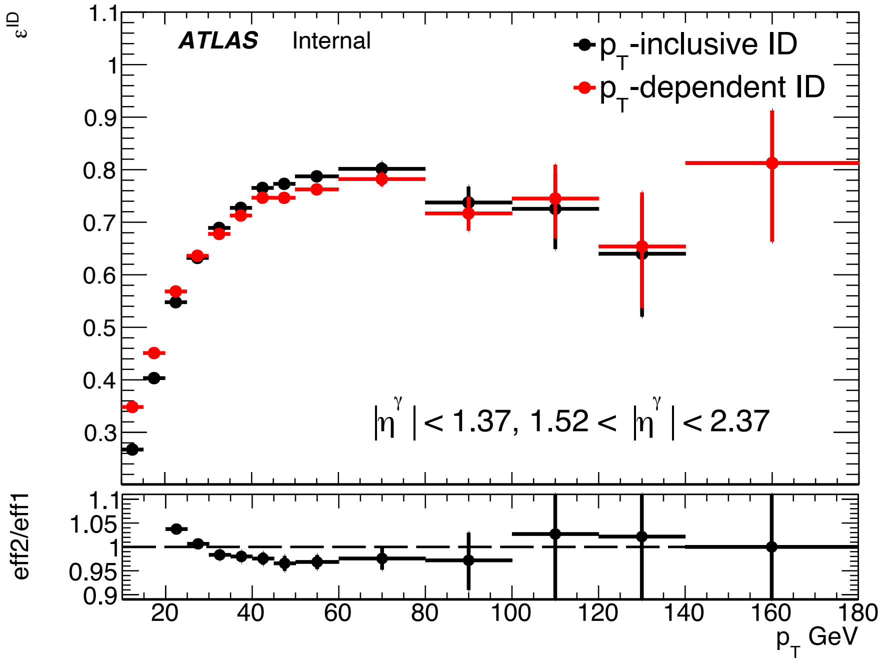
(b) $Z\gamma \rightarrow \mu\mu\gamma$

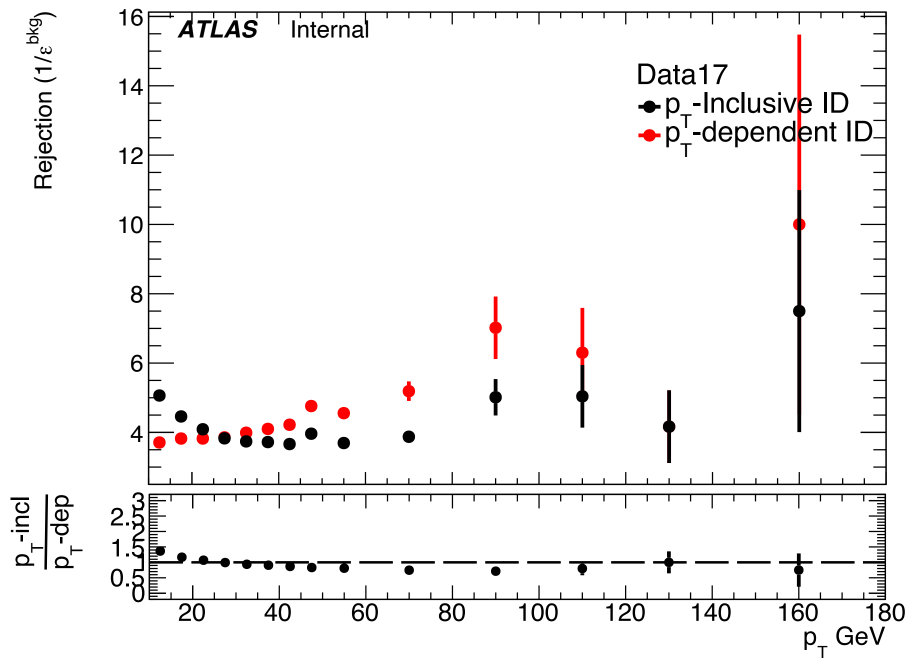
Photon final selections: photon pT / mZγ > 0.12, photon ID, Isolation

mZγ pre-selection: 105-160 GeV

Event selection: optimisation

- New version of tight ID menu was produced following feedback from H->Zy group:
  - increased signal efficiency at 10 < pT < 25 GeV
  - increased signal efficiency in middle pT range -> closer to the pT-inclusive dependency



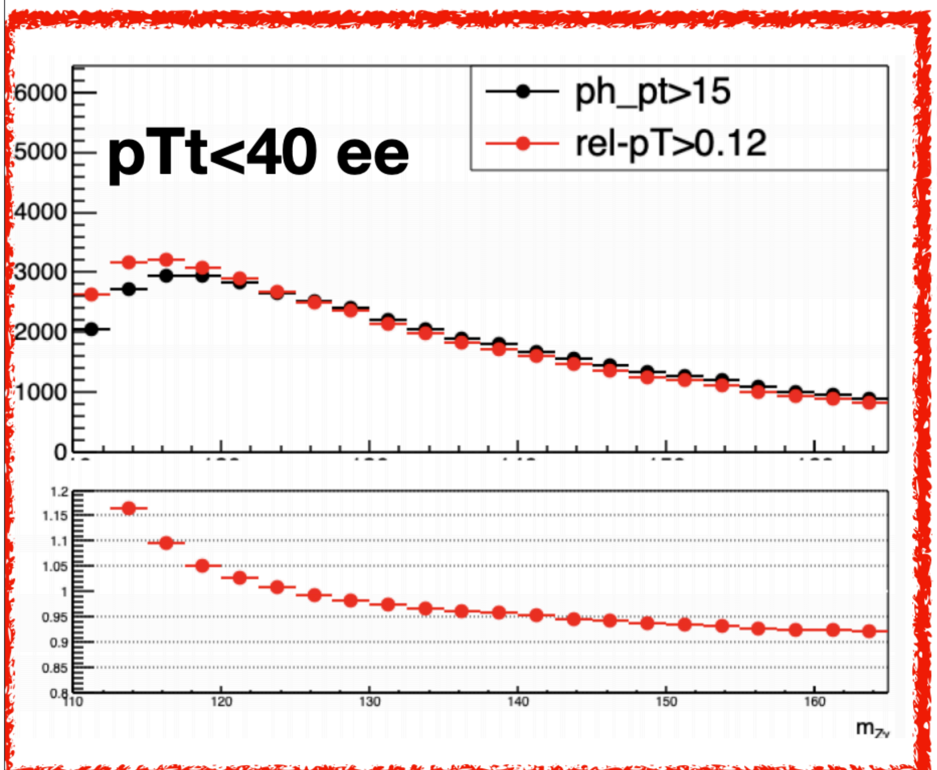


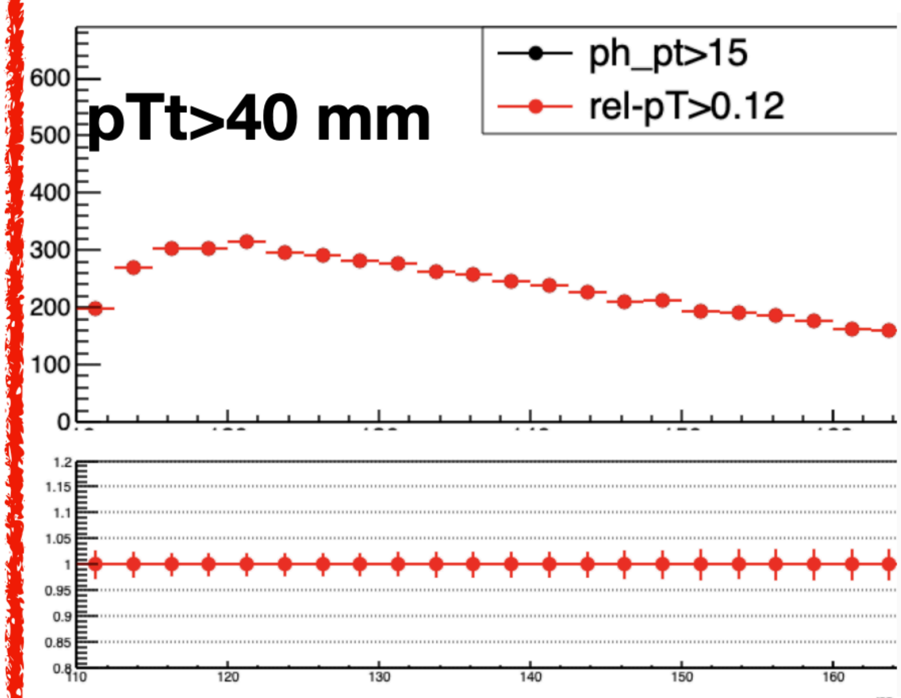
~20% higher efficiency at low-pT, ~4% improved significance

Event selection: optimisation

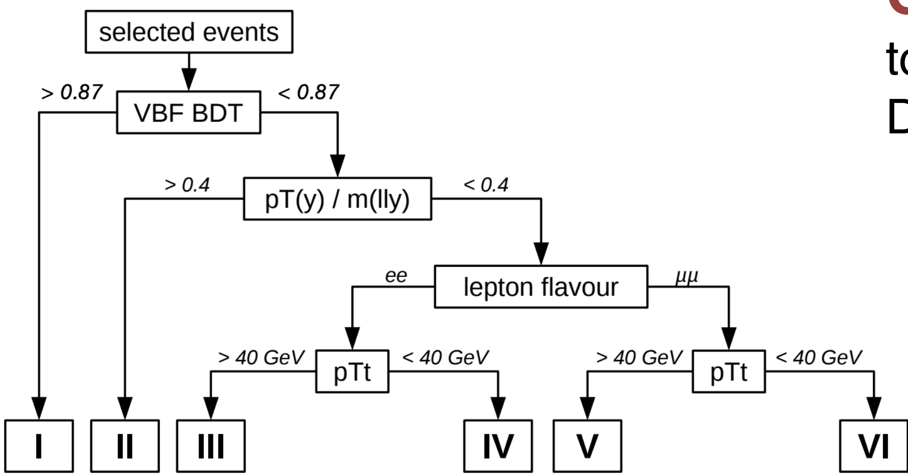
Final photon pT cut: update to photon pT / Mlly > 0.12 (from pT > 15 GeV) - mainly for background study:

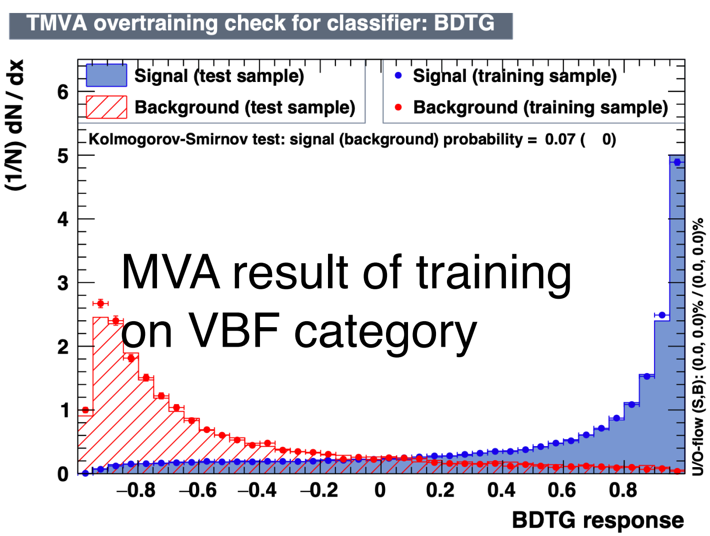
- Low pTt categories are influenced by a “shoulder” between 115-120 GeV with standard pT cut
- relpT>0.12 don’t have a “shoulder” at 115-120GeV
- Equivalent to a photon pT >15GeV cut at mZy = 125GeV





