Optimisation of the Higgs boson mass measurement in the diphoton channel with the ATLAS detector

Supervisors
Prof. Leonardo Carminati
Dott. Davide Mungo
Dott. Ruggero Turra

Candidate
Alessandro Demela

Università degli Studi di Milano

September 28, 2020

The Higgs boson

The Higgs boson

About the Higgs

- Particle in the SM
  - Theory of elementary particles and their interactions
- Discovered in 2012 by ATLAS and CMS collaborations
- Responsible for giving mass to all elementary particles
  - "Higgs mechanism"

The Higgs boson

About the Higgs

- Particle in the SM
  - Theory of elementary particles and their interactions
- Discovered in 2012 by ATLAS and CMS collaborations
- Responsible for giving mass to all elementary particles
  - "Higgs mechanism"

Masses are not predicted by SM $\Rightarrow$ measure m_H

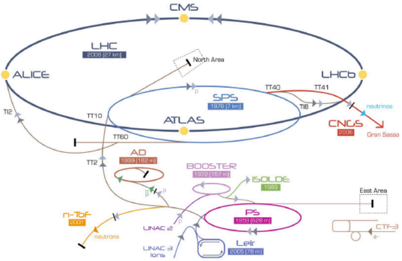
- Higgs properties $(\sigma, \mathrm{BR} \dots)$ are given as function of its mass
  - Compare them with predictions
- m_H, m_W, m_t are related
  - TEST SM consistency

Higgs boson @ LHC

Higgs boson @ LHC

The LHC is a *pp* collider at CERN

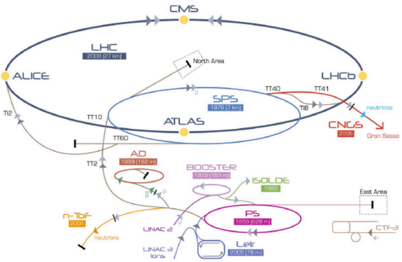
- Operates at 13 TeV
- Higgs production achieved in 4 main ways

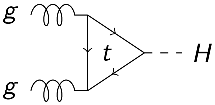


Higgs boson @ LHC

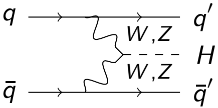
The LHC is a *pp* collider at CERN

- Operates at 13 TeV
- Higgs production achieved in 4 main ways

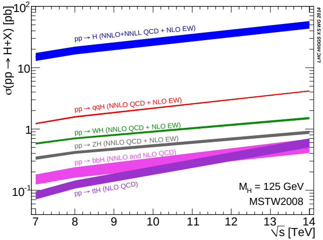


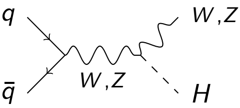


ggH

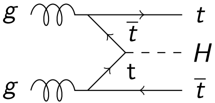


VBF





VH

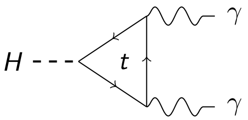


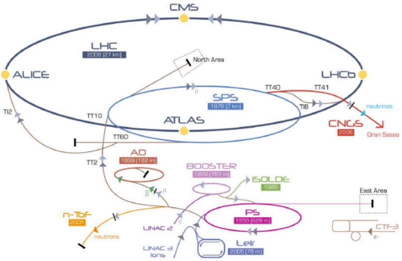
$t\bar{t}H$

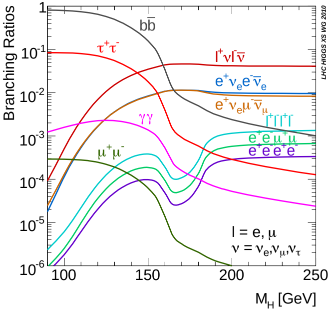
Higgs boson @ LHC

The LHC is a *pp* collider at CERN

- Operates at 13 TeV
- Higgs production achieved in 4 main ways
- Decays to every SM massive particles
  - $H \to \gamma\gamma$ with loop



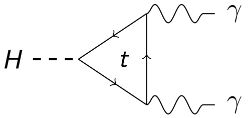




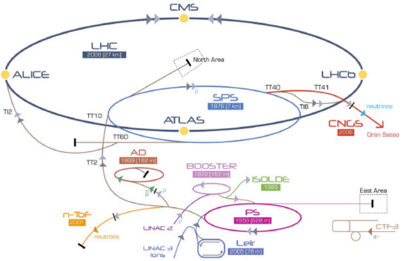
Higgs boson @ LHC

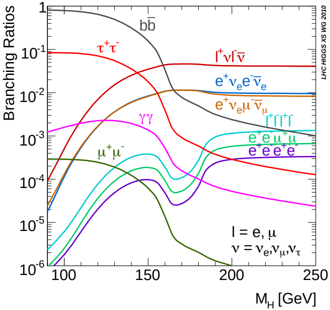
The LHC is a *pp* collider at CERN

- Operates at 13 TeV
- Higgs production achieved in 4 main ways
- Decays to every SM massive particles
  - $H \to \gamma\gamma$ with loop



- At ATLAS m_H is measured with $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$
  - Great resolution $\sim$1 GeV
  - Low background





Mass measurement

Mass measurement

Latest mass measurement from ATLAS @ 36.1 fb^{-1}

$$H \to \gamma\gamma \qquad m_H = 124.93 \pm 0.21(\text{stat}) \pm 0.34(\text{syst})\,\text{GeV}$$
$$H \to ZZ^* \to 4\ell \qquad m_H = 124.79 \pm 0.36(\text{stat}) \pm 0.09(\text{syst})\,\text{GeV}$$
$$\text{Combined} \qquad m_H = 124.86 \pm 0.18(\text{stat}) \pm 0.20(\text{syst})\,\text{GeV}$$

