



Measurement of Di-Higgs Production in the $b\bar{b}\gamma\gamma$ Channel with the ATLAS Experiment: Photon Identification Calibration

Hind Taibi

Supervisor: Marco Delmastro

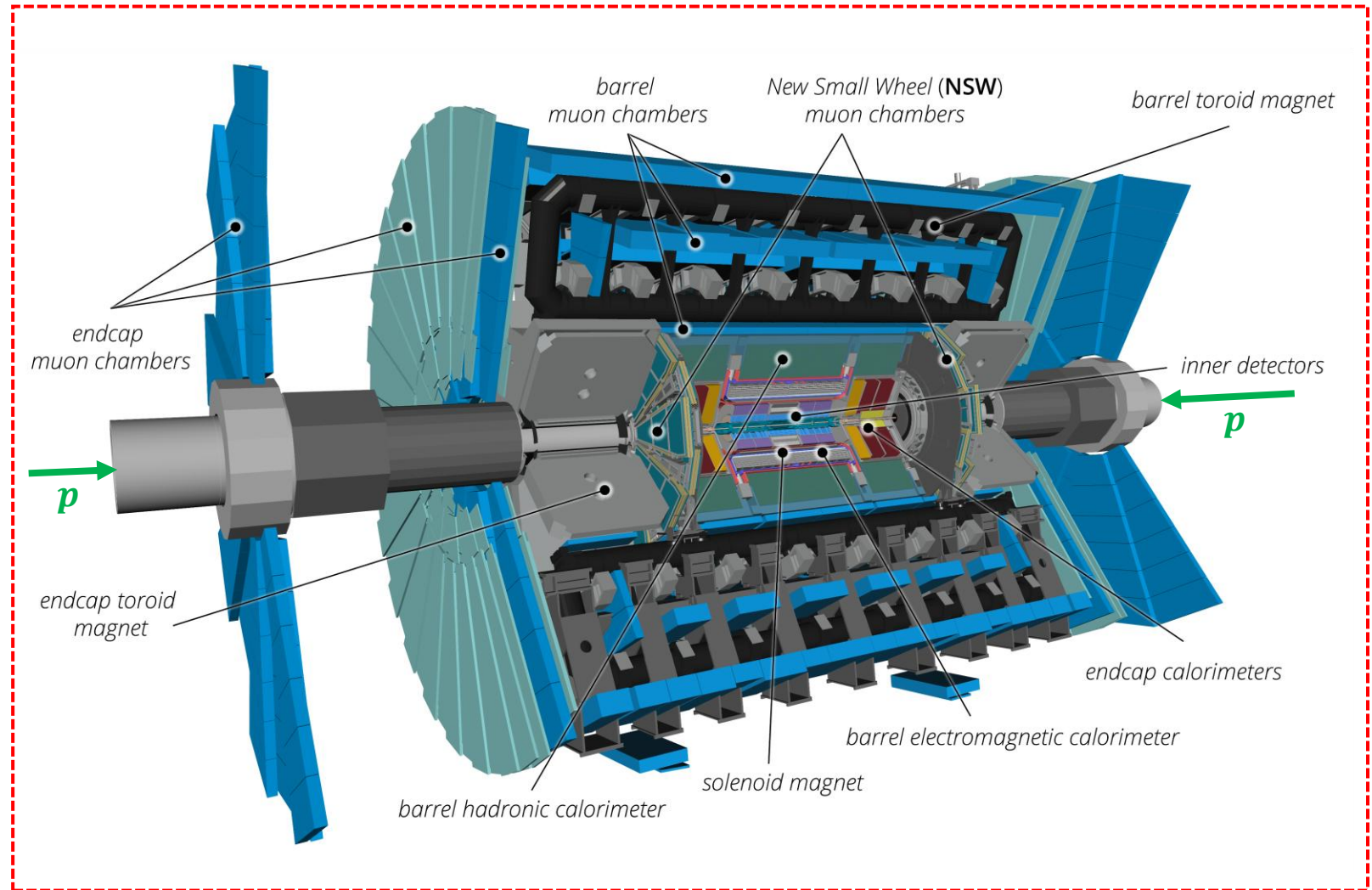
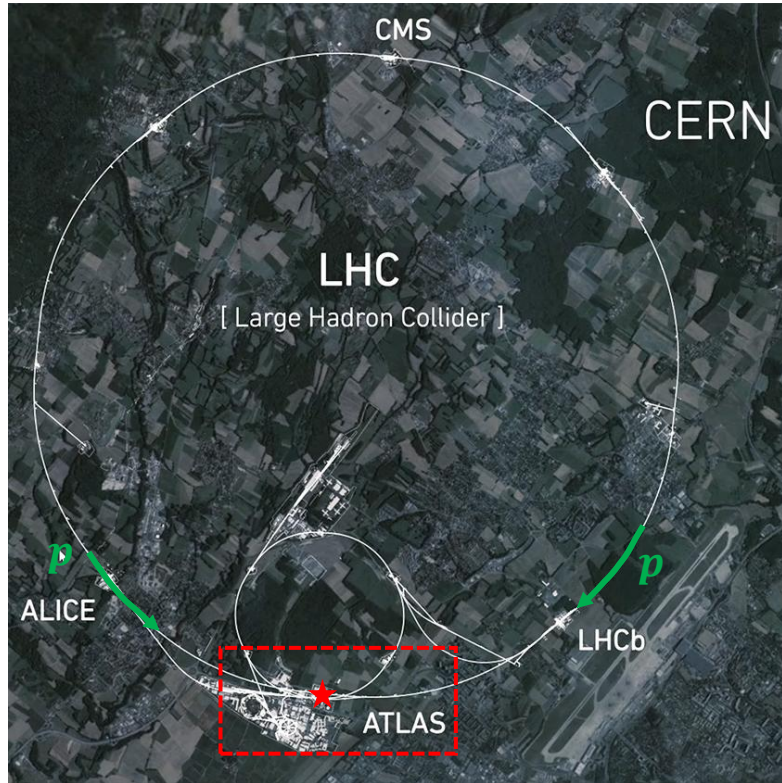
Comité de suivi individuel (CSI)

29th May 2026

Introduction

Introduction

The ATLAS Detector

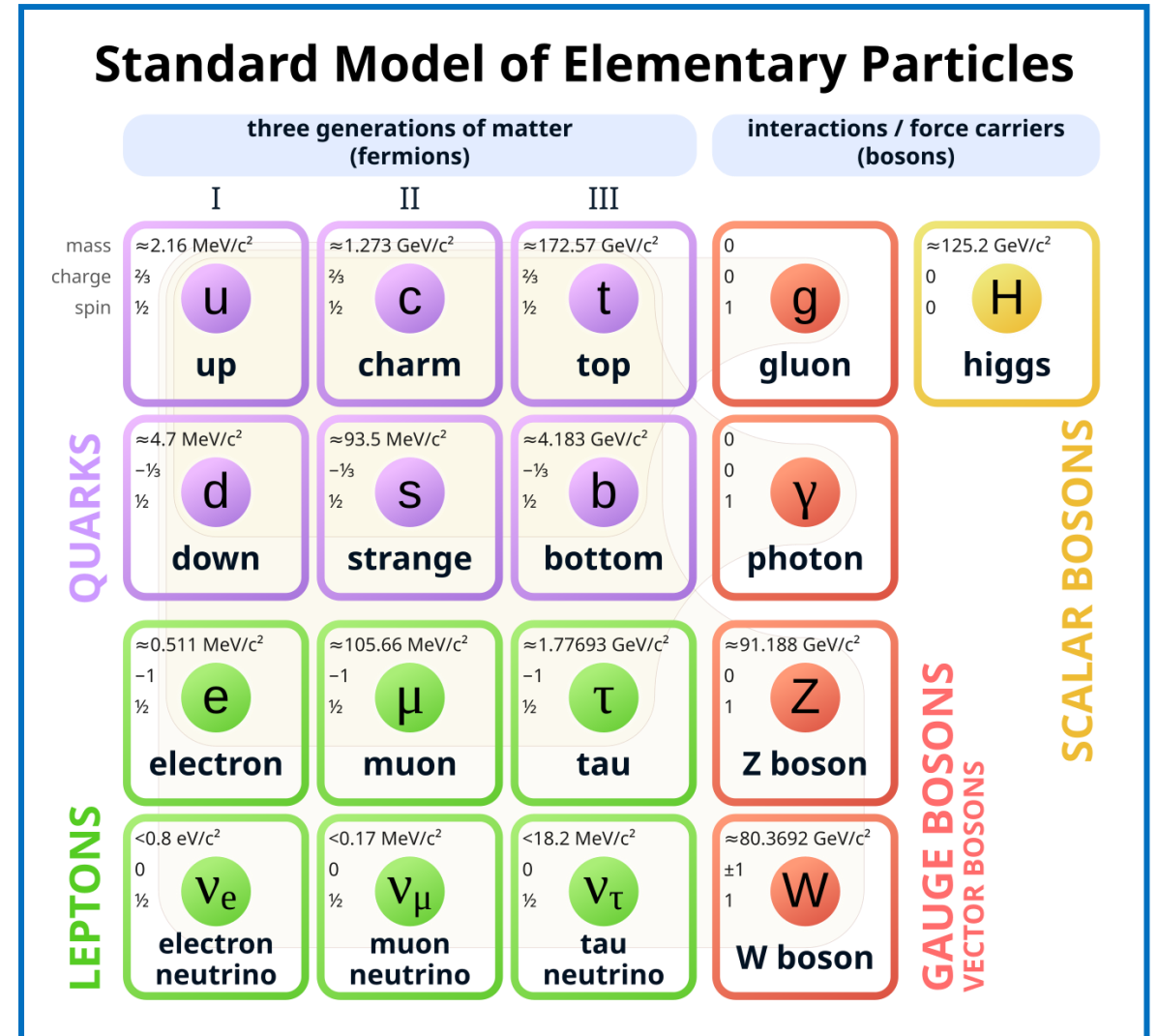


Introduction

The Standard Model of Particle Physics

\mathcal{L}^{SM} : describes the properties of the particles and their interactions