Categorisation





Categorisation is used to enhance the total sensitivity of the analysis

Different categorisation strategies were tried:

- <u>I: VBF category optimised with MVA method + 5 cut-based categories</u>
- II: VBF + ggF categorization with MVA method.
- III: categorisation based on Njet bins and MVA method

Strategy I is used as nominal one with provided highest sensitivity and clearly understood background shape distribution

- 7 variables: $\Delta\Phi_{Z,\gamma}$, $\Delta\eta_{j,j}$, $\Delta R^{min}_{\gamma or Z,j}$, m_{jj}, p_{Tt}, η^{Zepp}, $\Delta\Phi_{Z\gamma,jj}$

Category	Events	S_{68}	w_{68} [GeV]	S_{68}/B_{68} [10^{-2}]	$S_{68}/\sqrt{S_{68}+B_{68}}$
VBF-enriched	174	2.7	3.7	12.2	0.54
High relative p_T	2883	7.6	3.7	6.6	0.68
High p_{Tt} ee	5289	9.9	3.8	2.2	0.46
Low p_{Tt} ee	55092	34.5	4.1	0.5	0.43
High p_{Tt} $\mu\mu$	6606	12.0	3.9	2.0	0.48
Low p_{Tt} $\mu\mu$	73003	43.5	4.0	0.5	0.47
Inclusive	143047	110.2	4.0	0.7	0.85

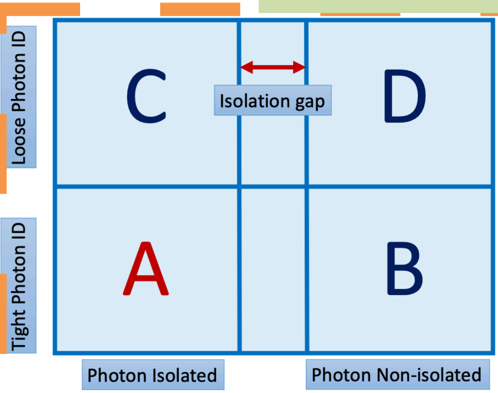
- Background yield is estimated with Zγ MC + data-driven Zjet
- The sensitivity is estimated in the mass region covering 68% signal

Background composition

The dominant background components: SM Zγ (irreducible) and Zjets (reducible)

2 different data-driven methods for the background decomposition:

ABCD method (ID/ISO)

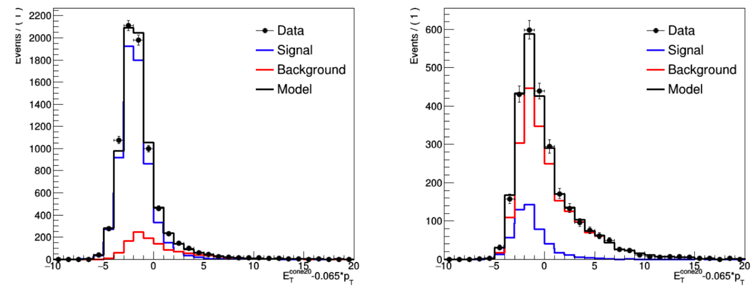


Possible purity variations:
- Vary ID boundaries of C,D with lp2, lp3, lp4, lp5
- Use Zjet MC with/w/o PU to estimate R in ABCD/A'B'C'D'
- Take into account the PU correction, use data in A'B'C'D' region to estimate R

Category	Nominal	Range
Inclusive	0.78	[0.71, 0.82]
e or μ, $BDTG > 0.87$	0.89	[0.65, 1.00]
e or μ, $BDTG < 0.87$, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} > 0.4$	0.92	[0.58, 0.95]
e , BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} > 40$	0.86	[0.76, 0.94]
e , BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} < 40$	0.76	[0.70, 0.81]
μ, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} > 40$	0.85	[0.80, 0.92]
μ, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} < 40$	0.77	[0.69, 0.80]

ABCD method is used as nominal method

isolation template fit method

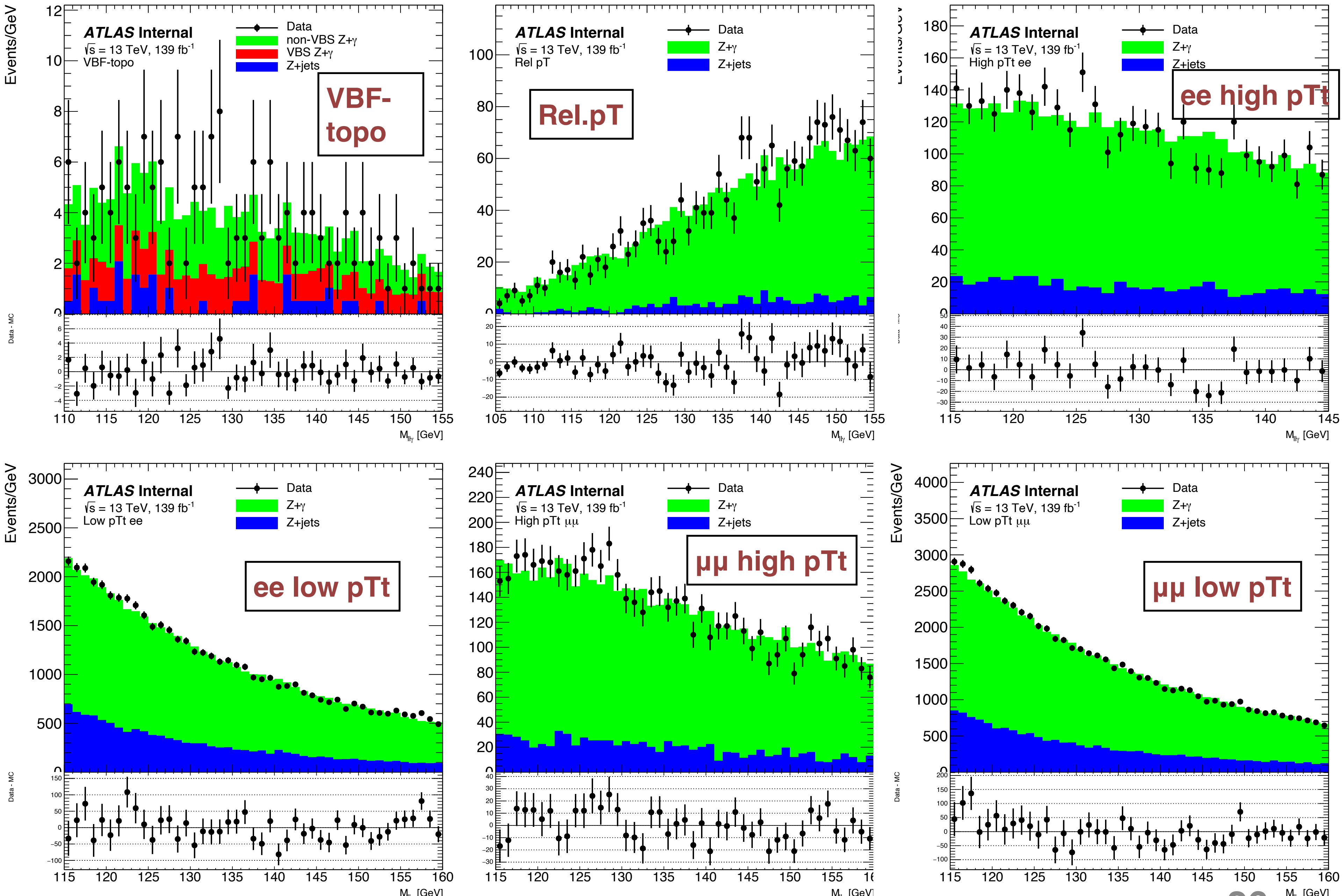


Purity variations are coming from:
- vary ID in bkg region with lp2, lp3, lp5
- vary isolation cone (20/40)

Category	Nominal purity	Range
Inclusive	0.73 ± 0.01	[0.71, 0.81]
e or μ, $BDTG > 0.87$	0.87 ± 0.05	[0.85, 0.87]
e or μ, $BDTG < 0.87$, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} > 0.4$	0.94 ± 0.03	[0.90, 0.96]
e, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} > 40$	0.87 ± 0.01	[0.76, 0.90]
e, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} < 40$	0.75 ± 0.01	[0.71, 0.81]
μ, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} > 40$	0.85 ± 0.01	[0.83, 0.90]
μ, BDTG < 0.87, $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} < 0.4$, $p_{\mathrm{Tt}} < 40$	0.75 ± 0.01	[0.70, 0.81]

- Considered purity variation impacts on the background modelling (little effect)
- We are taking the envelope of the variation ranges from both the 2D-sideband and the template-fit methods as the uncertainty of the purity in each category

Mass spectrum



Signal modeling

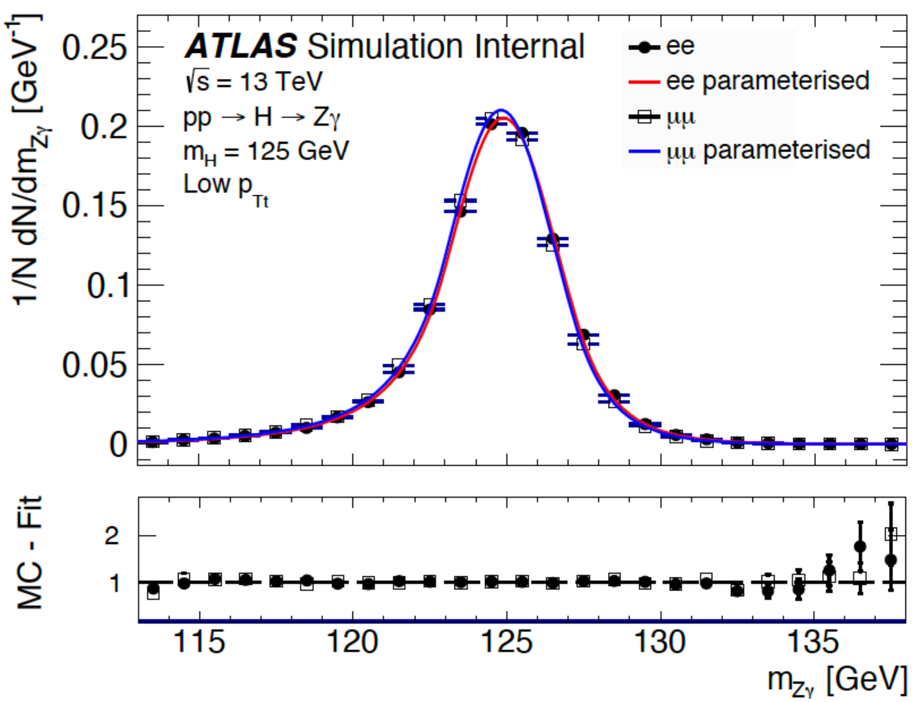
Signal shape:

- individual fit at the 125 GeV in each category on the Higgs MC shape, with **Double Sided Crystal Ball (DSCB)**, shift the mass point to 125.09 GeV

Signal efficiency:

- from Higgs MC (VBF, WpH, WmH, ZH, ttH).