Mass measurement

Latest mass measurement from ATLAS @ $36.1\,\text{fb}^{-1}$

$H \to \gamma\gamma$	$m_H = 124.93 \pm 0.21(\text{stat}) \pm 0.34(\text{syst})\,\text{GeV}$
$H \to ZZ^* \to 4\ell$	$m_H = 124.79 \pm 0.36(\text{stat}) \pm 0.09(\text{syst})\,\text{GeV}$
Combined	$m_H = 124.86 \pm 0.18(\text{stat}) \pm 0.20(\text{syst})\,\text{GeV}$

- In the **$H \to \gamma\gamma$ channel** systematic uncertainty dominates
  - Measurement @ $139\,\text{fb}^{-1}$ will be more
  - $H \to ZZ^* \to 4\ell$ is almost systematic free $\Rightarrow$ still competitive

Mass measurement

Latest mass measurement from ATLAS @ $36.1\,\text{fb}^{-1}$

$H \to \gamma\gamma$	$m_H = 124.93 \pm 0.21(\text{stat}) \pm 0.34(\text{syst})\,\text{GeV}$
$H \to ZZ^* \to 4\ell$	$m_H = 124.79 \pm 0.36(\text{stat}) \pm 0.09(\text{syst})\,\text{GeV}$
Combined	$m_H = 124.86 \pm 0.18(\text{stat}) \pm 0.20(\text{syst})\,\text{GeV}$

- In the **$H \to \gamma\gamma$ channel** systematic uncertainty dominates
  - Measurement @ $139\,\text{fb}^{-1}$ will be more
  - $H \to ZZ^* \to 4\ell$ is almost systematic free $\Rightarrow$ still competitive
- Systematic uncertainty must be reduced by a factor of 2
  - Precision from actual 0.27 % to 0.1 %

Mass measurement

Latest mass measurement from ATLAS @ 36.1 fb^{-1}

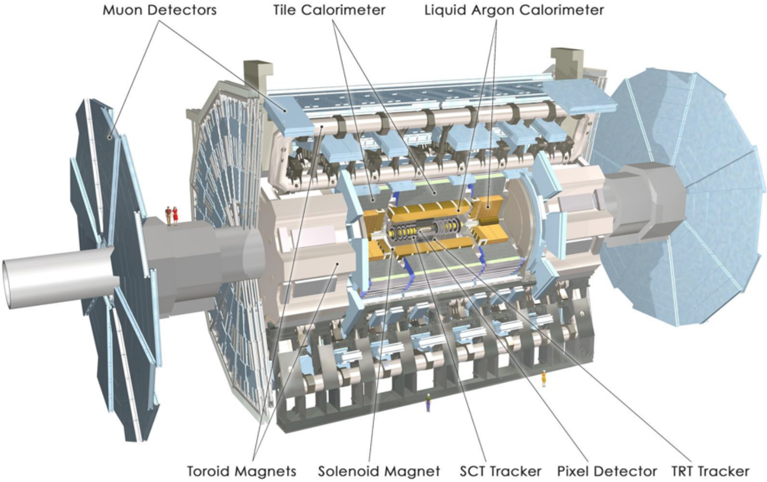
$H \to \gamma\gamma$	$m_H = 124.93 \pm 0.21(\text{stat}) \pm 0.34(\text{syst})$ GeV
$H \to ZZ^* \to 4\ell$	$m_H = 124.79 \pm 0.36(\text{stat}) \pm 0.09(\text{syst})$ GeV
Combined	$m_H = 124.86 \pm 0.18(\text{stat}) \pm 0.20(\text{syst})$ GeV

- In the $\boldsymbol{H \to \gamma\gamma}$ **channel** systematic uncertainty dominates
  - Measurement @ 139 fb^{-1} will be more
  - $H \to ZZ^* \to 4\ell$ is almost systematic free $\Rightarrow$ still competitive
- Systematic uncertainty must be reduced by a factor of 2
  - Precision from actual 0.27 % to 0.1 %

$\to$ **Main focus of this thesis**

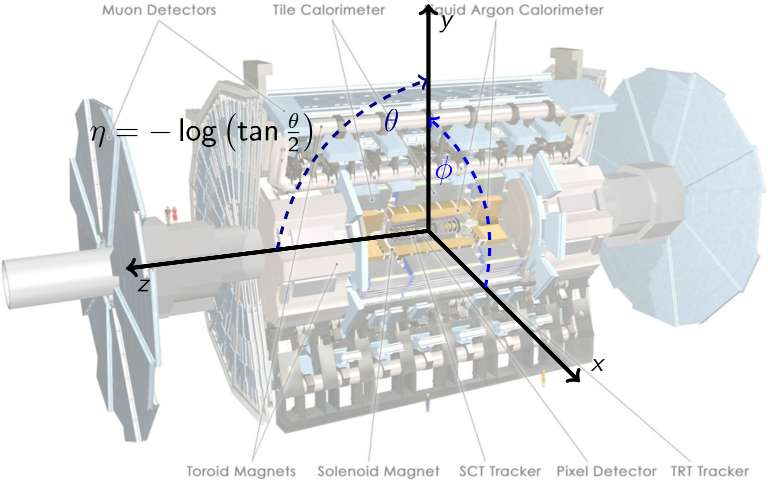
The ATLAS detector

ATLAS is a *general purpose* detector aiming to capture the products from *pp* and heavy ions collisions at the LHC