Standard Model
Lagrangian



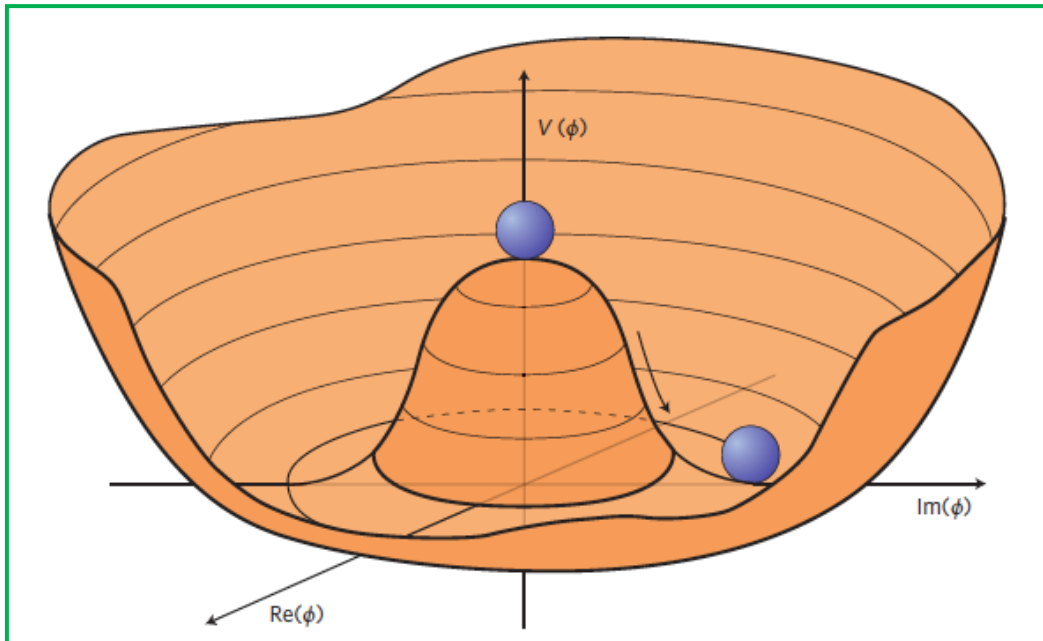
Introduction

The Higgs Potential

$$\mathcal{L}^{\text{SM}} \supset V(\Phi) = V_0 - \frac{m_H^2}{2} h^2 + \lambda_{\text{HHH}}^{\text{SM}} h^3 + \lambda_{\text{HHHH}}^{\text{SM}} h^4$$

↓
Higgs boson's self-coupling constant

↓
Not accessible at the LHC



Standard Model of Elementary Particles

	three generations of matter (fermions)			interactions / force carriers (bosons)	
	I	II	III		
mass	≈2.16 MeV/c ²	≈1.273 GeV/c ²	≈172.57 GeV/c ²	0	≈125.2 GeV/c ²
charge	2/3	2/3	2/3	0	0
spin	1/2	1/2	1/2	1	0
	u up	c charm	t top	g gluon	H higgs
	d down	s strange	b bottom	γ photon	
	e electron	μ muon	τ tau	Z Z boson	
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	W W boson	

QUARKS (left side of the table)

LEPTONS (left side of the table)

GAUGE BOSONS VECTOR BOSONS (bottom right of the table)

SCALAR BOSONS (right side of the table)

Introduction

Di-Higgs Searches

$$\mathcal{L}^{\text{SM}} \supset V(\Phi) = V_0 - \frac{m_H^2}{2} h^2 + \lambda_{HHH}^{\text{SM}} h^3 + \lambda_{HHHH}^{\text{SM}} h^4$$

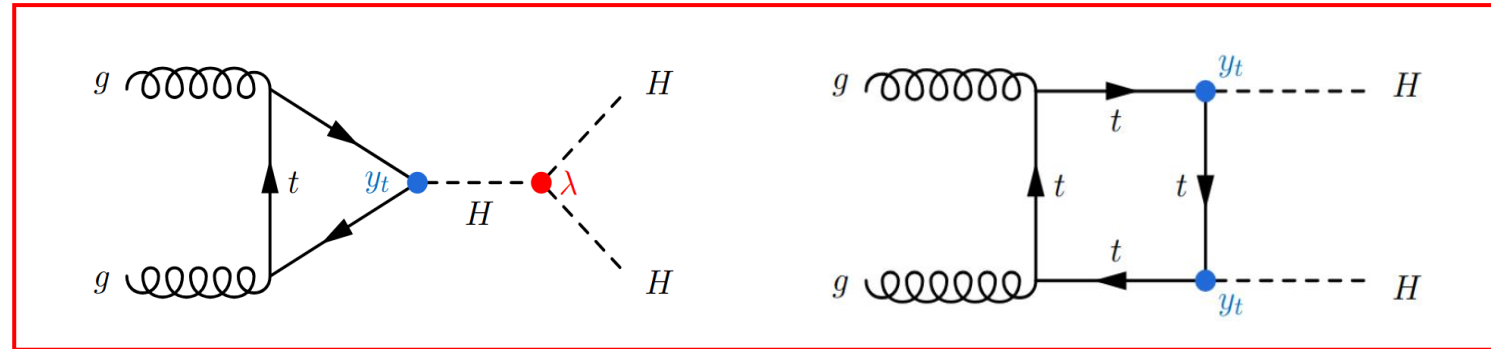
Latest constraints on the Higgs boson's self-coupling constant:

$$-1.3 < \kappa_\lambda < 6.7$$

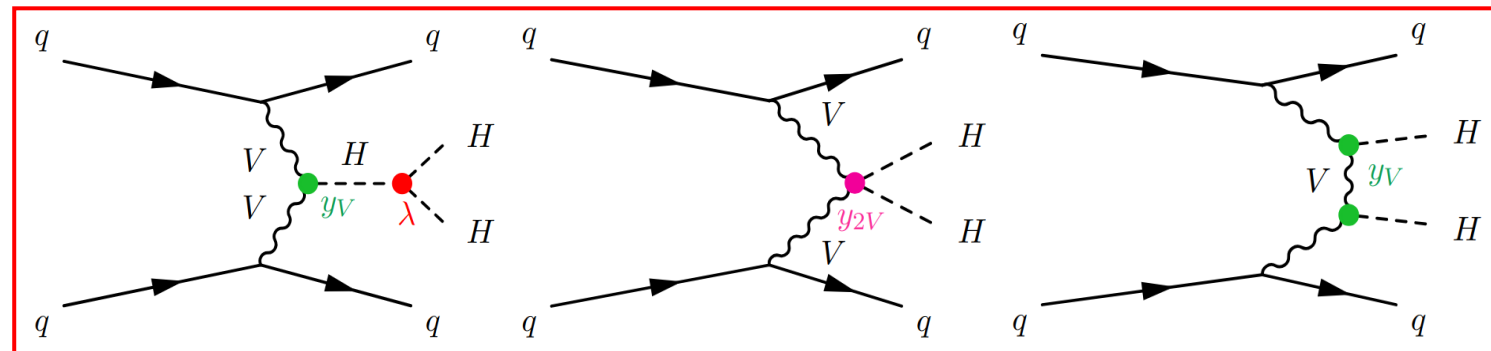
$$\equiv \lambda_{HHH} / \lambda_{HHH}^{\text{SM}}$$

[Combination of ATLAS and CMS searches for Higgs boson pair production at \$\sqrt{s} = 13\$ TeV with the ATLAS detector](#)

LO ggF Feynman Diagrams



LO VBF Feynman Diagrams



Introduction

The $HH \rightarrow b\bar{b}\gamma\gamma$ Channel

High contribution from $HH \rightarrow b\bar{b}\gamma\gamma$ in di-Higgs searches

Latest results of the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis using full Run 2 and partial Run 3 data:

$$-1.6 < \kappa_\lambda < 6.6$$

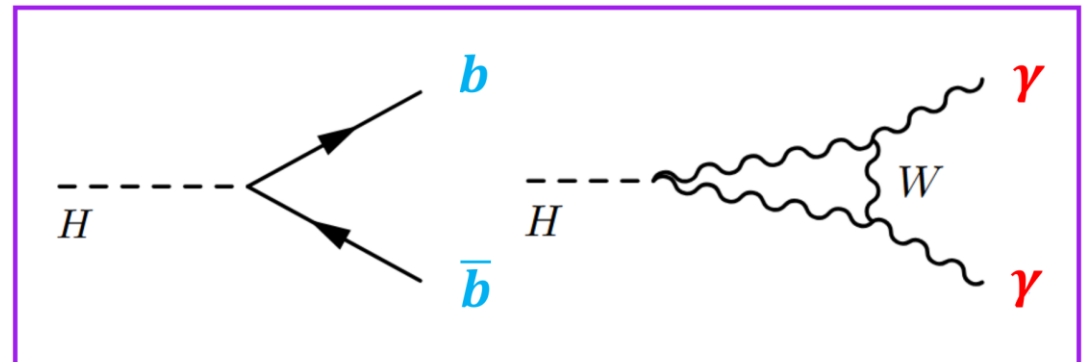
With $\left\{ \begin{array}{l} \text{New } b\text{-tagging algorithm} \\ \text{Same photo identification procedure} \end{array} \right.$

➔ Implement **new photon id.** for next analysis (e. g. BDT, gives **~5-10 % more efficiency**)

➔ **Calibrate** the new photon id.

