Central electrons resolution constant term bias study for the ATLAS electromagnetic calorimeter

Antinéa Guerguichon
LAL, Orsay
Supervisors: Louis Fayard, Christophe Goudet

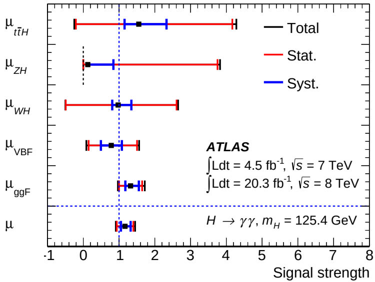
Master's internship defense
June 14, 2016

OUTLINE

Higgs boson: signal strength in $H \to \gamma\gamma$

Signal strength (i.e. N_{obs}/N_{SM}):

μ=1.17 $\pm$ 0.23 (stat) $^{+0.10}_{-0.08}$ (syst) $^{+0.12}_{-0.08}$ (theory)



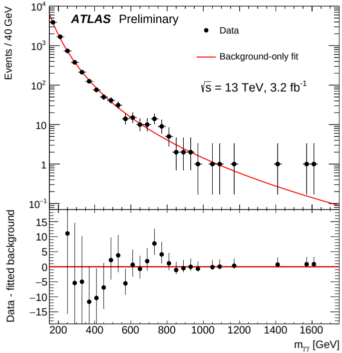
$\rightarrow$ 2018:

- number of events x9
  $\Rightarrow$ (stat)/3
- gluon-gluon fusion Higgs x-section newly computed at N3LO
  $\Rightarrow$ (theory)/2

$\Rightarrow$ systematic uncertainties will become dominant, **energy resolution** being the dominant systematics

X(750 GeV) width

X(750 GeV): excess seen in the $m_{\gamma\gamma}$ spectrum at 750 GeV (to be confirmed)



2 approximations:

- Free Width (**FWA**) $\rightarrow \simeq 45$ GeV
- Narrow Width (**NWA**): 4 MeV

2 different calibrations:

1. **December**, beginning Run II: using 2012 data

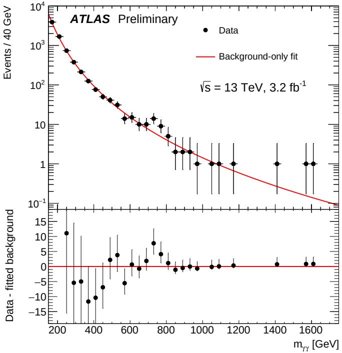
   (ATLAS-CONF-2015-081)

2. **March**: using 2015 data $\rightarrow$ reduce energy resolution uncertainty

   (CERN-EP-2016-120)

X(750 GeV) width

Significance $\sigma = \sqrt{2\Delta \ln L}$ (L = likelihood)

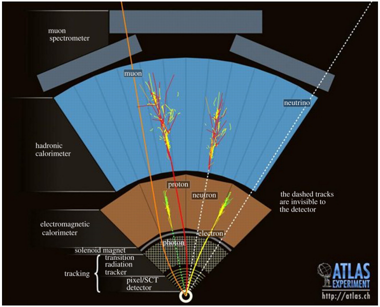


	FWA	NWA	FWA $\ominus$ NWA
calib 1 (December)	3.9σ	3.6σ	**1.5σ**
calib 2 (March)	3.9σ	2.9σ	**2.6σ**

$\Rightarrow$ after the new calibration systematic, the signal is less compatible with the NWA

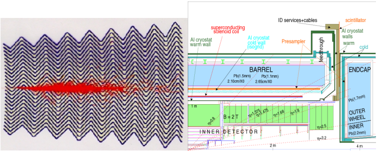
ATLAS EXPERIMENT

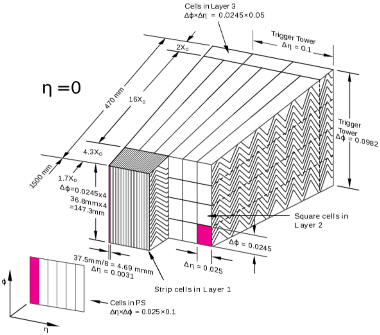
A Toroidal **L**HC **A**pparatu**S**: Multipurpose detector optimized for Higgs and BSM searches