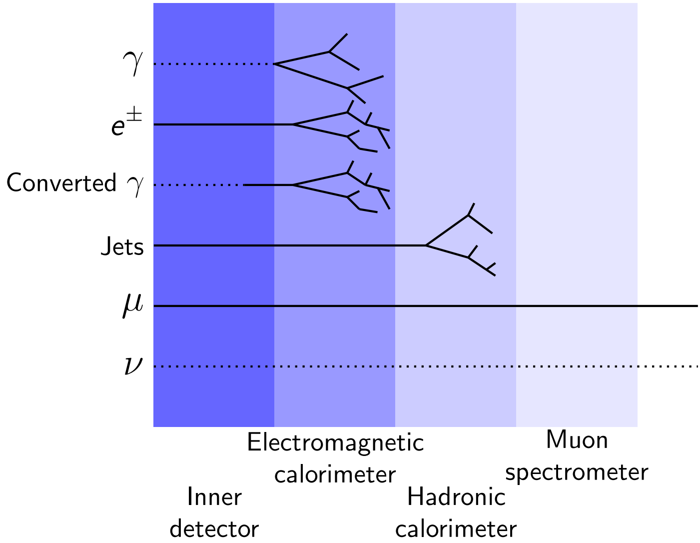


The ATLAS detector

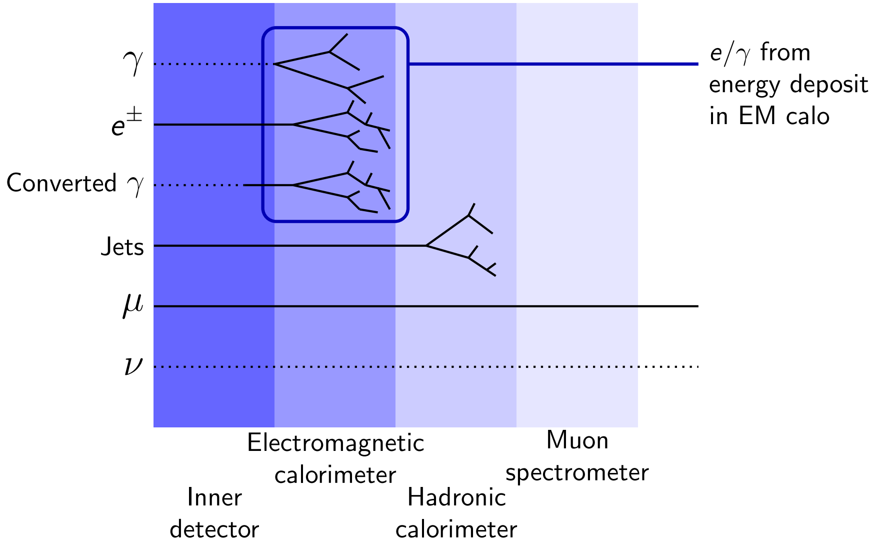
ATLAS is a *general purpose* detector aiming to capture the products from *pp* and heavy ions collisions at the LHC



The ATLAS detector



The ATLAS detector



$H \to \gamma\gamma$ mass analysis

$H \to \gamma\gamma$ mass analysis

- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]$GeV

$H \to \gamma\gamma$ mass analysis

- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]$GeV
- Events are classified in different categories

$H \to \gamma\gamma$ mass analysis

- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]$GeV
- Events are classified in different categories
- Signal Model
  - Fit MC signal sample
  - Resonant shape $\sim$125 GeV

$H \rightarrow \gamma\gamma$ mass analysis

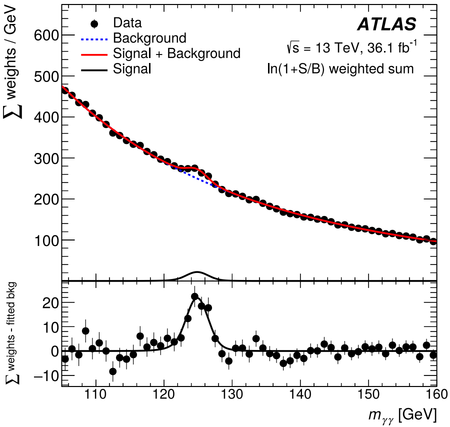
- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]$GeV
- Events are classified in different categories
- Signal Model
  - Fit MC signal sample
  - Resonant shape $\sim$125 GeV
- Background Model
  - Fit $m_{\gamma\gamma}$ in data in $[105, 120] \cup [130, 160]$ GeV
  - Main sources
    - QCD $\gamma\gamma$ continuum
    - Jets faking γ

$H \to \gamma\gamma$ mass analysis

- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]$GeV
- Events are classified in different categories
- Signal Model
  - Fit MC signal sample
  - Resonant shape $\sim$125 GeV
- Background Model
  - Fit $m_{\gamma\gamma}$ in data in $[105, 120] \cup [130, 160]$ GeV
  - Main sources
    - QCD $\gamma\gamma$ continuum
    - Jets faking γ
- Systematic Uncertainties
  - Generate MC signal variated sample