[Study of Higgs boson pair production in the \$HH \rightarrow b\bar{b}\gamma\gamma\$ final state with \$308 \text{ fb}^{-1}\$ of data collected at \$\sqrt{s} = 13 \text{ TeV}\$ and 13.6 by the ATLAS experiment](#)

	bb	WW	$\tau\tau$	ZZ	$\gamma\gamma$
bb	34%				
WW	25%	4.6%			
$\tau\tau$	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
$\gamma\gamma$	0.26%	0.10%	0.028%	0.012%	0.0005%



Photon Identification in ATLAS

Photon Identification in ATLAS

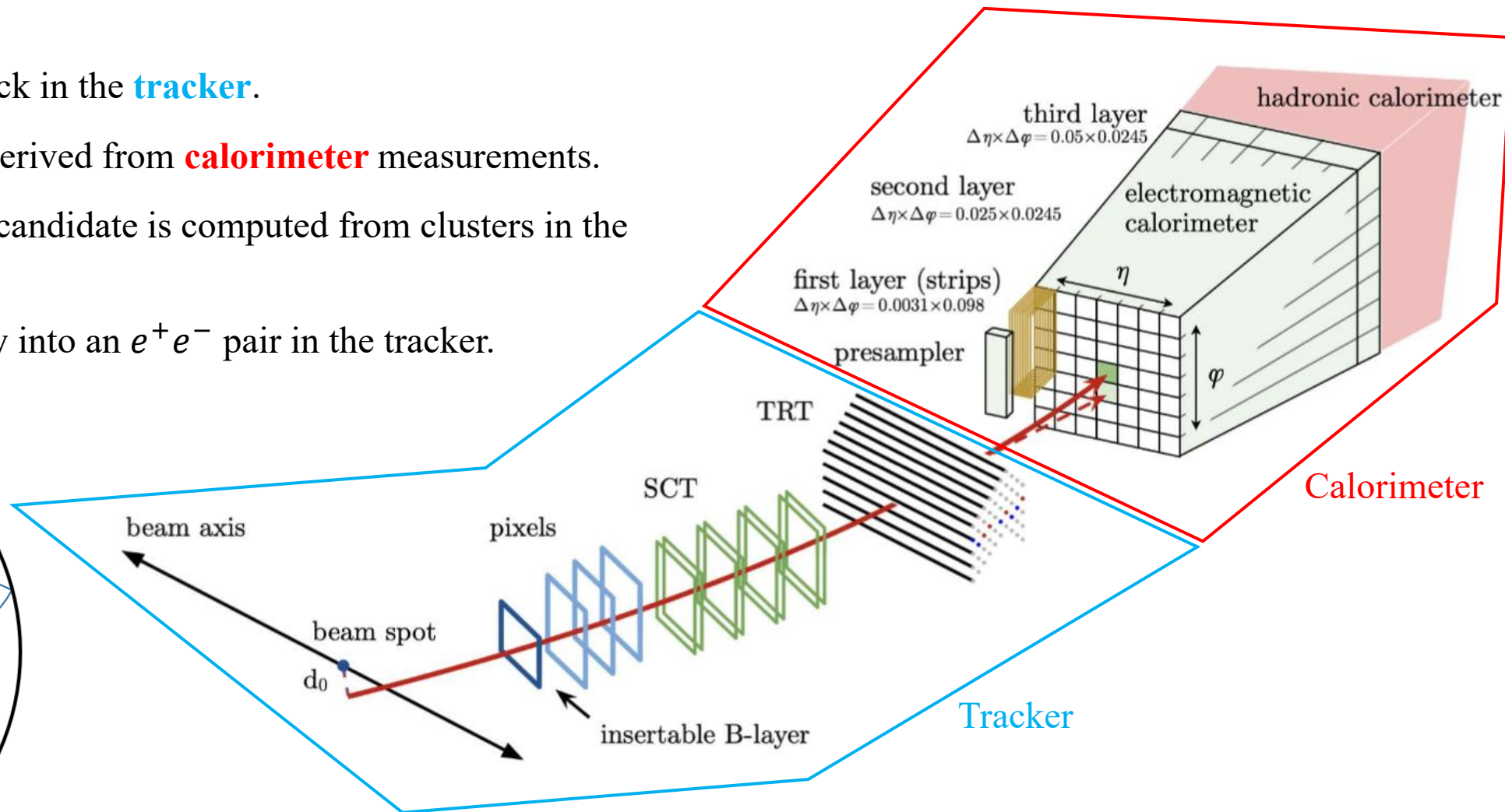
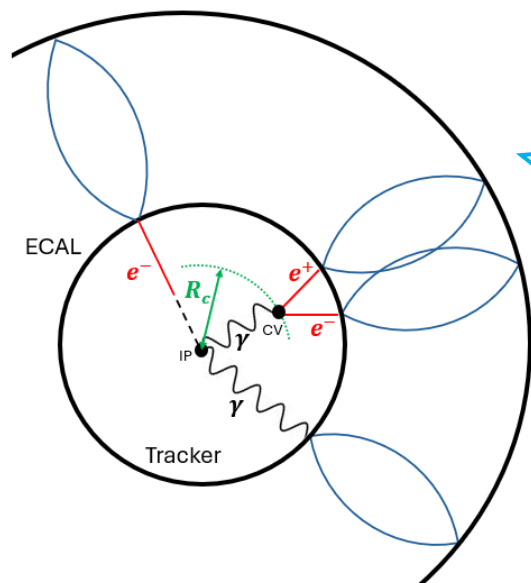
Photon Detection

Photons do not leave a track in the **tracker**.

Photon reconstruction is derived from **calorimeter** measurements.

The energy of the photon candidate is computed from clusters in the **EM calorimeter**.

Converted photons decay into an e^+e^- pair in the tracker.



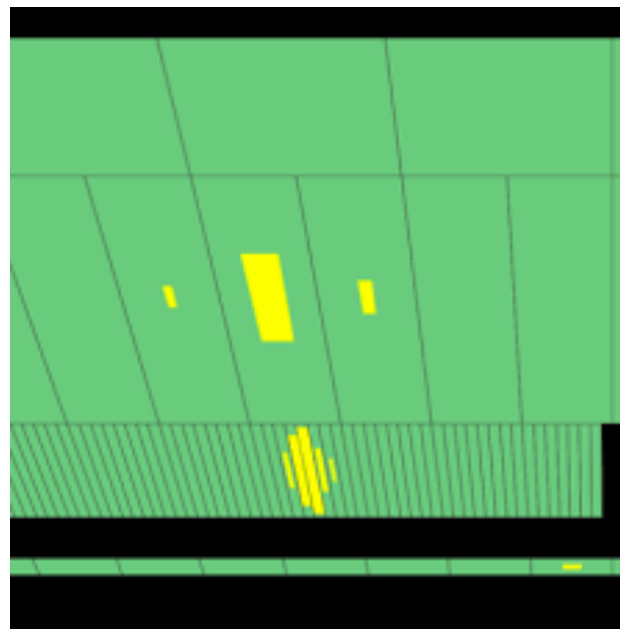
Photon Identification in ATLAS

Prompt (*True*) and Non-Prompt (*Fake*) Photons

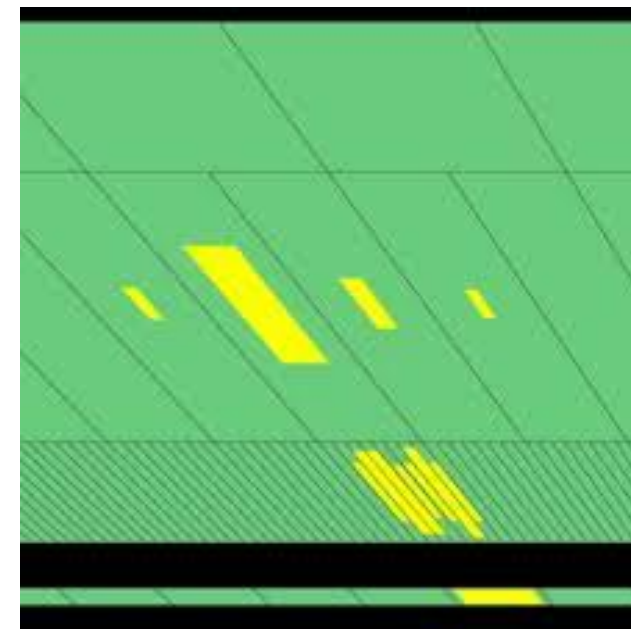
True photons: photons in the final state not coming from hadron decays ($H \rightarrow \gamma\gamma$)

Fake photons: photons coming from the decay of neutral mesons in QCD jets ($\pi^0 \rightarrow \gamma\gamma$)

- ➔ Highly collimated photon pair
- ➔ Dominant source of background in $H \rightarrow \gamma\gamma$



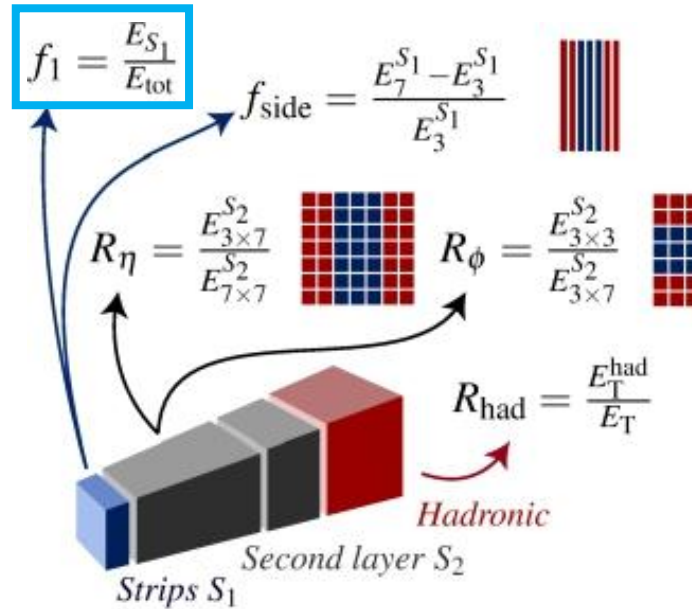
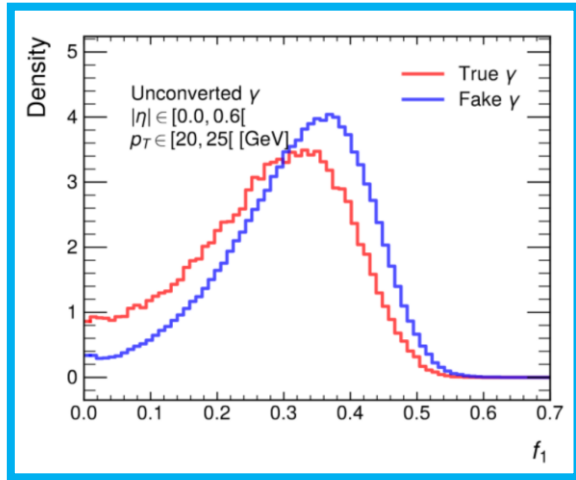
Display of **prompt-photon** event



Display of **non-prompt-photon** event

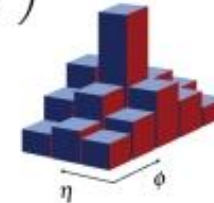
Photon Identification in ATLAS

Shower-Shape Variables for Photon Identification



$$w_{\eta_2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$$

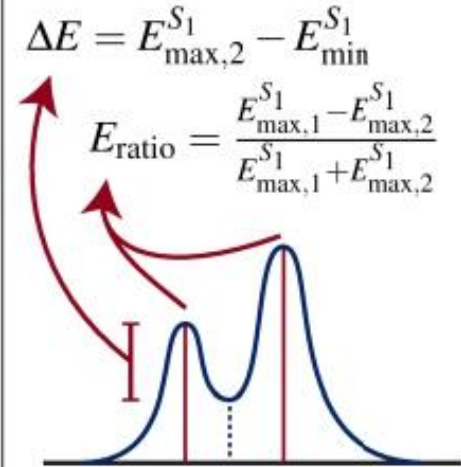
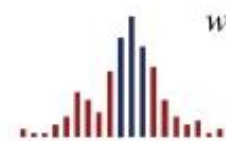
width in a 3×5 ($\Delta\eta \times \Delta\phi$) region of cells in S_2



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

w_{s3} uses 3×2 strips ($\eta \times \phi$)

$w_{s\text{tot}}$ is defined similarly but uses 20×2 strips



Legacy approach: rectangular cuts on shower-shape variables (SSV)

➡ **Loose** and **Tight** photon id.