- **Inner Tracker**: track reconstruction, momentum/vertex measurement
- **Electromagnetic CALorimeter (ECAL)**: energy/position of e, γ
- **Hadronic CALorimeter**: energy/position of jets
- **Muon Spectrometer**: momentum/trajectory of muons

ELECTROMAGNETIC CALORIMETER (ECAL)

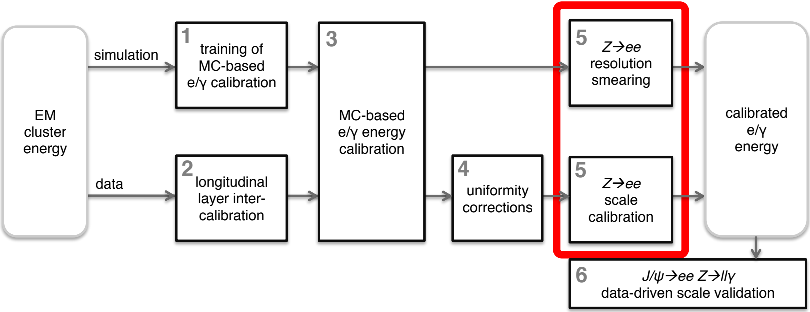




- Sampling calorimeter:
  - Absorber: lead
  - Active medium: liquid argon (LAr)
- Divided in 3 η ranges:
  - **Barrel**: central part $(0 < |\eta| < 1.475)$
  - **Crack**: lot of inactive material in front $(1.37 < |\eta| < 1.52)$
  - **Endcap**: $1.375 < |\eta| < 3.2$
- Transverse segmentation provides a good γ/π^0 separation

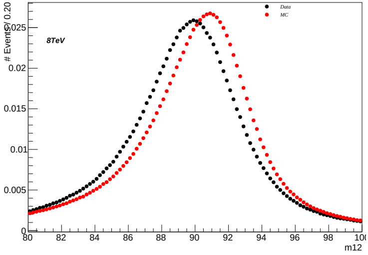
CALIBRATION PROCEDURE

Goal: correct the measured energy to get the true energy of the particle



Scale factors

After step 3 of calibration, MC and data $Z \to ee$ mass distributions **still have a discrepancy**.



Data-driven analysis (step 5)
$\to$ match the data with the MC distribution, using 2 η-dependant corrections: scale factors α (shifts data) and c (enlarges MC)
$\to$ measured with the template method

Energy scale factor α is applied on data:

$$E^{corr} = E^{data} = E^{true}(1 + \alpha)$$

RESOLUTION CONSTANT TERM

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

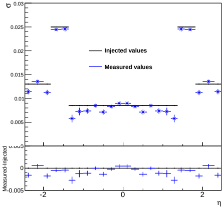
- **a**: **sampling/stochastic term**, linked to the development of the EM shower in the ECAL
- **b**: **electronic noise and pile-up term**
- **c: constant term**, describes non-uniformities in the detector and electronics

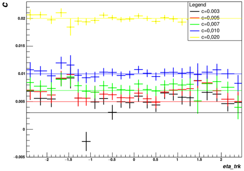
Data distribution larger than the MC: additional constant term c used to enlarge the MC width up to the data one with:

$$E^{corr} = E^{true}(1 + \mathcal{N}(0,1) * c)$$

with $\mathcal{N}(0,1)$ a Gaussian distributed random number

2012 DATA



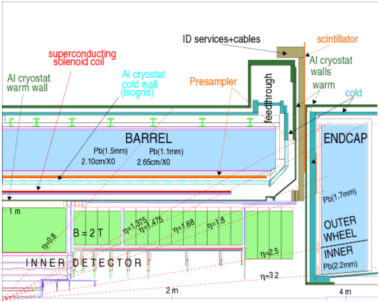


- **Closure**: c^{input} injected into a MC dataset used as pseudo-data + measured with the template method
- Closure systematic defined in each η bin as $\delta_{closure} = |c^{meas} - c^{input}|$
- Bias corresponds to $\delta_{closure}$ averaged over the number of closures: 1 closure before, now many (~ 1000)