$H \to \gamma\gamma$ mass analysis

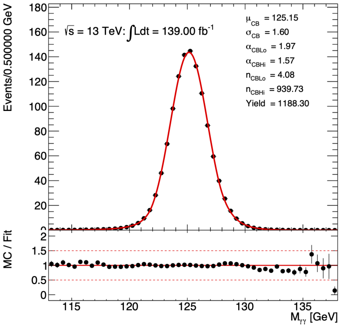
- 2 photons requested in final state
  - $m_{\gamma\gamma} \in [105, 160]\text{GeV}$
- Events are classified in different categories
- Signal Model
  - Fit MC signal sample
  - Resonant shape ∼125 GeV
- Background Model
  - Fit $m_{\gamma\gamma}$ in data in $[105, 120] \cup [130, 160]$ GeV
  - Main sources
    - QCD $\gamma\gamma$ continuum
    - Jets faking γ
- Systematic Uncertainties
  - Generate MC signal variated sample



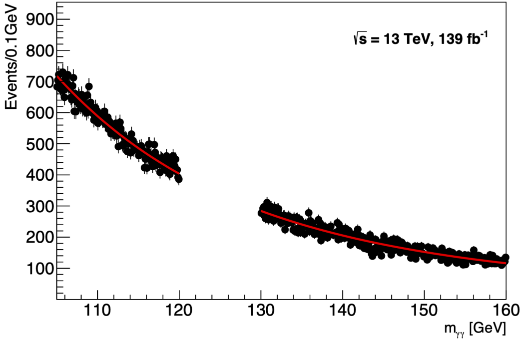
Fit result

Result for Signal and Background Fit in one event category

Signal Fit



Background Fit



Systematic uncertainties

Systematic uncertainties

Main contribution from γ energy determination

- EM calorimeter calibration with resonant processes
  - No candles $X \to \gamma\gamma$
  - Rely on e calibration

Systematic uncertainties

Main contribution from γ energy determination

- EM calorimeter calibration with resonant processes
  - No candles $X \to \gamma\gamma$
  - Rely on e calibration
- Series of 69 uncertainties including
  - Knowledge of the material before the EM calorimeter
  - EM calorimeter calibration
  - Effects of $e \to \gamma$ calibration extrapolation

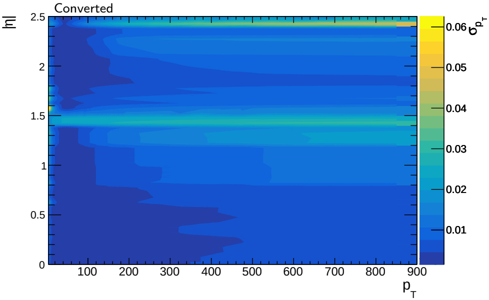
Systematic uncertainties

Main contribution from γ energy determination

- EM calorimeter calibration with resonant processes
  - No candles $X \rightarrow \gamma\gamma$
  - Rely on e calibration
- Series of 69 uncertainties including
  - Knowledge of the material before the EM calorimeter
  - EM calorimeter calibration
  - Effects of $e \rightarrow \gamma$ calibration extrapolation

How to evaluate (each systematic)

- Generate $m_{\gamma\gamma}^{\text{var}}$ distribution
  - From maps $f(p_{\text{T}}, \eta, \text{conv}) = \sigma_{p_{\text{T}}}$



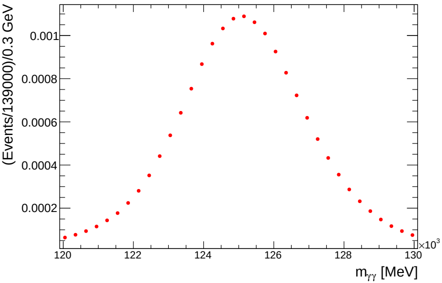
Systematic uncertainties

Main contribution from γ energy determination

- EM calorimeter calibration with resonant processes
  - No candles $X \to \gamma\gamma$
  - Rely on e calibration
- Series of 69 uncertainties including
  - Knowledge of the material before the EM calorimeter
  - EM calorimeter calibration
  - Effects of $e \to \gamma$ calibration extrapolation

How to evaluate (each systematic)

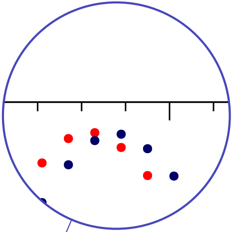
- Generate $m_{\gamma\gamma}^{\text{var}}$ distribution
  - From maps $f(p_T, \eta, \text{conv}) = \sigma_{p_T}$
- Generate $m_{\gamma\gamma}$ distribution MC samples



Systematic uncertainties

Main contribution from γ energy determination

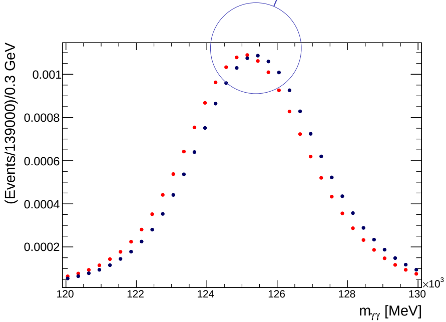
- EM calorimeter calibration with resonant processes
  - No candles $X \to \gamma\gamma$
  - Rely on e calibration
- Series of 69 uncertainties including
  - Knowledge of the material before the EM calorimeter
  - EM calorimeter calibration
  - Effects of $e \to \gamma$ calibration extrapolation



How to evaluate (each systematic)