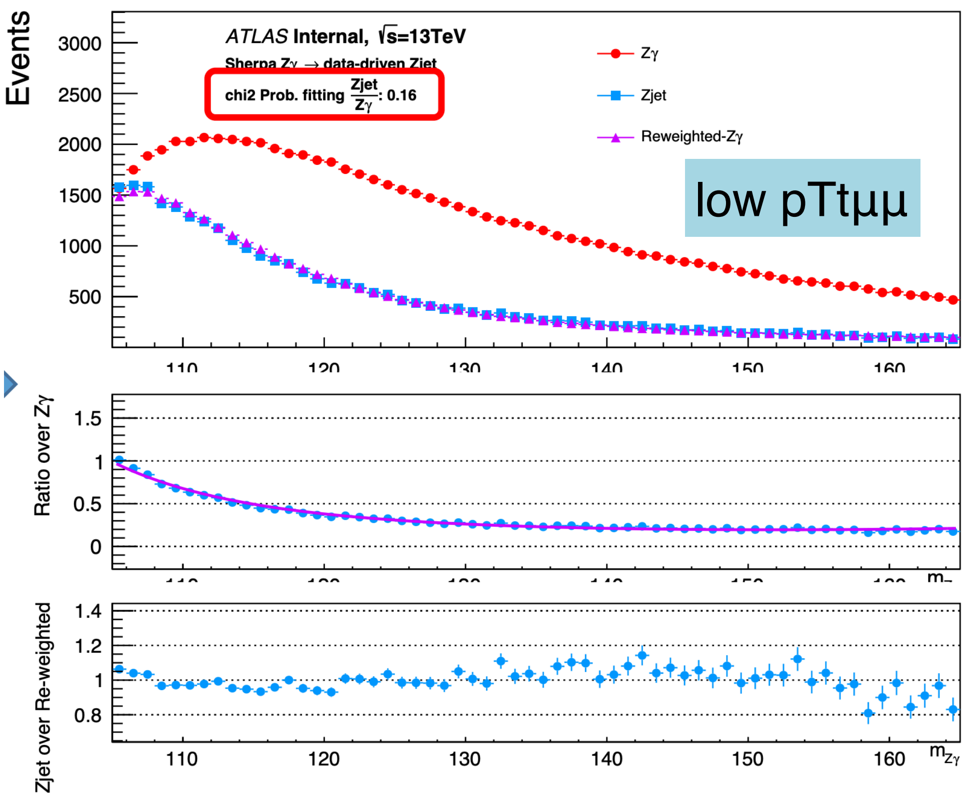
$$N \cdot \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{\mathrm{Lo}} \leq t \leq \alpha_{\mathrm{Hi}} \\ \dfrac{e^{-0.5\alpha_{\mathrm{Lo}}^2}}{\left[\frac{\alpha_{\mathrm{Lo}}}{n_{\mathrm{Lo}}}\left(\frac{n_{\mathrm{Lo}}}{\alpha_{\mathrm{Lo}}}-\alpha_{\mathrm{Lo}}-t\right)\right]^{n_{\mathrm{Lo}}}} & \text{if } t < -\alpha_{\mathrm{Lo}} \\ \dfrac{e^{-0.5\alpha_{\mathrm{Hi}}^2}}{\left[\frac{\alpha_{\mathrm{Hi}}}{n_{\mathrm{Hi}}}\left(\frac{n_{\mathrm{Hi}}}{\alpha_{\mathrm{Hi}}}-\alpha_{\mathrm{Hi}}+t\right)\right]^{n_{\mathrm{Hi}}}} & \text{if } t > \alpha_{\mathrm{Hi}}, \end{cases}$$



Bkg modelling: Zγ reweighing to Zjets

The procedure is done in order to get smooth Zjets mass shape:

- -Fast-sim samples are used to replace Zjets component
- -Define yield in each mass bin BZγ, BZjets for Zγ, Zjets mass distributions
- -Get the ratio R
- -Fit R vs mass by FK1 function $f_{k=1;d=1/3}(x; b, d, a_0, a_1) = (1 - x^{1/3})^b x^{a_0 + a_1 log(x)}$
- -Reweight Zγ samples by R*BZγ to replace Zjets component



The purity is applied and later Zγ+Zjets is normalised to data
It is possible to measure spurious signals in each category

Background function and its uncertainty

A loose spurious signal bias test

Fit range optimization: Lower bound -110, 115, 120; Higher bound - 140, 145, 150, 155, 160

The Fit range optimisation bring ~15% higher significance than fixed fit range of 115-150 GeV

Scan range: 123-129 GeV

Criteria : S/δS<50% χ2 prob. > 1% - depend on the MC statistic we can use, which bring acceptable Bkg Un. comparing to the the Stat. Un.

Both Fit range and Functions are decided by the significance while fitting SM expected Asimov data, and only with spurious signal uncertainty

	varying fit range, 0.25 GeV bin, 139.0/fb					
Event category	Function	Asimov sig.	$P(\chi^2)$	$\frac{S}{\delta S}$	max S	Fit Range
VBF-topo_BDTG	Pow2	0.51	0.602	24.3 %	1.5	110-155
high_rel_pt	ExpPoly2	0.60	0.265	-47.5 %	-7.3	105-155
high_pTt_ee	Bern2	0.37	0.964	-28.1 %	-10.4	115-145
low_pTt_ee	ExpPoly2	0.38	0.372	-21.4 %	-28.2	115-160
high_pTt_$\mu\mu$	Bern3	0.36	0.692	49.6 %	21.4	115-160
low_pTt_$\mu\mu$	Bern3	0.39	0.816	-25 %	-38.7	115-160
		Comb.=1.09				

Systematics: experimental sources

Spurious signal has impact on obs. μ of ~15%, others have impact < 3%

- Comparing to Statistical uncertainty impact of ~43%

Divided to the impacts on signal efficiency (object selections), signal shape (DSCB) mean and resolution, in additional with background uncertainty (spurious signal)

Sources	$H \to Z\gamma$
Luminosity [%]	
Luminosity	1.7
Signal efficiency [%]	
Modelling of pile-up interactions	0.01–0.2
Photon identification efficiency	0.8–1.8
Photon isolation efficiency	0.7–1.9
Electron identification efficiency	0.0–2.3
Electron isolation efficiency	0.0–0.1
Electron reconstruction efficiency	0.0–0.5
Electron trigger efficiency	0.0–0.1
Muon selection efficiency	0.0–0.6
Muon trigger efficiency	0.0–1.6
Jet energy scale,	0.0–3.5
Jet resolution	0.0–15
Jet pile-up	0.0–7.5
Jet flavor	0.0–11
Signal modelling on σ_{CB} [%]	
Electron and photon energy resolution	0.5–3.4
Muon ID resolution	0.0–1.2
Muon MS resolution	0.0–3.4
Signal modelling on μ_{CB} [%]	
Electron and photon energy scale	0.09–0.15
Muon momentum scale	0.0–0.03
Higgs boson mass measurement	0.19
Background modelling [number of spurious signal events]	
Spurious signal	1.5–39

Systematics: theoretical sources

QCD scale and Br has larger impact of 6% on the observed μ

Divided to the impacts on total efficiency and category acceptance

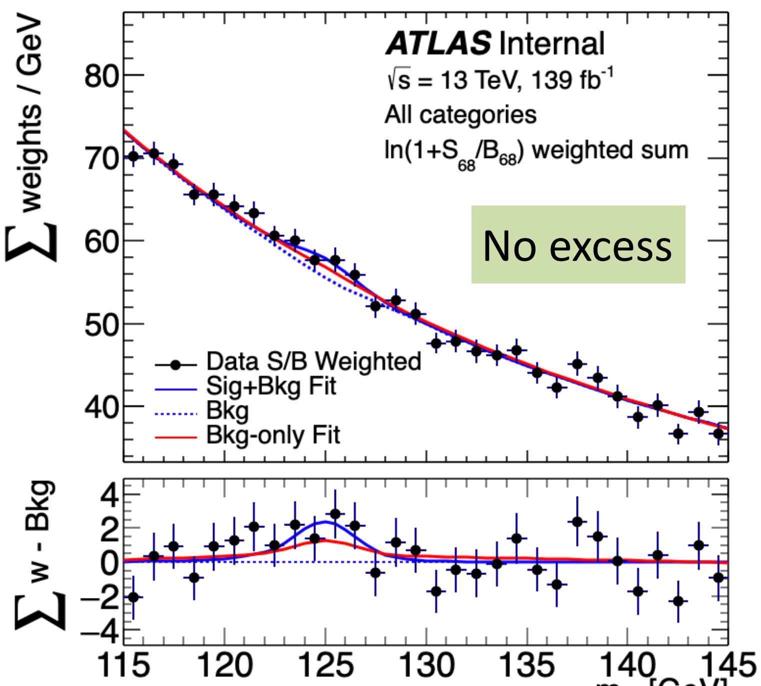
- **Yield uncertainty**: QCD-scale, PDF, Branching ratio, underlying event (MPI-off)
- **Category acceptance:**
  - Modeling on photon pT / mZγ
  - Modeling on pTt in ee/μμ channels
  - Modeling on the VBF BDT score