New approach: use SSV to train a machine-learning (ML) algorithm (e. g. BDT)

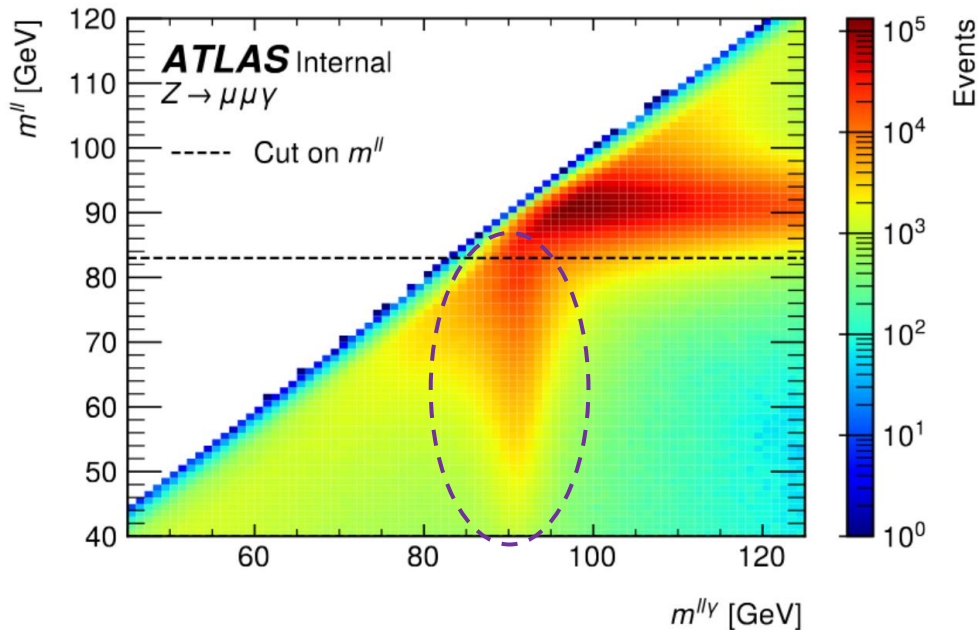
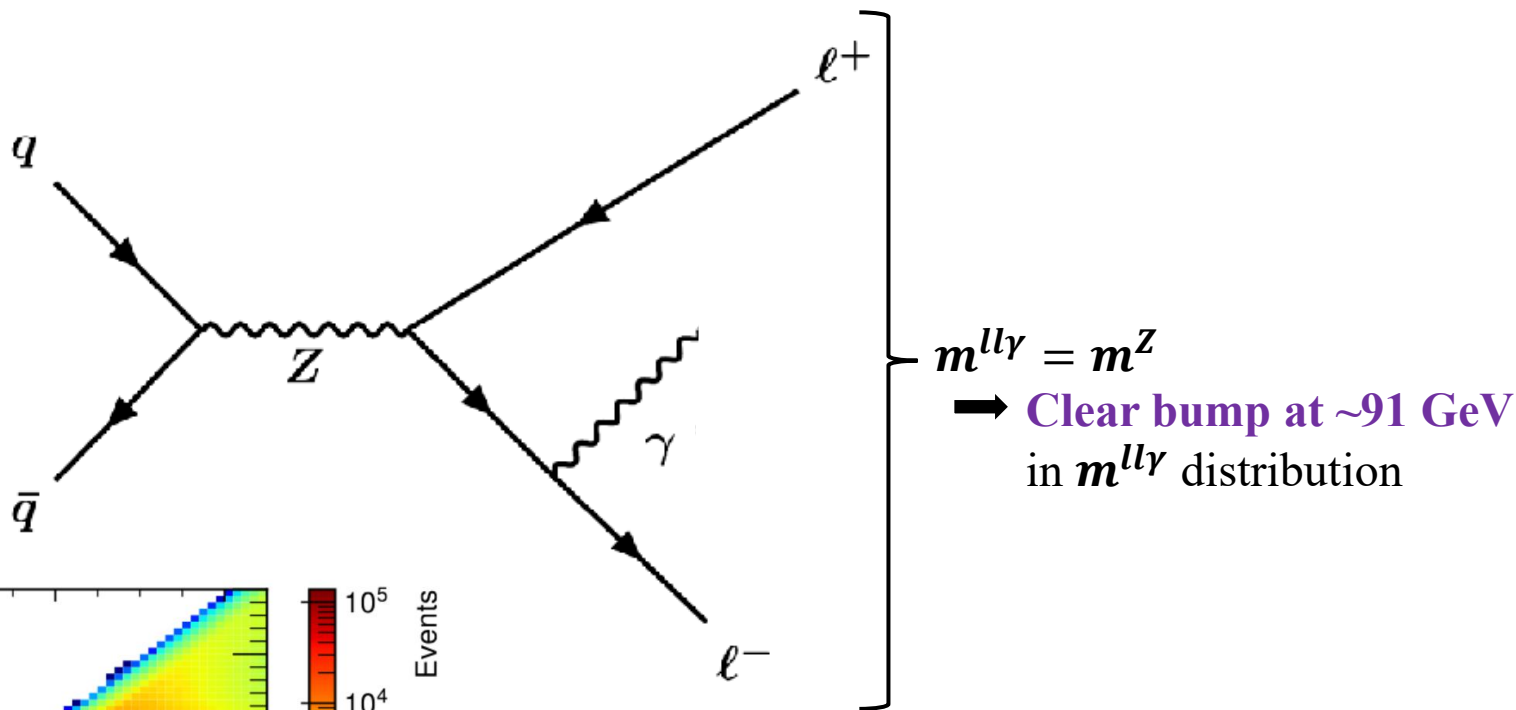
➡ **ML-based** photon id.

Must be calibrated to be used in analyses

Tight Identification Calibration with Radiative Z

Tight Identification Calibration with Radiative Z

Radiative Z Method



A cut on the m^{ll} distribution, alongside other selections on the leptons, removes initial state radiation (ISR)

- Pure samples with a true photon in the final state
- Photon id. efficiency can be estimated from the photon id. performance on these samples

Tight Identification Calibration with Radiative Z

Used Samples

Radiative Z ($Z \rightarrow ll\gamma$) data

2022, 2023 and 2024 (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

- Targeted $Z \rightarrow ll\gamma$ events
- Reconstructed from detector information
- Contains background (mainly $Z + j$ after ISR removal)

Tight Identification Calibration with Radiative Z

Used Samples

Radiative Z ($Z \rightarrow ll\gamma$) data

2022, 2023 and 2024 (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

Radiative Z ($Z \rightarrow ll\gamma$) MC

mc23a, e and d (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

- Simulated $Z \rightarrow ll\gamma$ events
- Truth information available
- Doesn't contain background

Tight Identification Calibration with Radiative Z

Used Samples

Contains $Z + j$ as dominant background source



Radiative Z ($Z \rightarrow ll\gamma$) data

2022, 2023 and 2024 (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

Radiative Z ($Z \rightarrow ll\gamma$) MC

mc23a, e and d (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

$Z + j$ MC

mc23a, e and d (Run 3)
 $Z(\rightarrow ee) + j$ and $Z(\rightarrow \mu\mu) + j$

- **Simulated $Z + j$ events**
- Low statistics
- ➔ Cannot be directly used for background estimation