- Generate $m_{\gamma\gamma}^{\text{var}}$ distribution
  - From maps $f(p_{\mathrm{T}}, \eta, \text{conv}) = \sigma_{p_{\mathrm{T}}}$
- Generate $m_{\gamma\gamma}$ distribution MC samples
- Evaluate

$$\delta = \frac{\langle m_{\gamma\gamma}^{\text{var}} \rangle}{\langle m_{\gamma\gamma} \rangle} - 1$$



Likelihood Fit

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} P\left(n_c | s_c(m_H, \boldsymbol{\theta}) + b_c\right) \prod_{i=1}^{n_c} f_c\left(m^i_{\gamma\gamma}, m_H, \boldsymbol{\theta}\right) K(\boldsymbol{\theta})$$

- Over 100 parameters are included

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} \boxed{P\left(n_c | s_c(m_H, \boldsymbol{\theta}) + b_c\right)} \prod_{i=1}^{n_c} f_c\left(m_{\gamma\gamma}^i, m_H, \boldsymbol{\theta}\right) K(\boldsymbol{\theta})$$

- Over 100 parameters are included
- Product for each category
  - Poissonian term
    - n_c events observed with $s_c + b_c$ expected

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} P\left(n_c | s_c\left(m_H, \boldsymbol{\theta}\right) + b_c\right) \boxed{\prod_{i=1}^{n_c} f_c\left(m^i_{\gamma\gamma}, m_H, \boldsymbol{\theta}\right)} K(\boldsymbol{\theta})$$

- Over 100 parameters are included
- Product for each category
  - Poissonian term
    - n_c events observed with $s_c + b_c$ expected
  - Diphoton mass probability density function f_c
    - Built with Signal and Background Model functions

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} P\left(n_c | s_c\left(m_H, \boldsymbol{\theta}\right) + b_c\right) \prod_{i=1}^{n_c} f_c\left(m^i_{\gamma\gamma}, m_H, \boldsymbol{\theta}\right) \boxed{K(\boldsymbol{\theta})}$$

- Over 100 parameters are included
- Product for each category
  - Poissonian term
    - n_c events observed with $s_c + b_c$ expected
  - Diphoton mass probability density function f_c
    - Built with Signal and Background Model functions
- $K(\boldsymbol{\theta}) = \prod_i \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta_i}{2}}$ are a set of gaussian constraints for NPs
  - A systematic uncertainty affecting quantity X, enters with a parameter δ in the likelihood as

$$X(\theta_i) = X \cdot r(\theta_i) = X \cdot \begin{cases} 1 + \delta\theta_i \\ e^{\delta\theta_i} \end{cases}$$

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} P\left(n_c | s_c(m_H, \boldsymbol{\theta}) + b_c\right) \prod_{i=1}^{n_c} f_c\left(m^i_{\gamma\gamma}, m_H, \boldsymbol{\theta}\right) K(\boldsymbol{\theta})$$

- Higgs mass m_H is the POI
  - Value and **uncertainty** are extracted with likelihood maximisation

Likelihood Fit

Statistical strategy is based on the definition of likelihood function

$$\mathcal{L}\left(m_H, \boldsymbol{\theta}; \boldsymbol{n}_c, \boldsymbol{m}_{\gamma\gamma}\right) = \prod_{c=1}^{N_{\text{cat}}} P\left(n_c | s_c(m_H, \boldsymbol{\theta}) + b_c\right) \prod_{i=1}^{n_c} f_c\left(m^i_{\gamma\gamma}, m_H, \boldsymbol{\theta}\right) K(\boldsymbol{\theta})$$

- Higgs mass m_H is the POI
  - Value and **uncertainty** are extracted with likelihood maximisation
- Test Statistic is the **PLR**

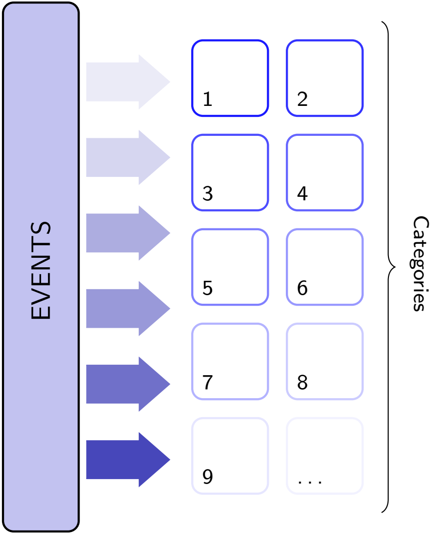
$$\lambda(m_H) = -2\log\left(\frac{\mathcal{L}(m_H, \hat{\boldsymbol{\theta}}_{m_H})}{\mathcal{L}(\hat{m}_H, \hat{\boldsymbol{\theta}})}\right)$$

  - May be interpreted as change of χ^2 wrt minimum
- Performed on an Asimov dataset

Event Categorisation

Event Categorisation

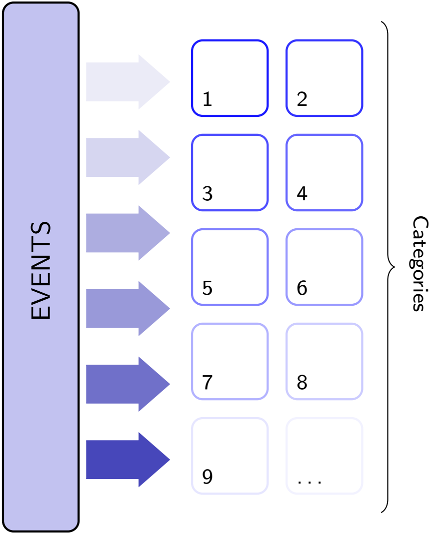
Events are classified according to kinematic requirements (p_{T}, η, ...)