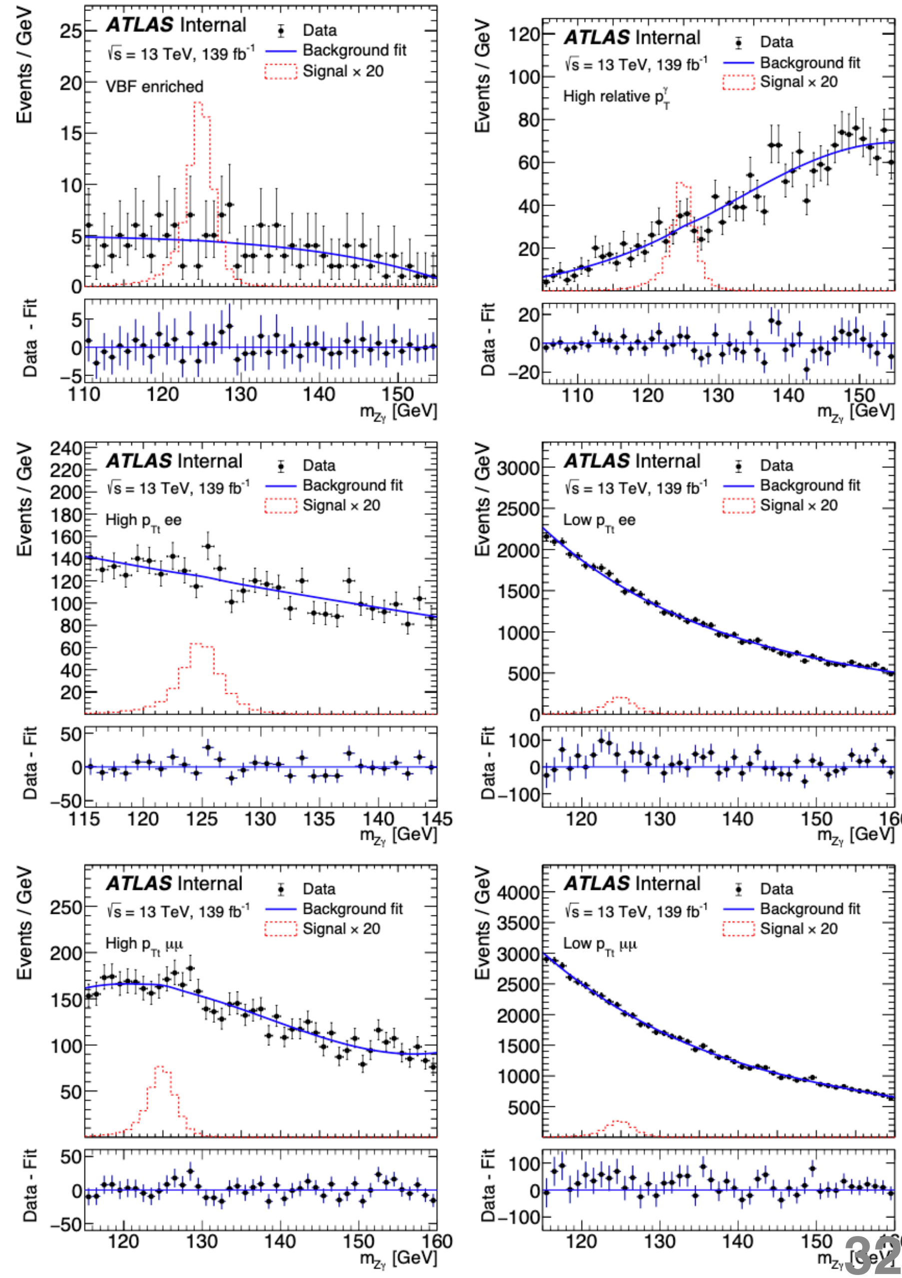
Sources	
Total cross-section and efficiency [%]	
ggF Underlying event	1.3
perturbative order	4.7–9.6
PDF and α_s	1.8–2.8
$B(H \to Z\gamma)$	5.7
Total (total cross-section and efficiency)	7.5–11
Category acceptance [%]	
ggF Underlying event	0.1–11
ggF H p_T perturbative order	0.3–0.4
ggF in VBF-enriched category	34
ggF in high relative p_T^{γ}	17
ggF in other categories	6.7–14
Other production modes	1.0–15
PDF and α_s	0.4–3.5
Total (category acceptance)	11–37

Results

Limits on μ at 95% CL:
Obs.: μ < 3.6
Exp: μ < 1.7 (2.6) assuming no (SM) H→Zγ

BR(H→Zγ) < 0.55% at 95% CL
σ∗BR < 305 fb at 95% CL





Conclusion

Re-optimized full Run2 H→Zγ search is performed, where Z is decaying into $e^-e^+/\mu^-\mu^+$

Several improvements on event selection, categorisation and background model bring higher significance than previous strategies

No evident deviation from background only assumption, limits set on signal strength, σ*Br and Br (branching ratio) with SM cross-section assumption

Results are compatible with expectation.

Statistical uncertainty is still driving the total uncertainty, where the leading systematic uncertainties are from background shapes and theoretical sources

Personal contribution & thesis

Personal contribution to the analysis:

- event selection check & optimisation
- one of few categorisation strategy optimisation
- one of two bkg estimation methods
- signal modeling
- theoretical uncertainties (underlying event, modeling on the VBF BDT score)

Thesis progress:

- Advisors:

- Nicolo de Groot (Prof., Radboud university)
- Anatoli Romaniouk (Prof., MEPhI)

- Time scale:

- ~3 months to finish the draft

Backup Slides

Photon reconstruction and identification

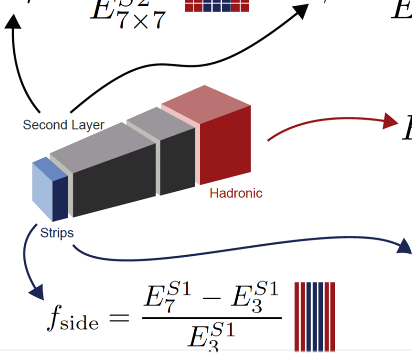
Variables and Position

	Strips	2nd	Had.
Ratios	f_1, f_{side}	R_η*, R_ϕ	$R_{Had.}$*
Widths	$w_{s,3}$, $w_{s,tot}$	$w_{\eta,2}$*	-
Shapes	ΔE, E_{ratio}		

* Used in PhotonLoose.

Energy Ratios

$$R_\eta = \frac{E^{S2}_{3\times7}}{E^{S2}_{7\times7}} \qquad R_\phi = \frac{E^{S2}_{3\times3}}{E^{S2}_{3\times7}}$$



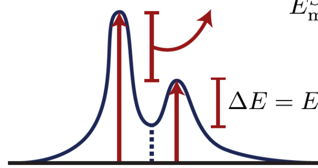
$$R_{Had} = \frac{E^{Had}_T}{E_T}$$

$$f_1 = \frac{E_{S1}}{E_{Tot.}}$$

$$f_{side} = \frac{E^{S1}_7 - E^{S1}_3}{E^{S1}_3}$$

Shower Shapes

$$E_{ratio} = \frac{E^{S1}_{max,1} - E^{S1}_{max,2}}{E^{S1}_{max,1} + E^{S1}_{max,2}}$$

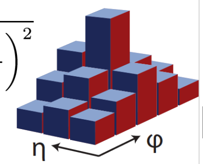


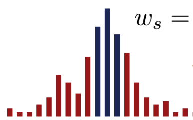
$$\Delta E = E^{S1}_{max,2} - E^{S1}_{min}$$

Widths

$$w_{\eta,2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$$

Width in a 3x5 (Δη×Δφ) region of cells in the second layer.

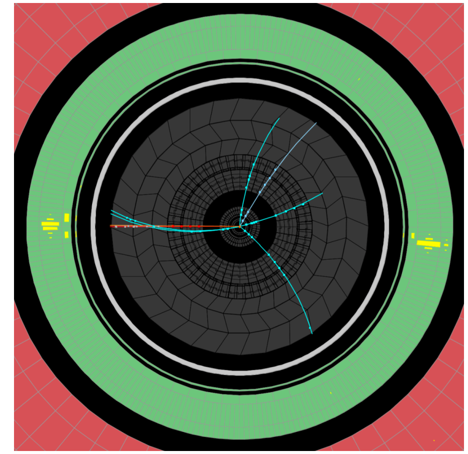




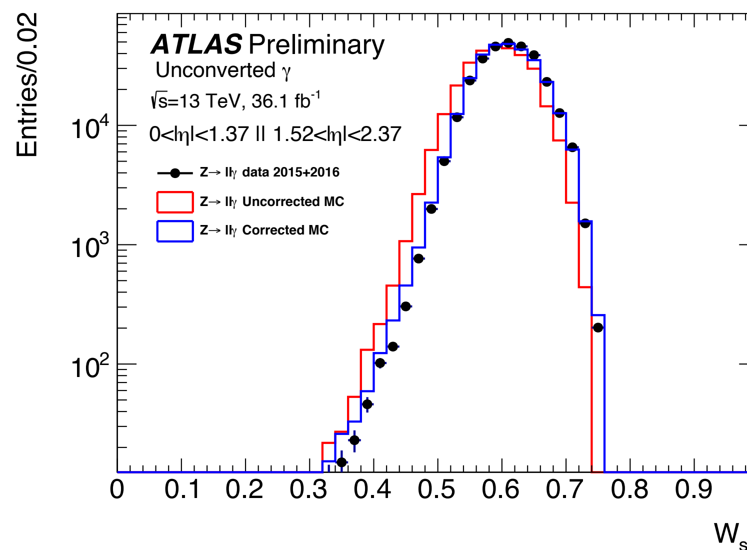
$$w_s = \sqrt{\frac{\sum E_i (i - i_{max})^2}{\sum E_i}}$$

ws3 = w_1 uses 3 strips in η; wstot is defined similarly, but uses 20 strips.

ID: 9 discriminating variables (DVs) based on energy in cells of ECAL and leakage in hadronic calorimeter HCAL



left photon: converted (two tracks)
right photon: unconverted



Shower shapes in MC are different than in data → MC is shifted to reproduce the data

Photon discriminating variables

Category	Description	Name	*loose*	*tight*
Acceptance	$\|\eta\| < 2.37$, with $1.37 < \|\eta\| < 1.52$ excluded	–	✓	✓
Hadronic leakage	Ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{had_1}	✓	✓
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{had}	✓	✓
EM Middle layer	Ratio of 3×7 $\eta \times \phi$ to 7×7 cell energies	R_η	✓	✓
	Lateral width of the shower	w_{η_2}	✓	✓
	Ratio of 3×3 $\eta \times \phi$ to 3×7 cell energies	R_ϕ		✓
EM Strip layer	Shower width calculated from three strips around the strip with maximum energy deposit	$w_{s\,3}$		✓
	Total lateral shower width	$w_{s\,\mathrm{tot}}$		✓
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips	F_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima	ΔE		✓
	Ratio of the energy difference associated with the largest and second largest energy deposits to the sum of these energies	E_{ratio}		✓

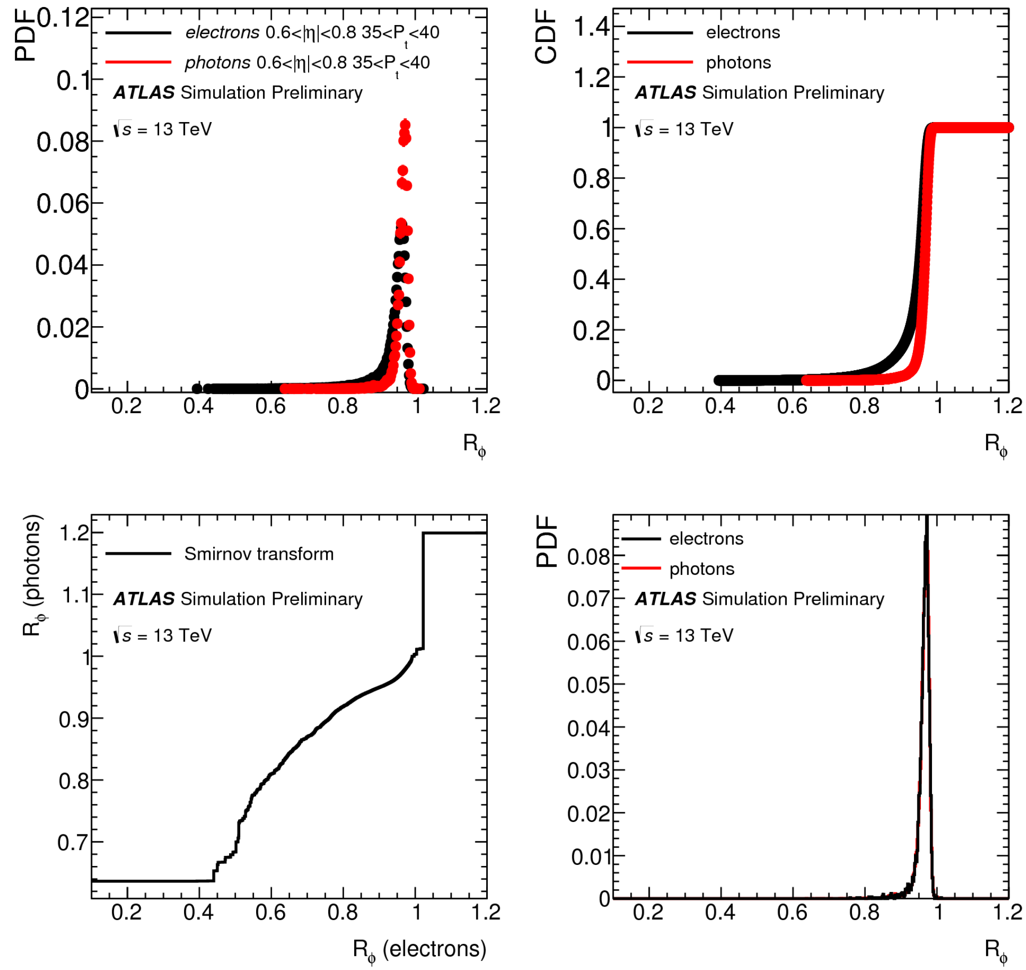
Table 1: Discriminating variables used for *loose* and *tight* photon identification.