Tight Identification Calibration with Radiative Z

Used Samples

Radiative Z ($Z \rightarrow ll\gamma$) data

2022, 2023 and 2024 (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

Radiative Z ($Z \rightarrow ll\gamma$) MC

mc23a, e and d (Run 3)
 $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$

Used for signal and background estimation

$Z + j$ MC

mc23a, e and d (Run 3)
 $Z(\rightarrow ee) + j$ and $Z(\rightarrow \mu\mu) + j$

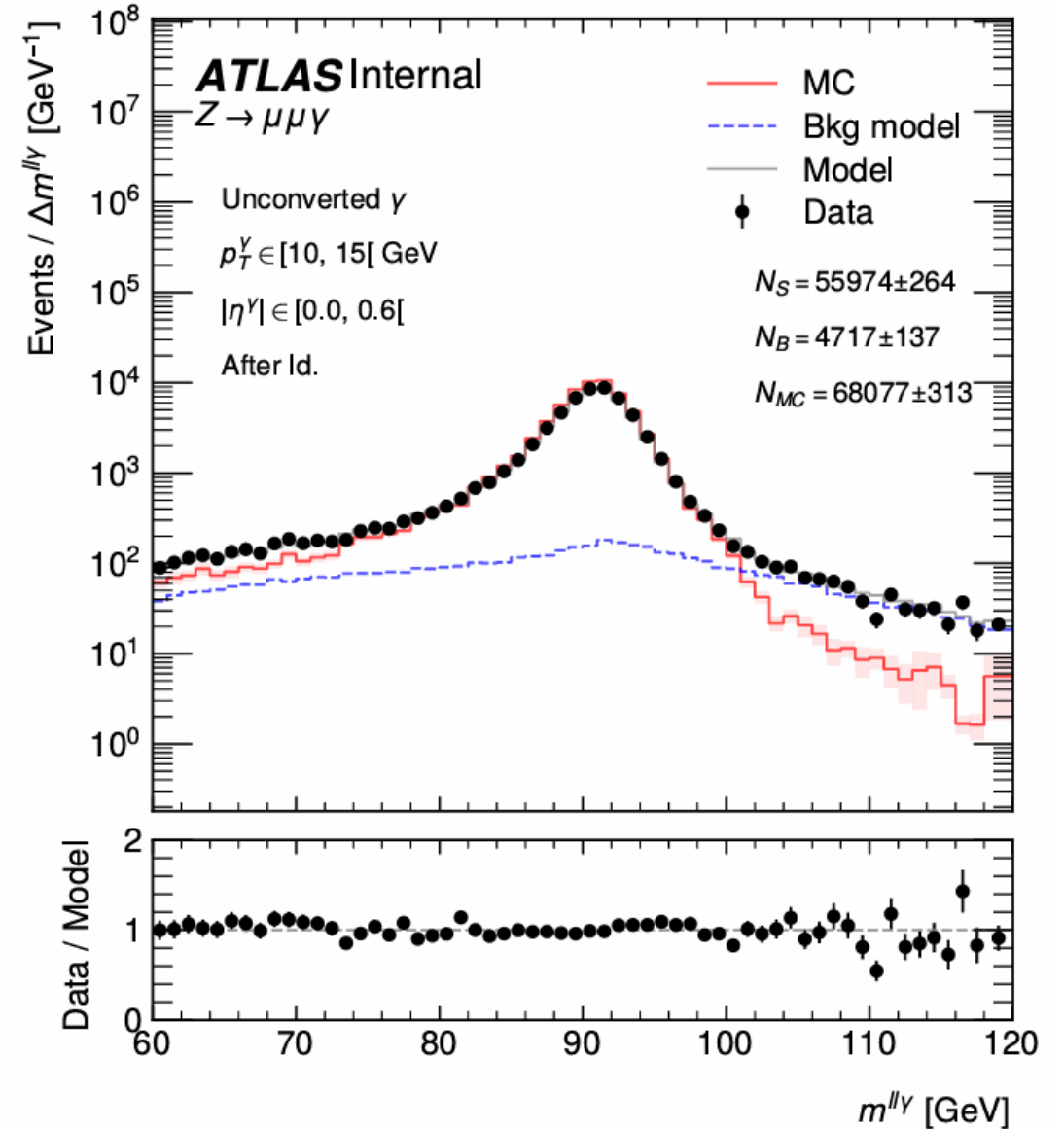
Used to cross check of the
estimated background

Tight Identification Calibration with Radiative Z

General Procedure

We use the distribution of the invariant mass $m^{ll\gamma}$ in $Z \rightarrow ll\gamma$ decays to estimate the number of signal events before and after Tight id.

$$\epsilon_{\text{Id.}} = \frac{N_S^{\text{Id.}}}{N_S}$$



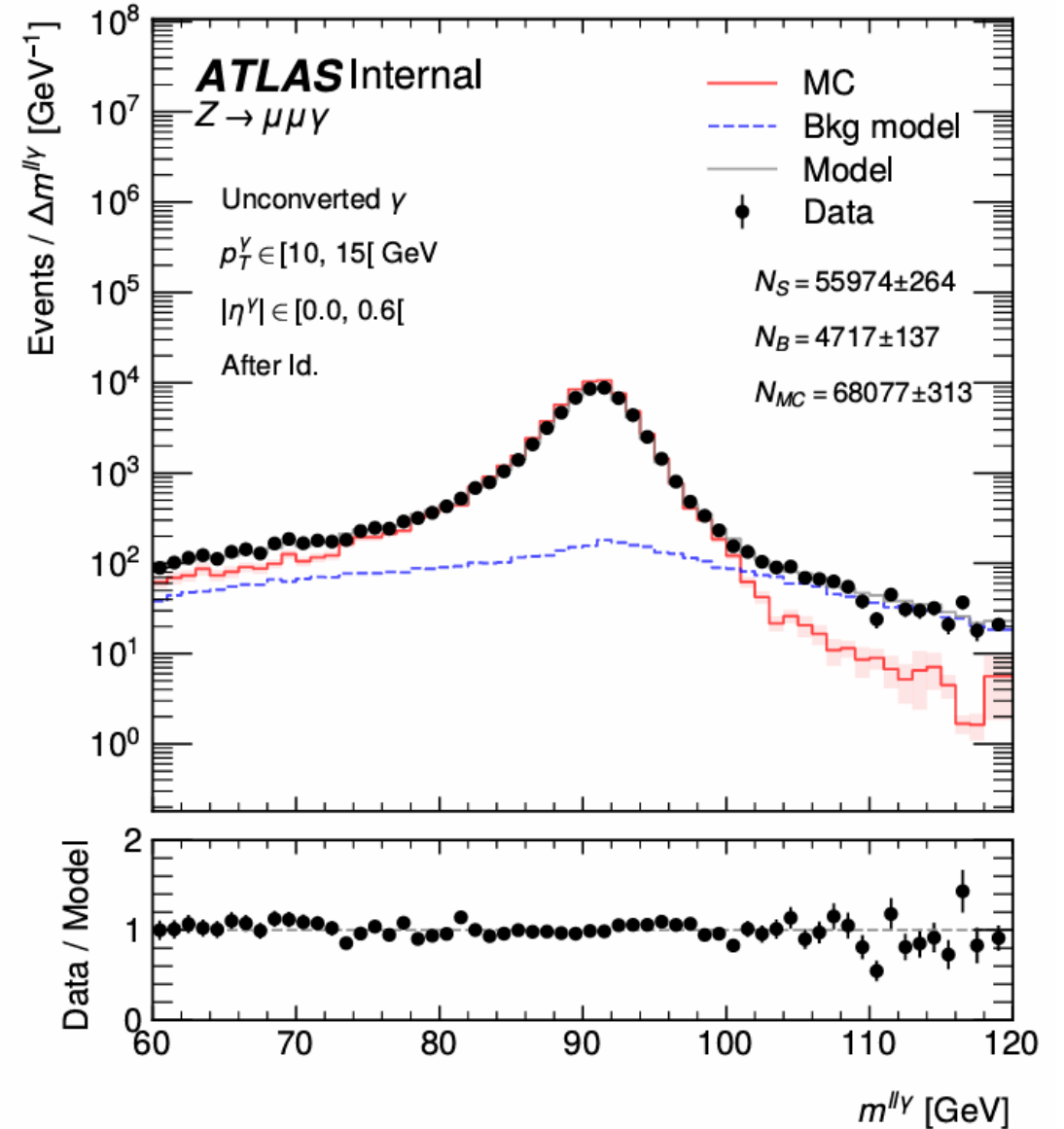
Tight Identification Calibration with Radiative Z

General Procedure

We use the distribution of the invariant mass $m^{ll\gamma}$ in $Z \rightarrow ll\gamma$ decays to estimate the number of signal events before and after Tight id.

$$\epsilon_{\text{Id.}} = \frac{N_S^{\text{Id.}}}{N_S}$$

In MC the number of signal events is obtained by integrating the $m^{ll\gamma}$ distribution.



Tight Identification Calibration with Radiative Z

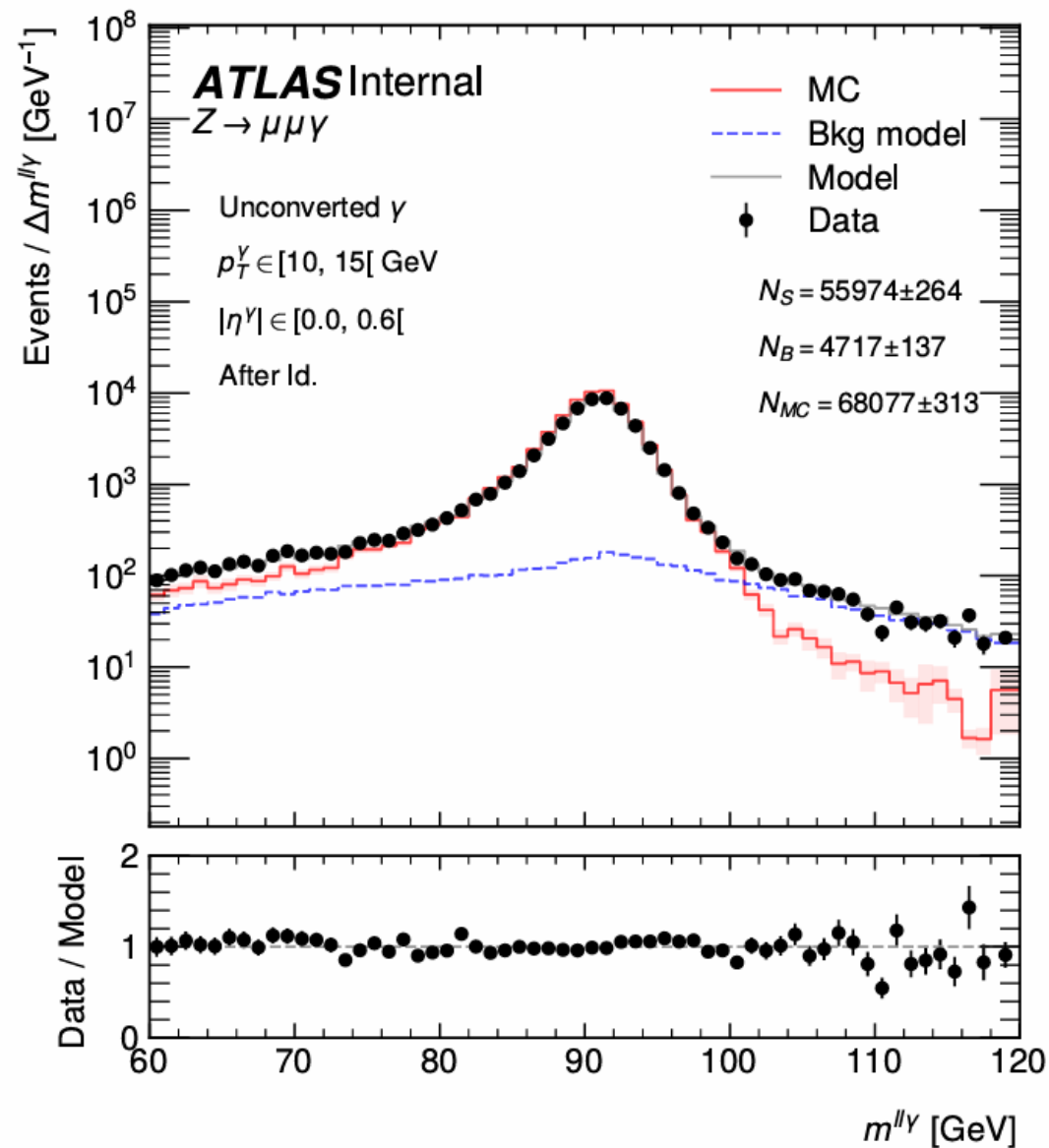
General Procedure

We use the distribution of the invariant mass $m^{l\ell\gamma}$ in $Z \rightarrow l\ell\gamma$ decays to estimate the number of signal events before and after Tight id.

$$\epsilon_{\text{Id.}} = \frac{N_S^{\text{Id.}}}{N_S}$$

In MC the number of signal events is obtained by integrating the $m^{l\ell\gamma}$ distribution.