Event Categorisation

Events are classified according to kinematic requirements (p_T, η, ...)

- Categories are orthogonal
- Enhance the sensitivity of the analysis

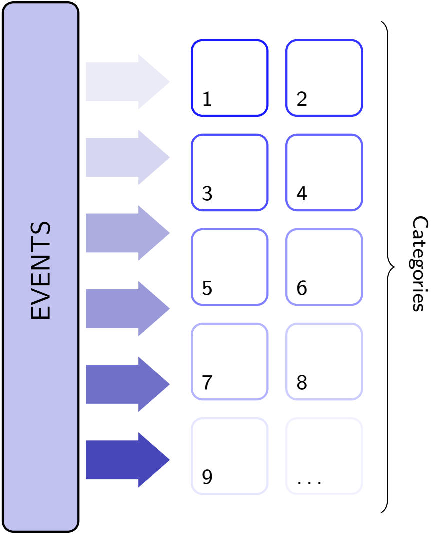


Event Categorisation

Events are classified according to kinematic requirements (p_T, η, ...)

- Categories are orthogonal
- Enhance the sensitivity of the analysis

Systematic uncertainties must be reduced as seen in slide 2



Event Categorisation

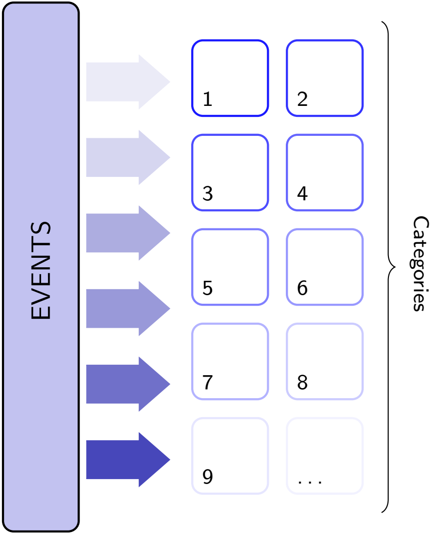
Events are classified according to kinematic requirements (p_T, η, ...)

- Categories are orthogonal
- Enhance the sensitivity of the analysis

Systematic uncertainties must be reduced as seen in slide 2

By means

- Performance studied of the EM calorimeter



Event Categorisation

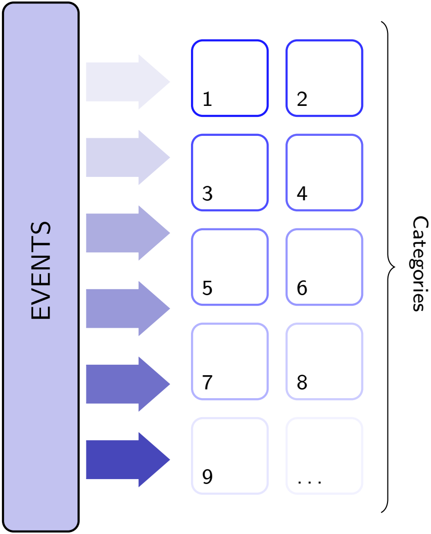
Events are classified according to kinematic requirements (p_T, η, ...)

- Categories are orthogonal
- Enhance the sensitivity of the analysis

Systematic uncertainties must be reduced as seen in slide 2

By means

- Performance studied of the EM calorimeter
- Build a **systematic uncertainty oriented** categorisation
  - Select events with low systematic value
- Suitable choice can reduce the systematic uncertainty
- 4 different categorisation are studied



Coupling analysis

Coupling analysis

First categorisation tested is the one used for the 2018 Higgs Coupling analysis

- 29 categories built to target different Higgs boson production modes

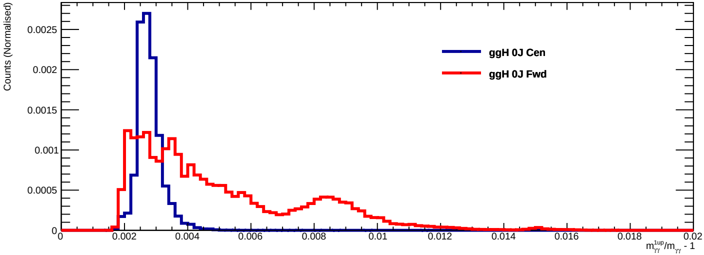
Coupling analysis

First categorisation tested is the one used for the 2018 Higgs Coupling analysis

- 29 categories built to target different Higgs boson production modes

Attempt to optimise the uncertainty on coupling categorisation

- Analyse 2 *ggH* categories
- $m_{\gamma\gamma}^{\text{var}}/m_{\gamma\gamma} - 1$ distribution, for total systematic variation
  - ggH 0J Cen is sharply peaked
  - ggH 0J Fwd is broader



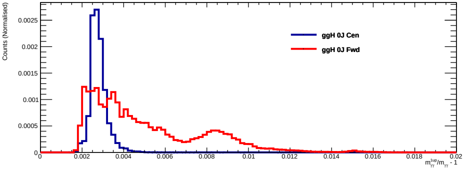
Coupling analysis

First categorisation tested is the one used for the 2018 Higgs Coupling analysis

- 29 categories built to target different Higgs boson production modes

Attempt to optimise the uncertainty on coupling categorisation

- Analyse 2 *ggH* categories
- $m_{\gamma\gamma}^{\text{var}}/m_{\gamma\gamma} - 1$ distribution, for total systematic variation
  - ggH 0J Cen is sharply peaked
  - ggH 0J Fwd is broader
- Address a new categorisation according to systematic value