Photon ID: Electron extrapolation method

- Shower shape distributions of electrons and photons γ are similar due to similar interactions of photons and electrons in the detector
- Select a pure sample of electrons from Z decays using a tag-and-probe method and transform their shower shape distributions such that the resulting object has photon properties:

$$s' = \mathrm{CDF}_{\gamma}{}^{-1}(\mathrm{CDF}_{\mathrm{e}}(s))$$

- Typical E_T of electrons from Z decays of order $m_Z/2$ -> measurement in range:
    - **E_T in [25; 150] GeV**



Photon ID: Matrix method

- Sample of inclusive photons collected with a single-photon trigger
- Large kinematic range: **E_T in [25; 1500] GeV**
- ID efficiency can be computed by employing an additional discriminating variable: track isolation (assumed uncorrelated with shower shape variables) which is applied before and after ID cuts

$$\varepsilon_{\mathrm{ID}} = \frac{N^S_{\mathrm{ID}}}{N^S}$$

$$\hat{N}_{\mathrm{ID}} = \hat{\varepsilon}^S_{\mathrm{ID}} \cdot N^S_{\mathrm{ID}} + \hat{\varepsilon}^B_{\mathrm{ID}} \cdot N^B_{\mathrm{ID}}$$

$$\hat{N} = \hat{\varepsilon}^S \cdot N^S + \hat{\varepsilon}^B \cdot N^B$$

$$\boxed{\varepsilon_{\mathrm{ID}} = \frac{\frac{\hat{\varepsilon}_{\mathrm{ID}} - \hat{\varepsilon}^B_{\mathrm{ID}}}{\hat{\varepsilon}^S_{\mathrm{ID}} - \hat{\varepsilon}^B_{\mathrm{ID}}} \cdot N_{\mathrm{ID}}}{\frac{\hat{\varepsilon} - \hat{\varepsilon}^B}{\hat{\varepsilon}^S - \hat{\varepsilon}^B} \cdot N}}$$

Track-isolation efficiencies are obtained:

- from MC for signal (photons)
- from data for background, making use of low correlation between strip layer variables and track isolation

Framework, samples and preselection

Where to find:

Analysis package is based on EGamma ZllgAnalysis - git

eos ntuples:

/eos/atlas/atlascerngroupdisk/perf-egamma/photonID/NTUP_ZLLG/

Contains: data15/16/17/18, MC15/16

Samples:

- MC: MC16a/d/e Zeeγ, Zmumuγ (Sherpa, PowhegPythia), EGAM3/4
- Data: 2015-2018 (~140 fb-1), EGAM3/4

Preselection:

Leptons: Et(el,mu) > 10 GeV, $|\eta_{el}| < 2.47$, $|\eta_{mu}| < 2.7$, Loose isolation

Photons: Et > 10 GeV, Loose OR Tight ID, deltaR(el/mu, γ) > 0.4, $|\eta| < 2.37$

Event selection: 40 < Mll < 83 GeV; 80 < Mllγ < 100 GeV, trigger matching

Methods of background estimation

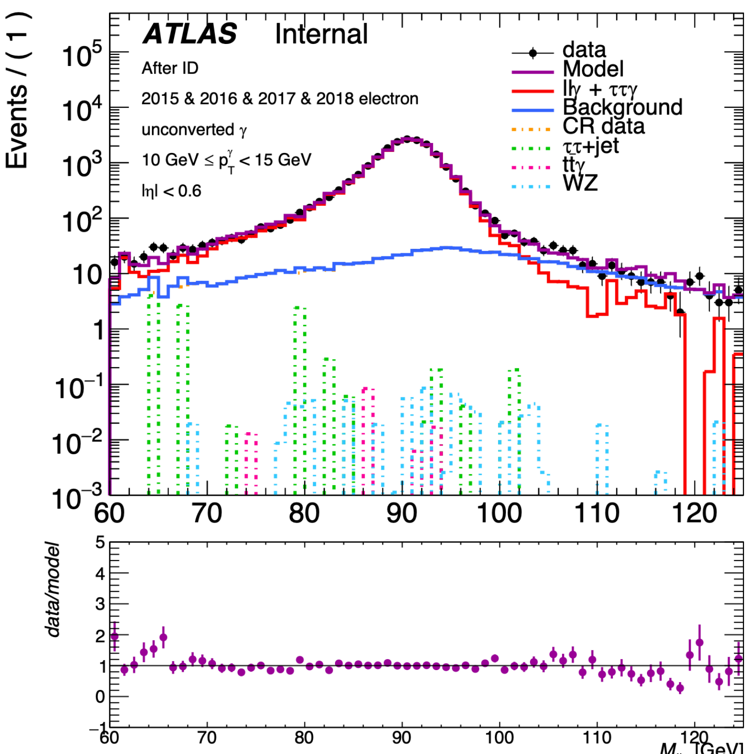
Purity estimation with a mass fit method (Zihang) - [details]

Model construction for template fit:

Model = signal (llgamma + tautaugamma scaled) + CR data + (tautaujet + ttbar+ WZ)

Floating fraction 1 (signal) | Floating fraction 2 (CR data) | **Fixed** fraction (tautaujet + ttbar+ WZ)

A template fit to data for pT range [10, 35] GeV; efficiencies for higher bins are obtained by counting.



$$n_{sig} = N_{sig} \times f_{sig}$$

$$n_{bkg^i} = N_{bkg^i} \times f_{bkg^i} \quad (i = 1,2,3,4)$$

$$n_{total} = n_{sig} + \sum_i n_{bkg^i}$$

$$\sigma_{n_{sig}}{}^2 = N_{sig}{}^2 \times \sigma_{f_{sig}}{}^2 + f_{sig}{}^2 \times \sigma_{N_{sig}}{}^2$$

$$\sigma_{n_{bkg^i}}{}^2 = N_{bkg^i}{}^2 \times \sigma_{f_{bkg^i}}{}^2 + f_{bkg^i}{}^2 \times \sigma_{N_{bkg^i}}{}^2$$

$$\sigma_{total}^2 = \sigma_{n_{sig}}{}^2 + \sum_i \sigma_{n_{bkg^i}}{}^2$$

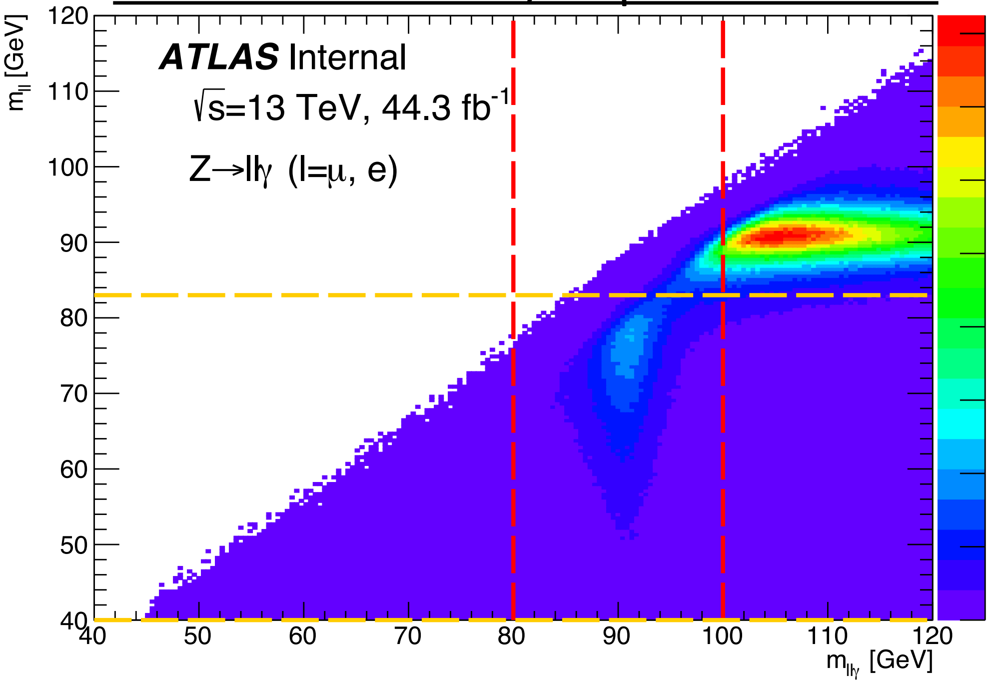
$$P = \frac{n_{sig}}{n_{sig} + \sum_i n_{bkg^i}}$$

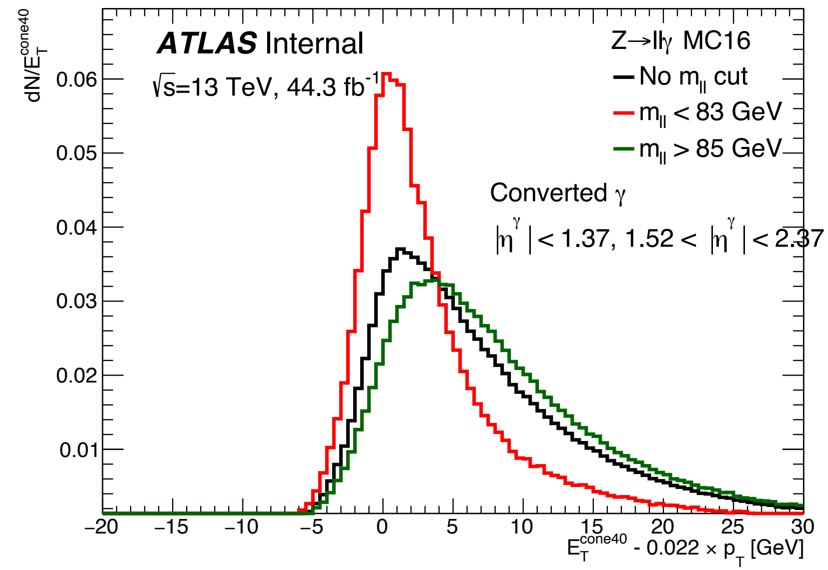
$$\sigma_P{}^2 = \frac{\sigma_{n_{sig}}{}^2}{n_{total}{}^2} + \frac{n_{sig}{}^2 \times \sigma_{total}^2}{n_{total}{}^4}$$