In data the number of signal events is estimated with a template fit where $\text{PDF} = \text{PDF}_{\text{MC}} + \text{PDF}_{\text{Bkg}}$ is fitted to the data and yields N_S and N_B are the free parameters of the fit.



Tight Identification Calibration with Radiative Z

General Procedure

We use the distribution of the invariant mass $m^{ll\gamma}$ in $Z \rightarrow ll\gamma$ decays to estimate the number of signal events before and after Tight id.

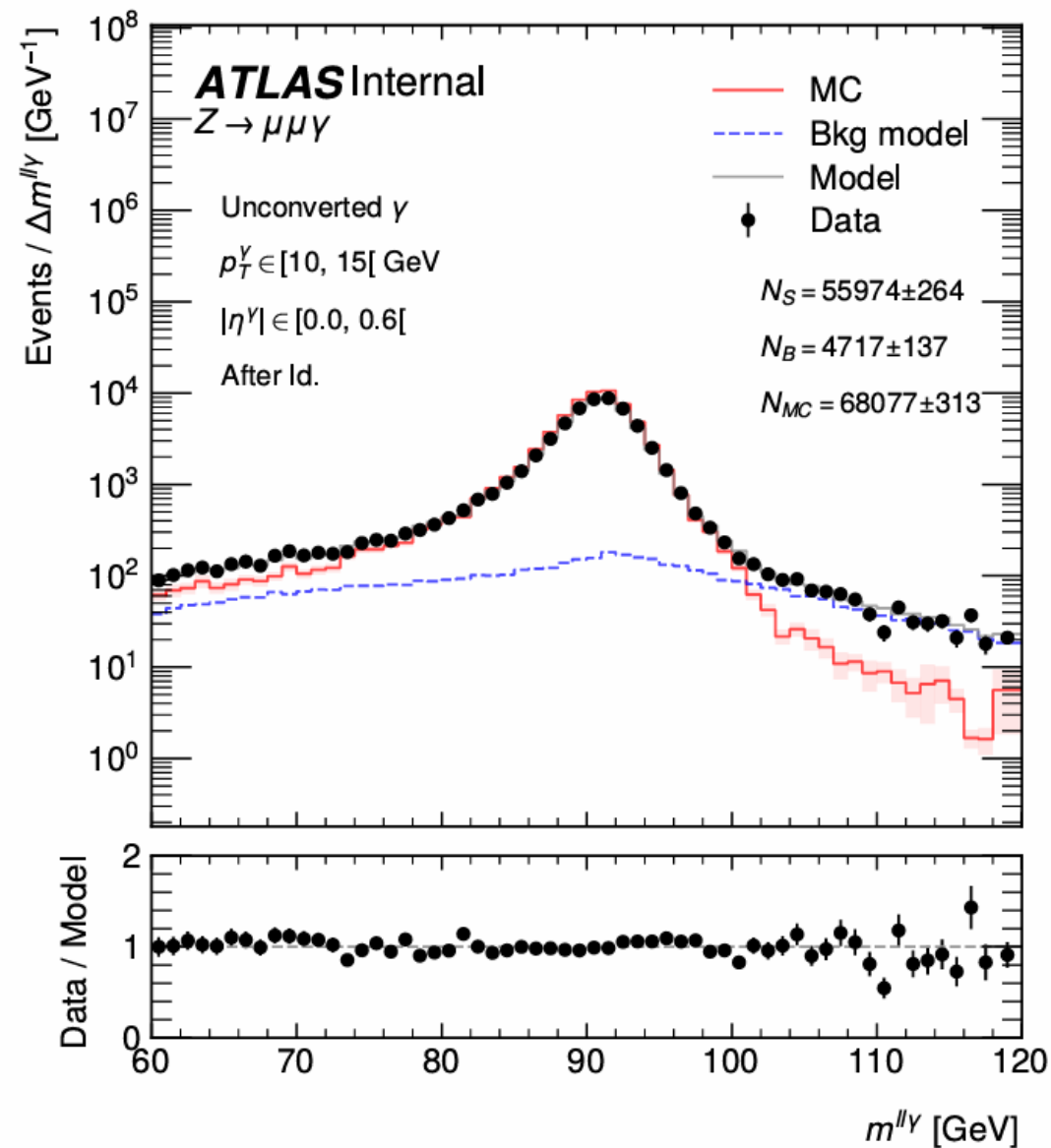
$$\epsilon_{\text{Id.}} = \frac{N_S^{\text{Id.}}}{N_S}$$

In MC the number of signal events is obtained by integrating the $m^{ll\gamma}$ distribution.

In data the number of signal events is estimated with a template fit where $\text{PDF} = \text{PDF}_{\text{MC}} + \text{PDF}_{\text{Bkg}}$ is fitted to the data and yields N_S and N_B are the free parameters of the fit.

Scale factors to correct the MC difference w. r. t. data:

$$SF = \frac{\epsilon_{\text{Id.}}^{\text{Data}}}{\epsilon_{\text{Id.}}^{\text{MC}}}$$



Tight Identification Calibration with Radiative Z

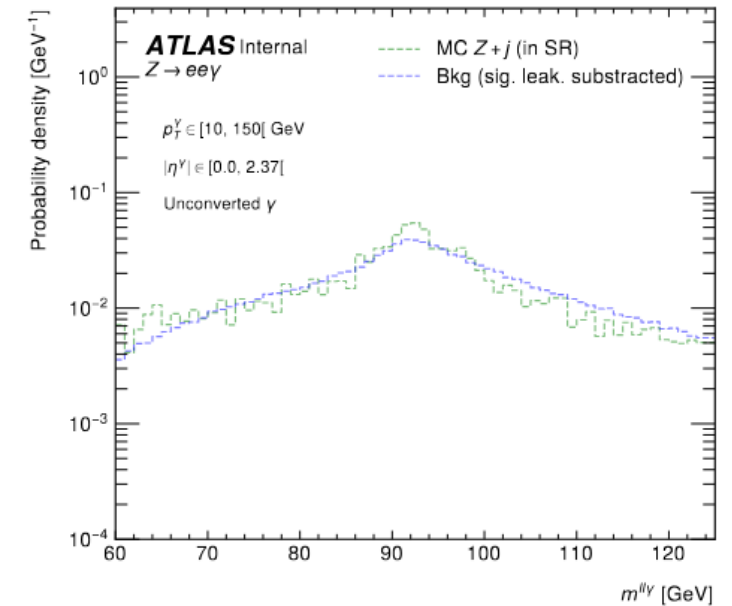
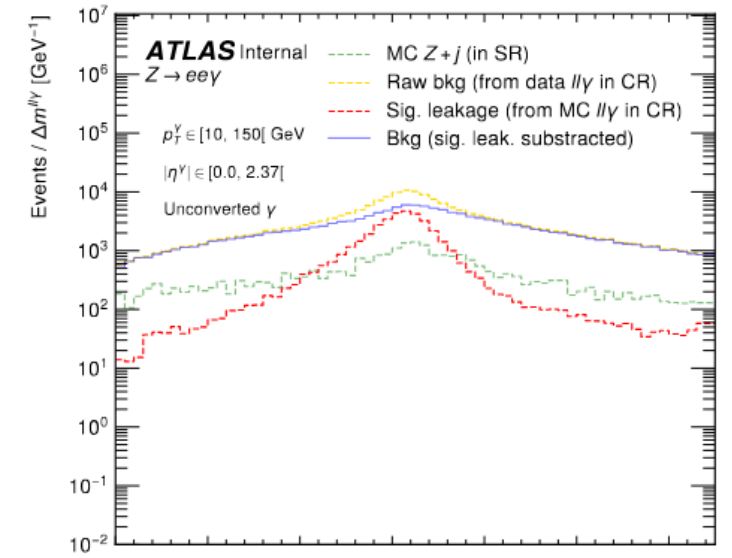
Background Template

Not enough statistics of MC $Z + j$

➡ We use a **data-driven** background estimation approach

*Control region (CR): anti-id.-and-anti-iso. region

**Signal region (SR): id.-and-iso. region



Tight Identification Calibration with Radiative Z

Background Template

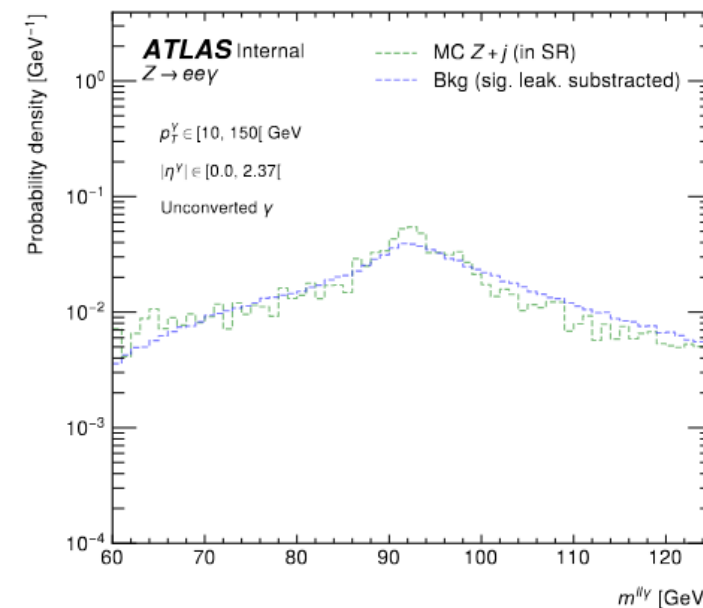
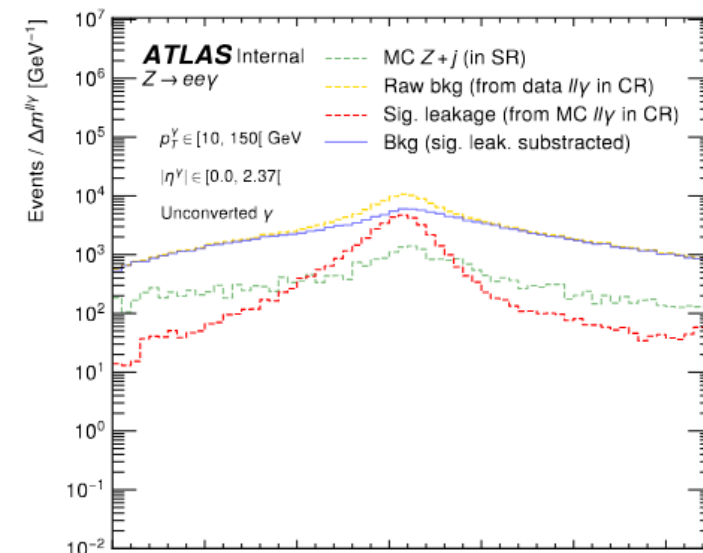
Not enough statistics of MC $Z + j$