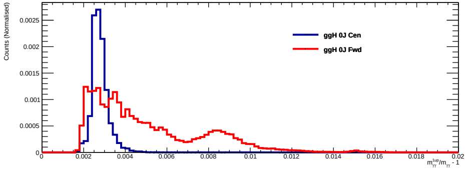
Coupling analysis

First categorisation tested is the one used for the 2018 Higgs Coupling analysis

- 29 categories built to target different Higgs boson production modes

Attempt to optimise the uncertainty on coupling categorisation

- Analyse 2 *ggH* categories
- $m_{\gamma\gamma}^{\text{var}}/m_{\gamma\gamma} - 1$ distribution, for total systematic variation
  - ggH 0J Cen is sharply peaked
  - ggH 0J Fwd is broader
- Address a new categorisation according to systematic value
- Higher systematic values are seen in
  - Unconverted photons
  - Photons at $|\eta| \sim 1.5$



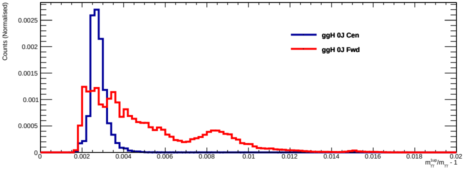
Coupling analysis

First categorisation tested is the one used for the 2018 Higgs Coupling analysis

- 29 categories built to target different Higgs boson production modes

Attempt to optimise the uncertainty on coupling categorisation

- Analyse 2 *ggH* categories
- $m_{\gamma\gamma}^{\text{var}}/m_{\gamma\gamma} - 1$ distribution, for total systematic variation
  - ggH 0J Cen is sharply peaked
  - ggH 0J Fwd is broader
- Address a new categorisation according to systematic value
- Higher systematic values are seen in
  - Unconverted photons
  - Photons at $|\eta| \sim 1.5$
- The division follows those criteria
- 4 new categories are introduced



Production mode categorisation

Production mode categorisation

A different approach in categorising events

- Aiming to reduce systematic uncertainty as well
- Based on Higgs production mode

Production mode categorisation

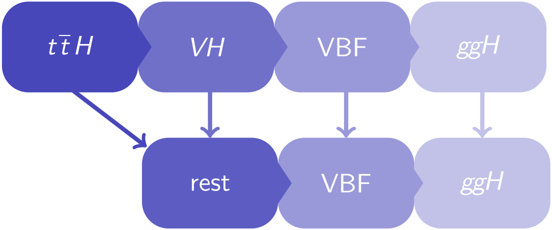
A different approach in categorising events

- Aiming to reduce systematic uncertainty as well
- Based on Higgs production mode

Production mode categorisation

A different approach in categorising events

- Aiming to reduce systematic uncertainty as well
- Based on Higgs production mode



- ggH being the most populated

Production mode and systematic value categorisation

Production mode and systematic value categorisation

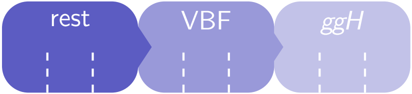
Natural subdivision of production mode categorisation

- Focus on total systematic uncertainty
- Create 3 subcategories

Production mode and systematic value categorisation

Natural subdivision of production mode categorisation

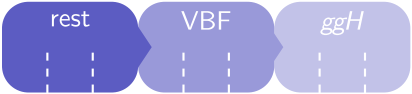
- Focus on total systematic uncertainty
- Create 3 subcategories



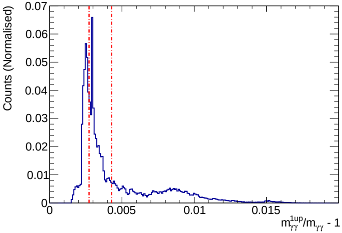
Production mode and systematic value categorisation

Natural subdivision of production mode categorisation

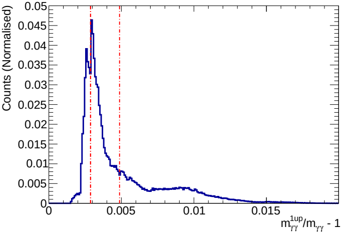
- Focus on total systematic uncertainty
- Create 3 subcategories



Total systematic value distribution



ggH

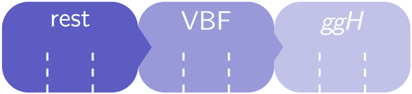


VBF

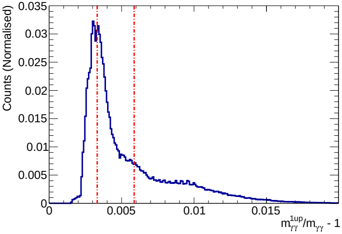
Production mode and systematic value categorisation

Natural subdivision of production mode categorisation

- Focus on total systematic uncertainty
- Create 3 subcategories



Total systematic value distribution



rest

Results

For every categorisation

- Models are re–built (Signal, Background, Systematics)
- Asimov fit is performed
  - $\Rightarrow$ extract the expected uncertainty on m_H