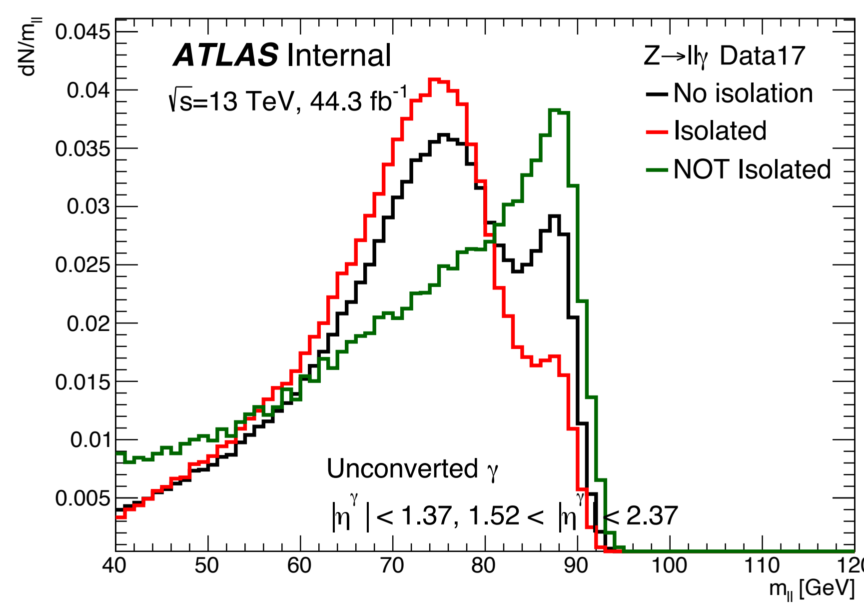
2d sideband method

Method validation

Current cuts on mll/mllγ can be varied







Isolated: cut on isolation is applied

NOT isolated: revered cut on isolation is applied

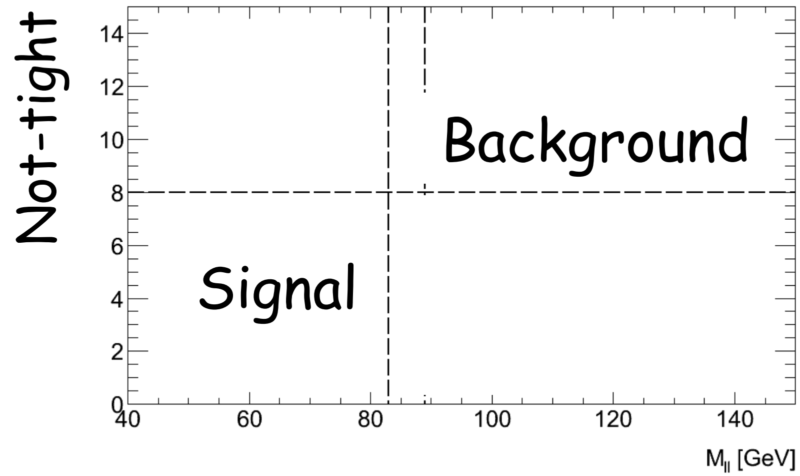
Isolation: Methods of background estimation

2d sideband method

- <u>Use of loose prime photons</u>: bits 17, 19, 20 and 21 are removed (tight–4 – less correlated with the isolation)

- Reversing Mll cut: for Mll > 85 GeV = almost all photons should be jets->γ

- N_{bkg} can be estimated by normalizing control bkg shape to the tail of signal (tight) distribution



Method allows to estimate background contamination up to ~50 GeV

Used for <u>isolation only</u>

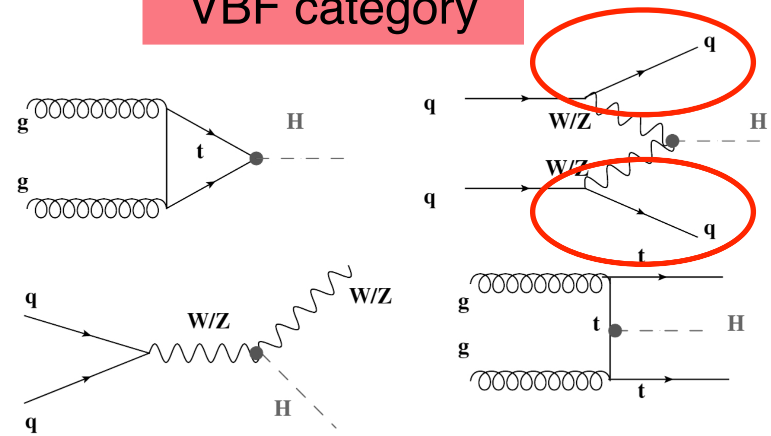
H->Zy analysis chain

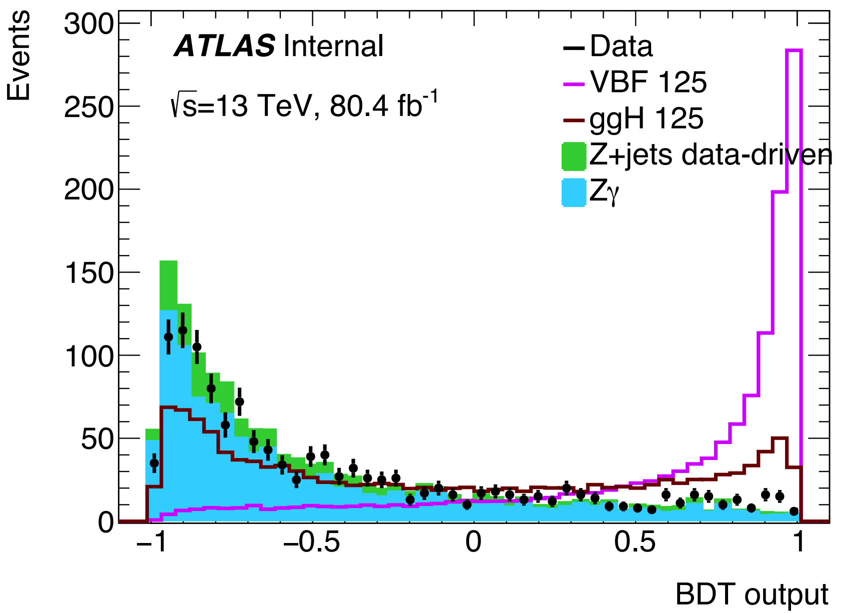
Variables used in MVA training:

$\Delta\Phi_{Z,\gamma}$	Azimuthal angle between di-lepton system and photon
$\Delta\eta_{jj}$	Pseudo-rapidity separation of dijet
$\Delta R^{min}_{\gamma or Z,j}$	Minimum ΔR between one object of the Zgamma and jets
m_{jj}	Invariant mass of dijet
p_{Tt}	Zgamma p_{T} projected perpendicular to the Zgamma thrust axis
$\eta^{Zeppenfeld}$	$\lvert\eta_{Z\gamma} - 0.5 * (\eta_{j1} + \eta_{j2})\rvert$
$\Delta\Phi_{Z\gamma,jj}$	Azimuthal angle between Zgamma and dijet system

$$p_{\mathrm{T}t} = 2\frac{|p_x^Z p_y^\gamma - p_x^\gamma p_y^Z|}{p_{\mathrm{T}}^{Z\gamma}}$$