➡ We use a **data-driven** background estimation approach

Data $Z \rightarrow ll\gamma$ in CR* ➡ **Raw background**
(= **Background** + **signal leakage**)

*Control region (CR): anti-id.-and-anti-iso. region

**Signal region (SR): id.-and-iso. region



Tight Identification Calibration with Radiative Z

Background Template

Not enough statistics of MC $Z + j$

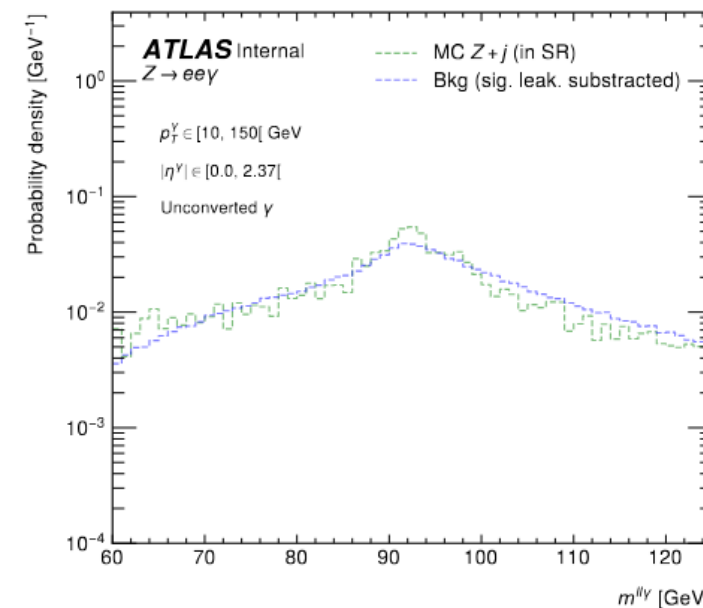
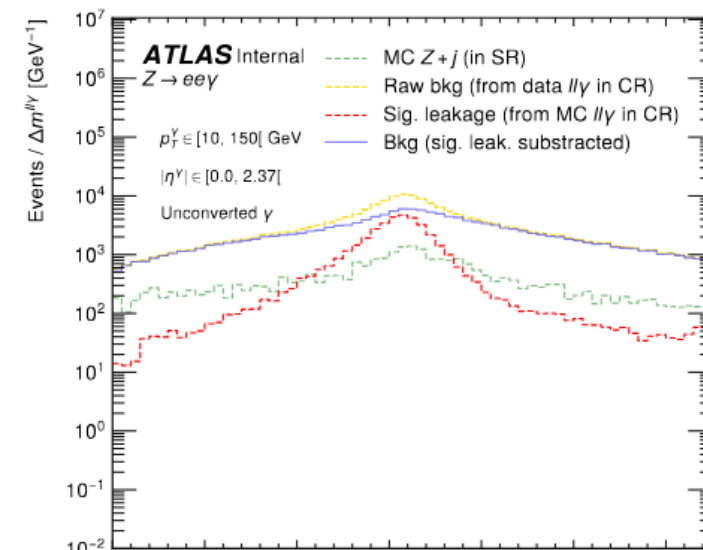
➡ We use a **data-driven** background estimation approach

Data $Z \rightarrow ll\gamma$ in CR* ➡ **Raw background**
(= **Background** + **signal leakage**)

MC $Z \rightarrow ll\gamma$ in CR* ➡ **Signal leakage**

*Control region (CR): anti-id.-and-anti-iso. region

**Signal region (SR): id.-and-iso. region



Tight Identification Calibration with Radiative Z

Background Template

Not enough statistics of MC $Z + j$

➡ We use a **data-driven** background estimation approach

Data $Z \rightarrow ll\gamma$ in CR* ➡ **Raw background**
(= **Background** + **signal leakage**)

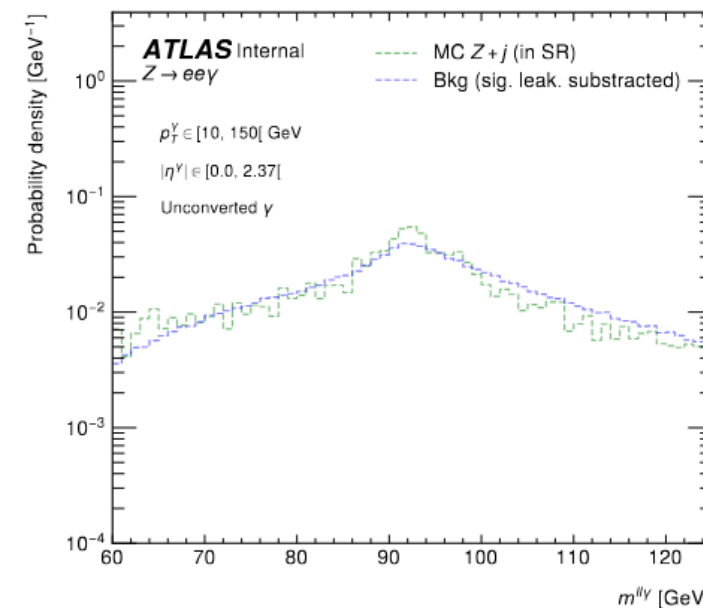
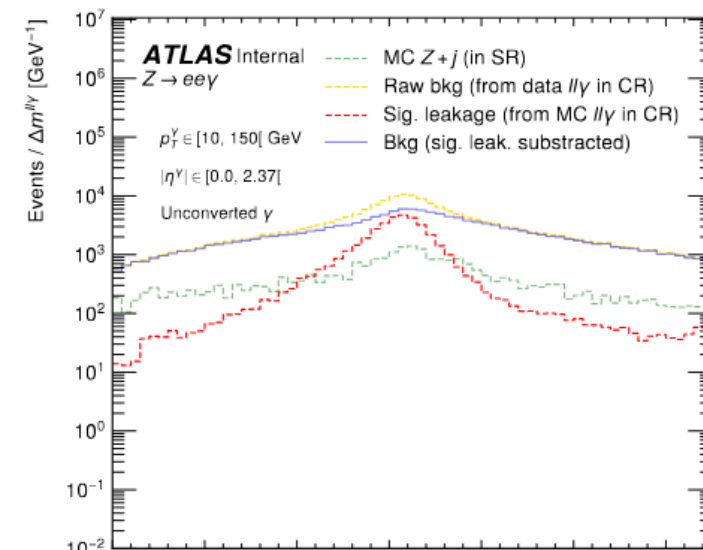
MC $Z \rightarrow ll\gamma$ in CR* ➡ **Signal leakage**

Raw background - **Signal leakage** = **Background template used for the fit**

➡ Shape to be compared with MC $Z + j$ in SR**

*Control region (CR): anti-id.-and-anti-iso. region

**Signal region (SR): id.-and-iso. region



Tight Identification Calibration with Radiative Z

Events Separation

Events are separated according to:

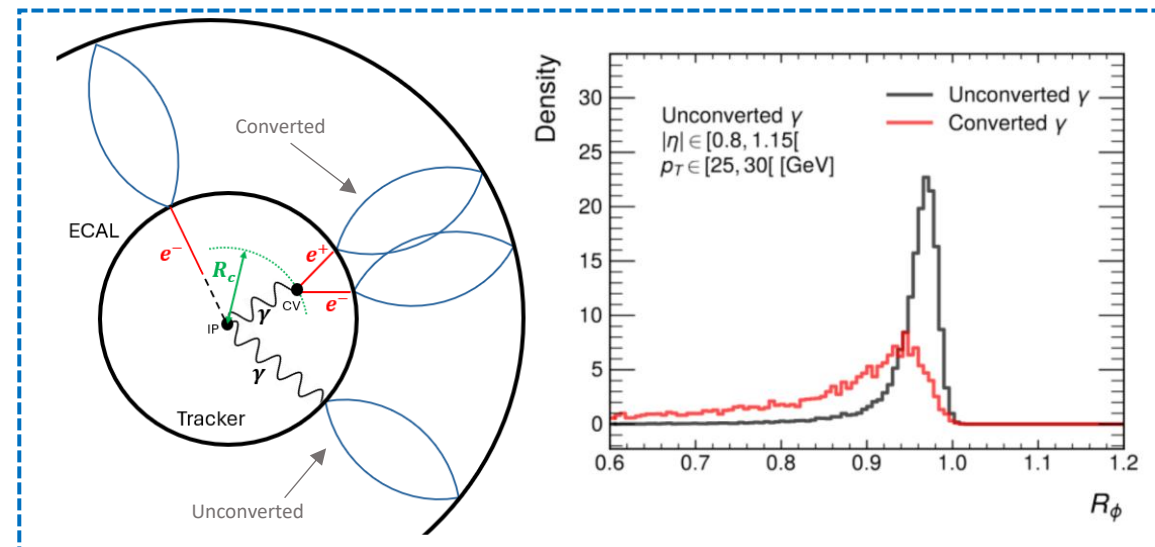
Decay type $Z \rightarrow ee\gamma$ or $Z \rightarrow \mu\mu\gamma$

Conversion status Converted or unconverted photons

$|\eta^\gamma|$ region $[0, 0.6, 0.8, 1.37[\text{ U }]1.52, 1.81, 2.01, 2.37[$

p_T^γ region $[10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 100[\text{ GeV}$

So the efficiency is estimated in **conversion- $|\eta^\gamma|$ - p_T^γ bins** for $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$ and is then combined to give an efficiency for $Z \rightarrow ll\gamma$.