Results

For every categorisation

- Models are re–built (Signal, Background, Systematics)
- Asimov fit is performed
  - ⇒ extract the expected uncertainty on m_H

Categorisation	Uncertainty [MeV]		
	Total	Statistic	Systematic
Coupling	298	129	268
Split “ggH 0J FWD”	294	126	266
Production mode	325	155	286
Production mode and systematic	280	152	235

Results

For every categorisation

- Models are re–built (Signal, Background, Systematics)
- Asimov fit is performed
  - $\Rightarrow$ extract the expected uncertainty on m_H

Categorisation	Uncertainty [MeV]		
	Total	Statistic	Systematic
Coupling	298	129	268
Split "ggH 0J FWD"	294	126	266
Production mode	325	155	286
Production mode and systematic	280	152	235

Reduction of the total (systematic) uncertainty by 6 % (12 %)

Conclusion

I performed an optimisation of the Higgs boson mass analysis

The last ATLAS measurement showed a significant contribution from systematic uncertainty

I addressed the optimisation of the event category

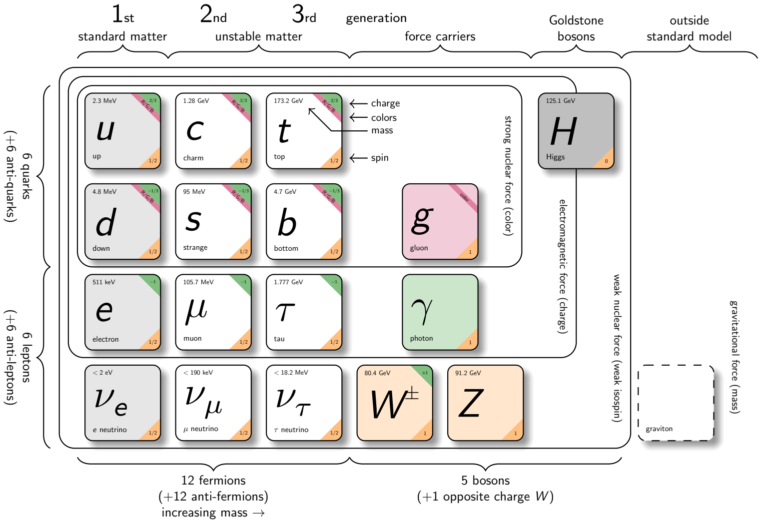
- Event division according to systematic value
- I noticed a possible improvement wrt the actual categorisation
  - Reduction of the systematic uncertainty of 12 %
  - Reduction of the total uncertainty of 6 %

Future improvements will include

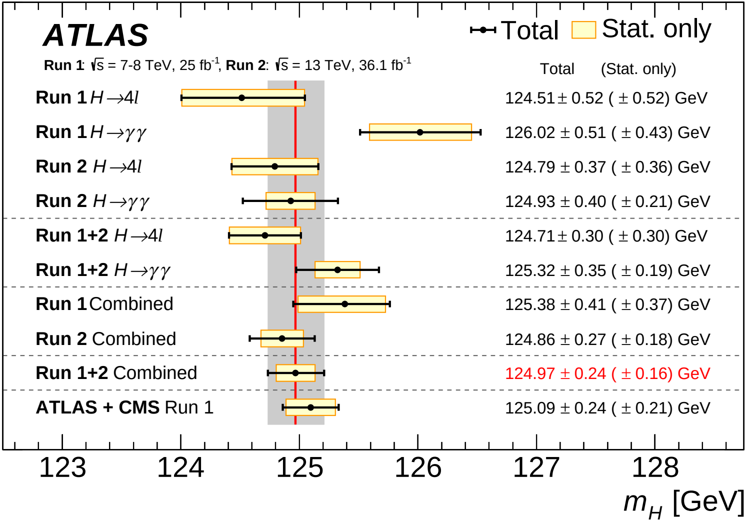
- Reduction of statistical uncertainty
  - Increasing number of categories
- Optimisation of boundaries between them

Backup

The Standard Model



Higgs boson mass



ggH 0J FWD division

From analysis of the systematic value as a function of kinematic variables

- Converted photons have lower systematic
- Higher systematic within the barrel–endcap EM calorimeter transition

Division proposed

ggH 0J FWD CC | ggH 0J FWD | ggH 0J FWD UC

ggH 0J FWD UU GR | ggH 0J FWD UU BR

ggH 0J FWD CC
- Both photons converted

ggH 0J FWD UC
- Only one photon converted

ggH 0J FWD UU GR
- Both photons unconverted
- Both photons away of the transition region

ggH 0J FWD UU BR
- Both photons unconverted
- At least on photons in the transition region

ggH 0J FWD division

From analysis of the systematic value as a function of kinematic variables

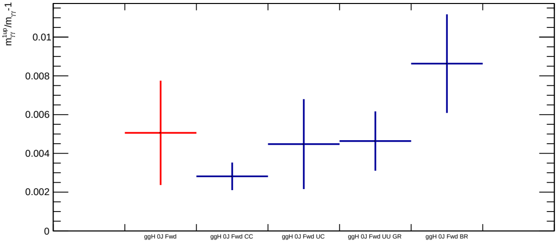
- Converted photons have lower systematic
- Higher systematic within the barrel–endcap EM calorimeter transition

Division proposed

ggH 0J FWD CC | ggH 0J FWD | ggH 0J FWD UC

ggH 0J FWD UU GR | ggH 0J FWD UU BR

Total systematic, mean and RMS



EM calibration