VBF category

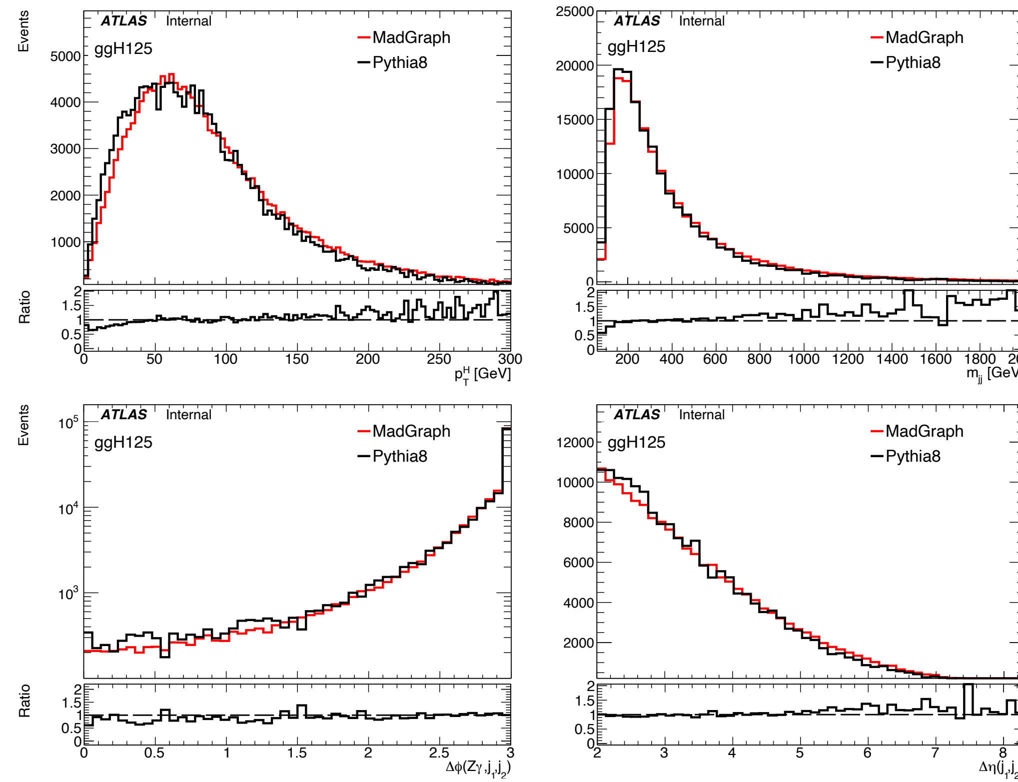


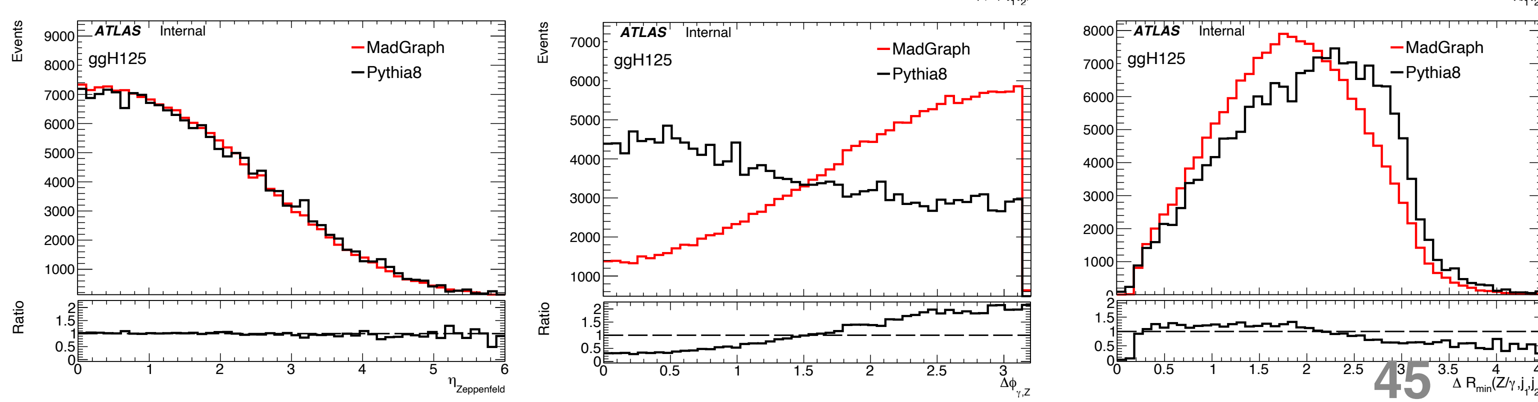


Systematics: theoretical sources

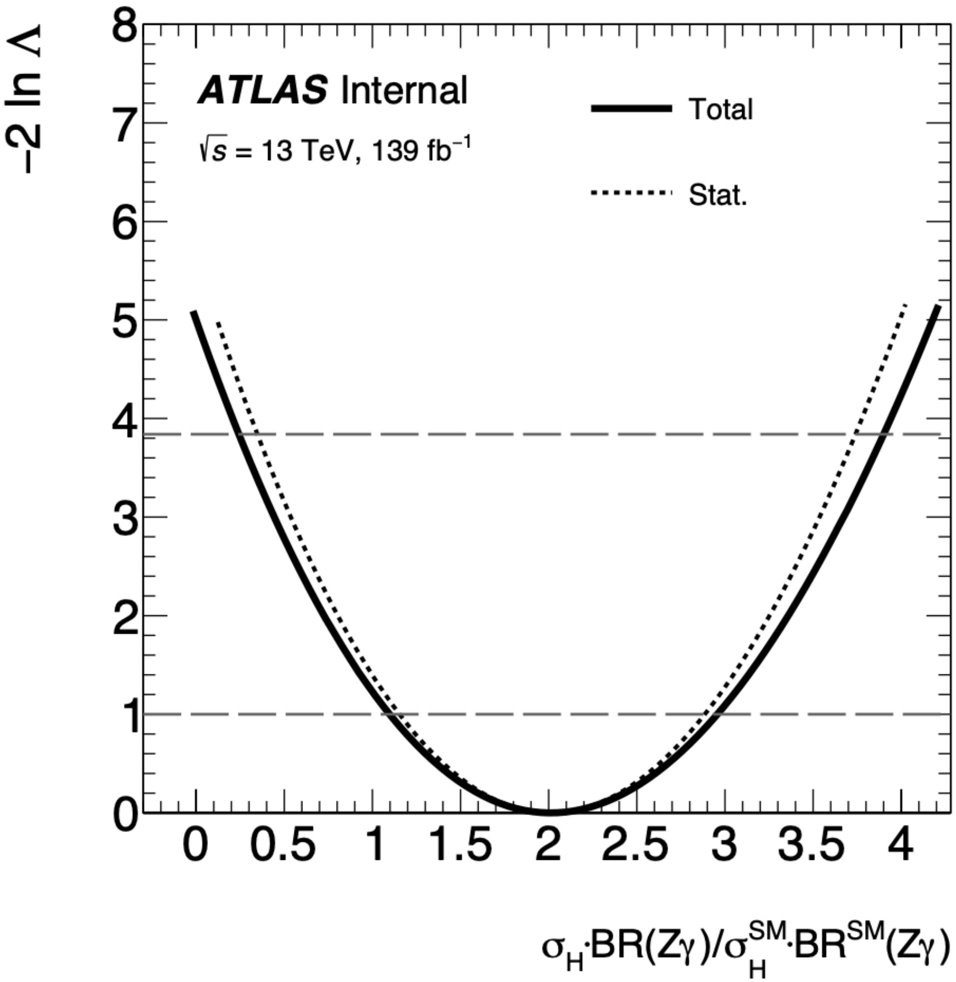
Example: modelling on the VBF BDT score

Delivered from MG5_aMC@NLO/Pythia8 sample using H → γγ events (H →Zγ is not available) at evgen/truth level
- difference between nominal and alternative sample is taken as an uncertainty





Results

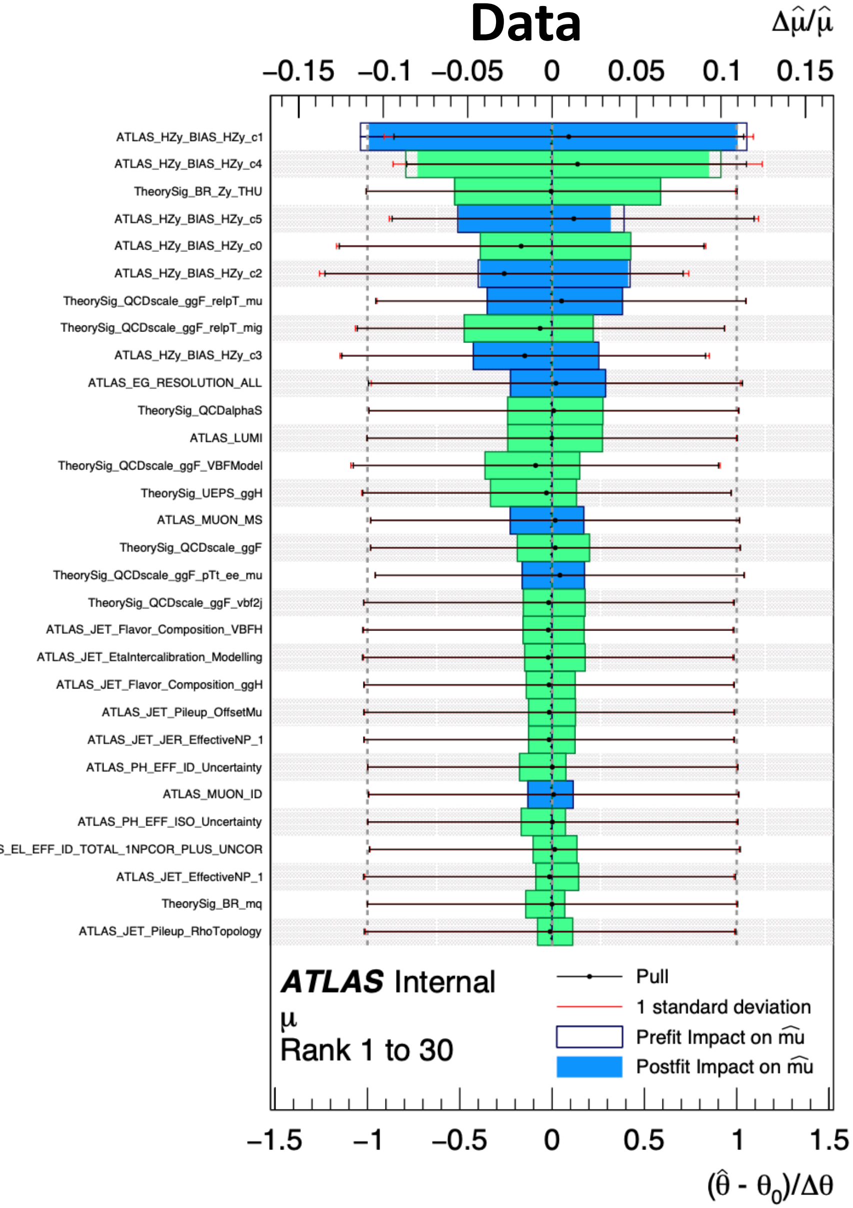


Category	μ	Significance
VBF-topo	$0.5^{+1.9}_{-1.7}$ $(1.0^{+2.0}_{-1.6})$	0.3 (0.6)
Rel. pT	$1.6^{+1.7}_{-1.6}$ $(1.0^{+1.7}_{-1.6})$	1.0 (0.6)
High pTt ee	$4.7^{+3.0}_{-2.7}$ $(1.0^{+2.7}_{-2.6})$	1.7 (0.4)
Low pTt ee	$3.9^{+2.8}_{-2.7}$ $(1.0^{+2.7}_{-2.6})$	1.5 (0.4)
High pTt $\mu\mu$	$2.9^{+3.0}_{-2.8}$ $(1.0^{+2.8}_{-2.7})$	1.0 (0.4)
Low pTt $\mu\mu$	$0.8^{+2.6}_{-2.6}$ $(1.0^{+2.6}_{-2.5})$	0.3 (0.4)
Combined	$2.0^{+1.0}_{-0.9}$ $(1.0^{+0.9}_{-0.9})$	2.2 (1.2)