Combined efficiency

$$\varepsilon = \frac{\frac{\varepsilon_e}{\sigma_e^2} + \frac{\varepsilon_\mu}{\sigma_\mu^2}}{\frac{1}{\sigma_e^2} + \frac{1}{\sigma_\mu^2}} \quad \sigma_\varepsilon = \frac{1}{\sqrt{\frac{1}{\sigma_e^2} + \frac{1}{\sigma_\mu^2}}}$$

Tight Identification Calibration with Radiative Z

Procedure Parameters

Parameters to vary for sys. unc. estimation

Fit Region (FR)

- $m^{l\gamma} \in [60, 120]$ GeV (nom.)
- $m^{l\gamma} \in [46, 125]$ GeV
- $m^{l\gamma} \in [70, 110]$ GeV

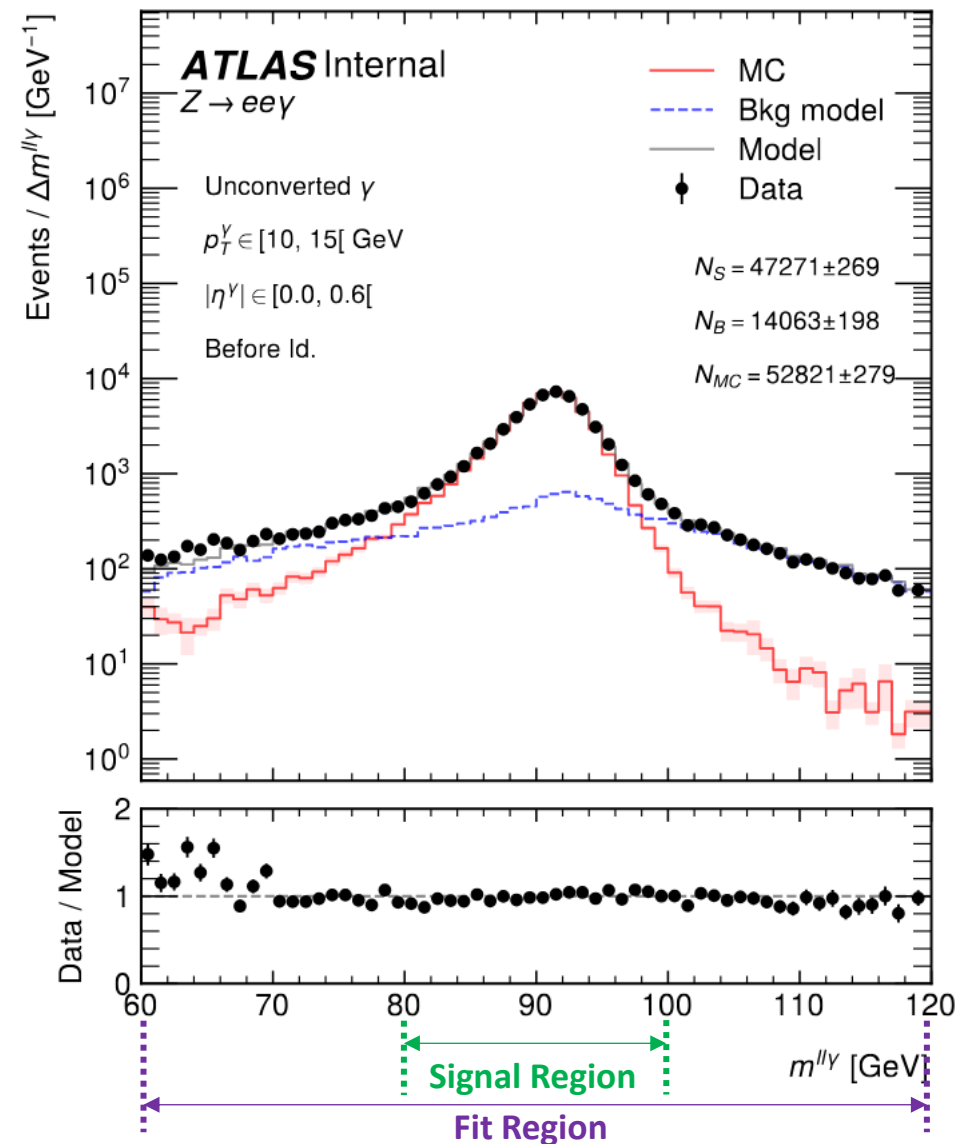
Signal Region (SR)

- $m^{l\gamma} \in [80, 100]$ GeV (nom.)
- $m^{l\gamma} \in [70, 110]$ GeV
- $m^{l\gamma} \in [85, 95]$ GeV

Bkg Template Shape (BS)

- Data - sig. leak. in CR (nom.)
- (Data - sig. leak. In CR) reweighted with MC Z + j in SR

➡ The maximum deviation from the nominal value is taken as a sys. uncertainty.



Tight Identification Calibration with Radiative Z

Results

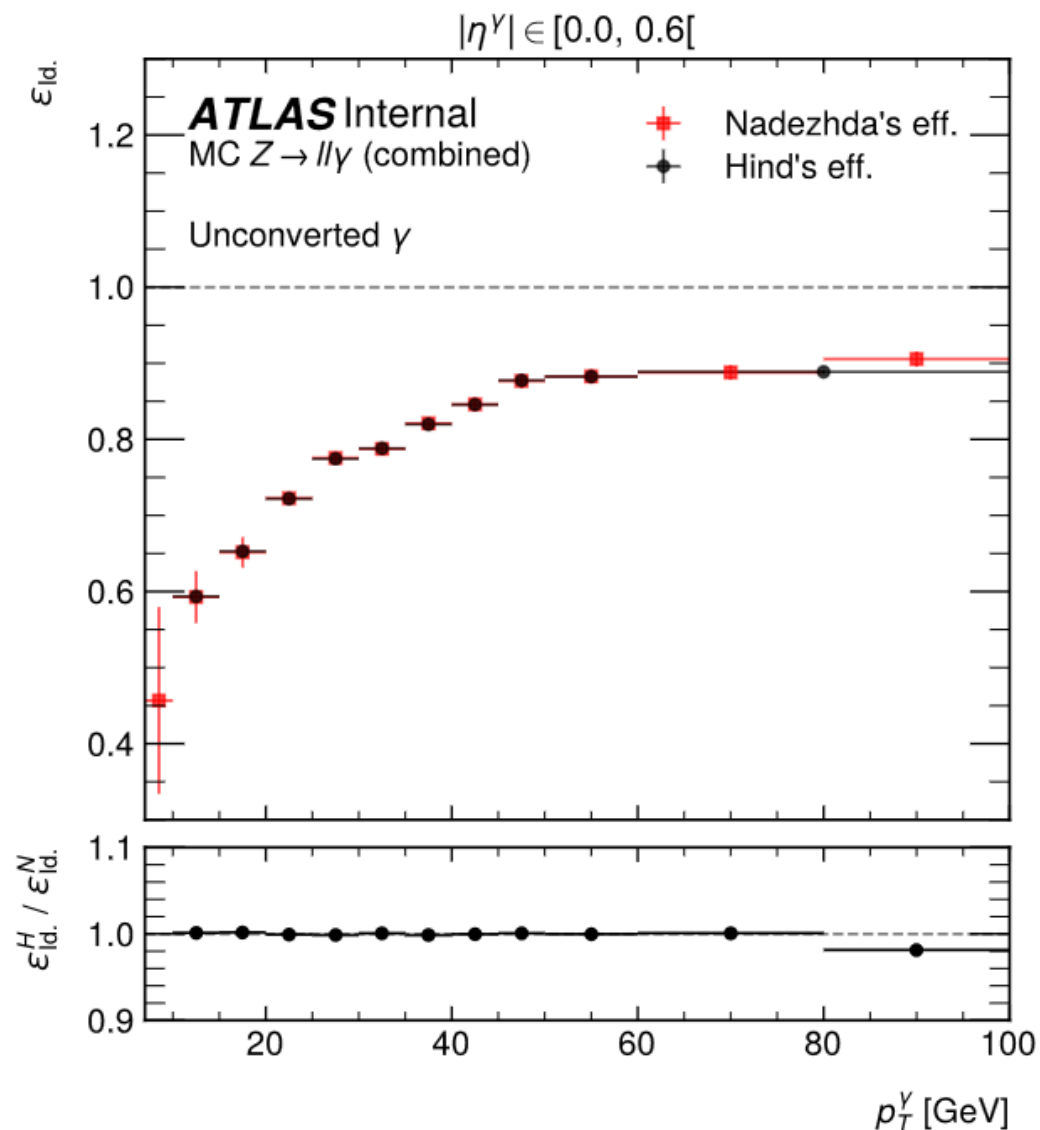
The estimated efficiencies are to be compared with the consolidated Run 3 efficiencies (computed by Nadezhda)

➡ Good marching between the two

~1-3 % total uncertainty on efficiency

Dominant source of uncertainty: systematic uncertainty due to background shape

➡ $\epsilon_{\text{Id.}}^H / \epsilon_{\text{Id.}}^N$ with an error bar of $\sigma_{\epsilon_{\text{Id.}}^H} / \epsilon_{\text{Id.}}^N$



Conclusion and Next Steps

Conclusion and Next Steps

Conclusion

- Generic procedure of photon id. efficiency estimation completed
- Efficiencies estimated with this procedure in agreement with consolidated Run 3 efficiencies

Next Steps

- Produce new radiative-Z ntuples with a BDT score associated to the photon (ongoing)
- Estimate efficiency of BDT-based photon id. for 1 working point
- Calibrate entire BDT score
- Move on to the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis by evaluating impact of new photon id. BDT and uncertainties on current analysis (targeting Moriond 2027)
- Begin to work on alternative categorization simultaneously fitting $HH \rightarrow b\bar{b}\gamma\gamma$ and single $H \rightarrow \gamma\gamma$
- On longer term (legacy analysis), possibly measure at the same time H and HH STXS

Conclusion and Next Steps

Upcoming Presentations in Workshops

- e/γ workshop (Edinburgh, 15th-19th June)
- Di-Higgs workshop (Grenoble, 8th-10th July)

Trainings

- Completed 61/120 h of training:
 - Encadrer efficacement des TD
 - MOOC Doctorat et poursuite de carrière/ PhD and career development
 - ATLAS Software Tutorials
- Completed 29 h of teaching (at IUT):
 - Introduction à l'analyse de données
 - TP de mathématiques
- Will be participating the European school of HEP (Ostroda, 19th June-2 July)

Backup