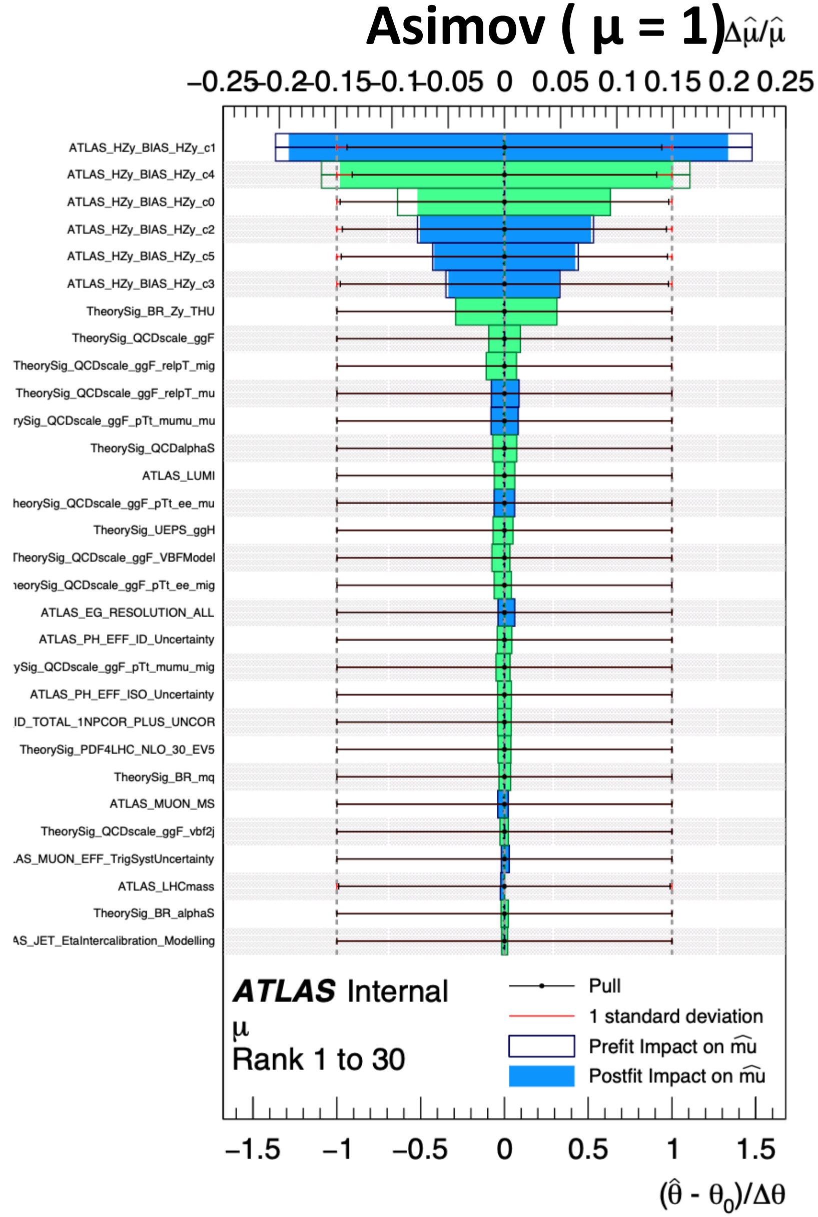
NP ranking

The changed orders are due to 2x significance in observed data

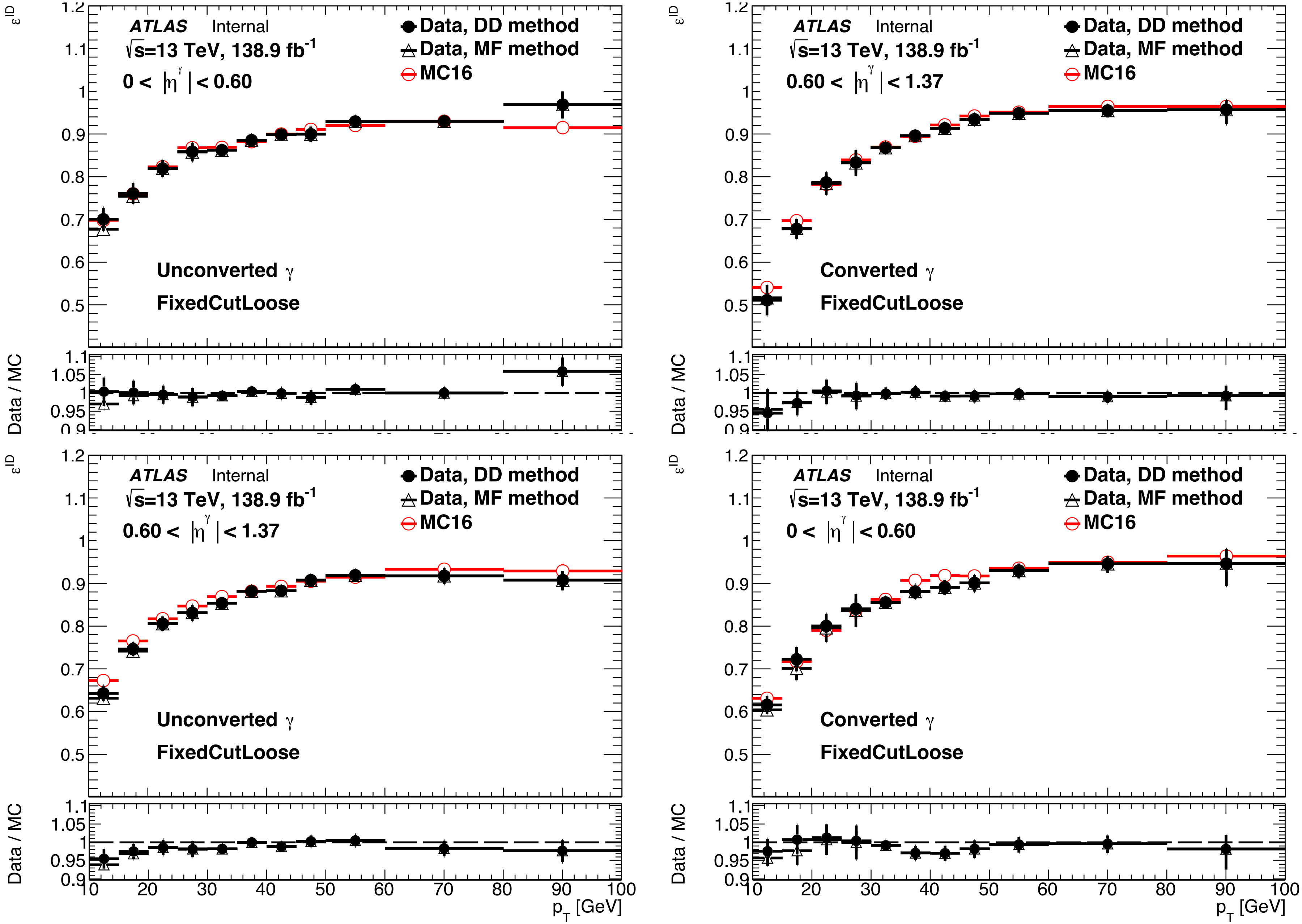
Data



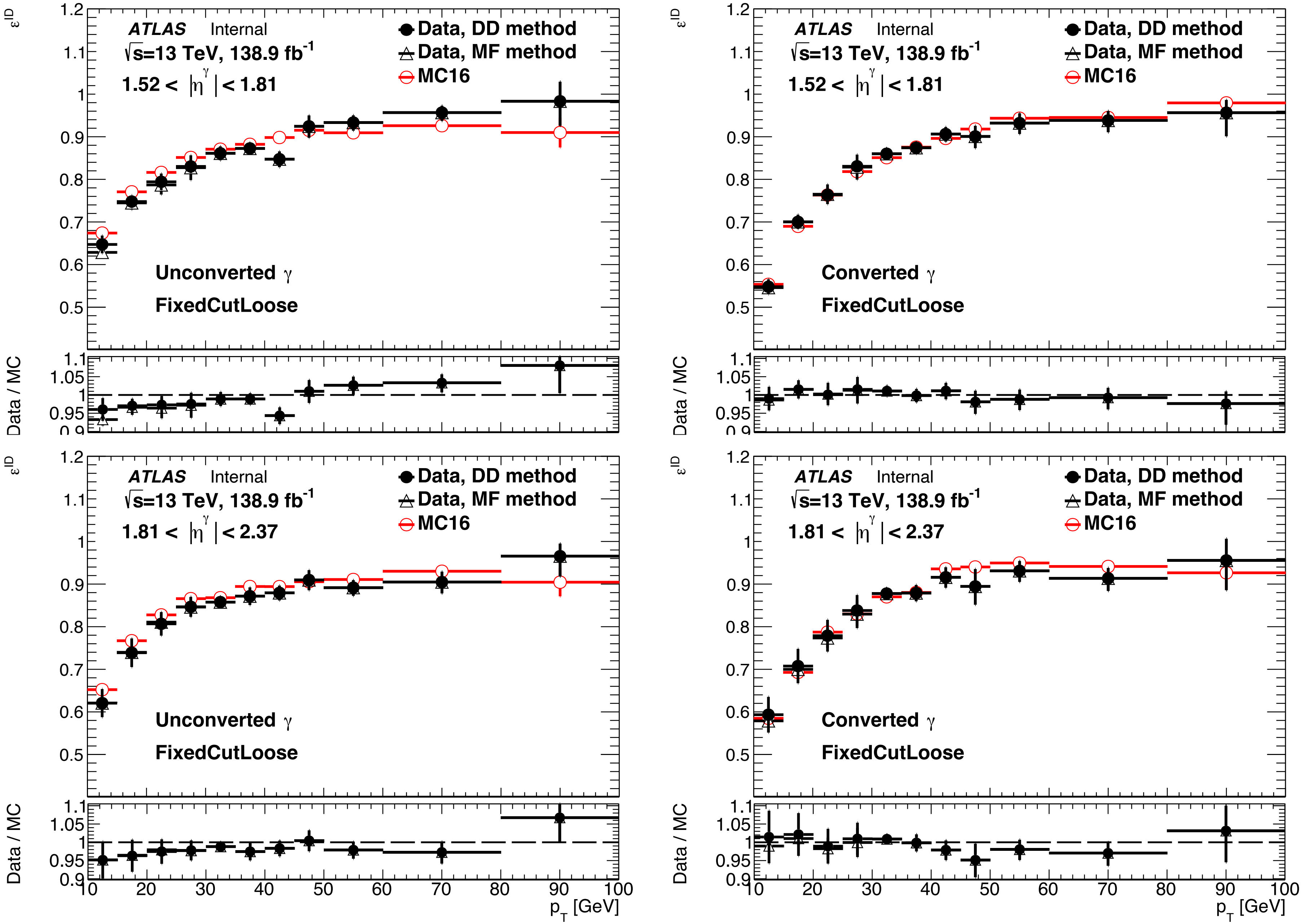
Asimov (μ = 1)



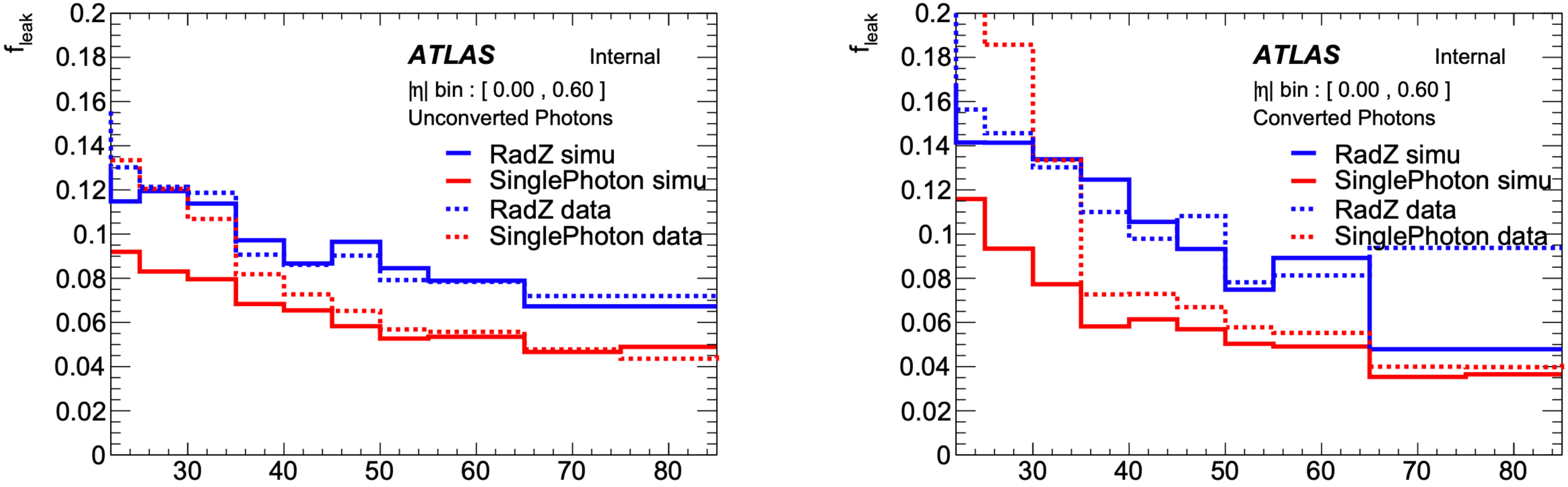
Photon ID efficiency vs pT



Photon ID efficiency vs pT



Photon iso leakage



similar behaviour in all eta bins

Leakage fraction obtained from MC:

$$f_{leak,MC} = \frac{loose'4_{MC}}{tight_{MC} + loose'4_{MC}}$$

Obtained directly from MC for radZ and SinglePhotons

Obtained from the fits on data for SinglePhotons,
directly from data for radZ