Goal: quantize and correct the bias arising from the template method to reduce the closure uncertainty ($\simeq 0.1\%$ in Run I)

DETECTOR SPLITTING IN η BINS

Electrons are labelled according to their η bin $\Rightarrow$ Z labelled by the combination of electrons bins $(\eta_1, \eta_2) = (i, j)$



- First, Z **mass corrections** α_{ij} and c_{ij} are measured independently for each (i, j) configuration
- Electrons scale factors α_i and c_i are inferred afterwards

6 bins used in this study: 0-5 encap, 1-4 crack, 2-3 barrel

*(**NB**: For the final calibration study: 68 bins for α , 24 bins for c)*

Fit at constant c

Modified MC datasets (=**templates**) are created with injected test values of c and α

$\Rightarrow$ χ^2 **between Z mass distribution of pseudo-data and templates is computed**

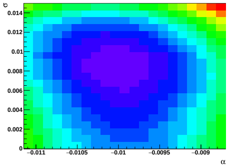
$\Rightarrow$ α_{ij} and c_{ij} most probable values correspond to the fitted minimum of the χ^2 scan $\rightarrow$ fit performed in 2 steps of 1D fits:

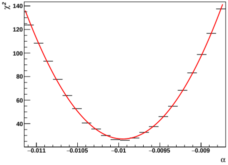
- **For a given** c **(line)**, the $\chi^2(\alpha)$ distribution is fitted with:

$$\chi^2(\alpha) = \chi^2_{min} + \frac{(\alpha - \alpha_{min})^2}{(\Delta\alpha_{min})^2}$$

$\rightarrow$ α_{min} and χ^2_{min} extracted

In this study, α not fitted and set to 0 because α^{input}=0.



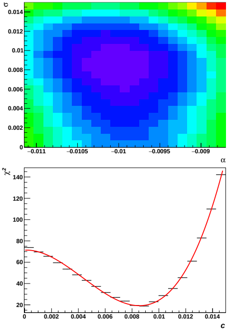


FIT AS A FUNCTION OF C

Then, the $\chi^2_{min}(c)$ distribution is not parabolic (i.e. non Gaussian-like) and fitted with:

$$\chi^2_{min}(c) = \chi^2_{min,min} + \frac{(c - c_{ij})^2}{(\Delta c_{ij})^2} + a_3 \frac{(c - c_{ij})^3}{(\Delta c_{ij})^3}$$

- c_{ij} is the measured constant term in the (η_1, η_2) configuration: given by the minimum of the fit
- Δc_{ij} is its statistical uncertainty (in the Gaussian approximation)



INVERSION PROCEDURE

Inversion procedure = getting the c_i from the c_{ij}

= getting e^- scale factors from the Z ones

→ This requires the minimization of the following χ^2,
assuming α_{ij} and c_{ij} are Gaussian distributed:

$$\alpha_i + \alpha_j = 2\alpha_{ij} \quad \Rightarrow \quad \chi^2 = \sum_{i,j\leq i} \frac{(\alpha_i + \alpha_j - 2\alpha_{ij})^2}{(\Delta\alpha_{ij})^2}$$

$$c_i^2 + c_j^2 = 2c_{ij}^2 \quad \Rightarrow \quad \chi^2 = \sum_{i,j\leq i} \frac{\left(\sqrt{\frac{c_i^2+c_j^2}{2}} - c_{ij}\right)^2}{\Delta^2 c_{ij}}$$

Constant term bias measure

Using a closure, the official MC dataset containing 5.4M events is split in 2:

- 2.7M events: **pseudo-data** $\rightarrow$ smeared with a chosen c^{input} that is measured with the template method
- 2.7M events: **MC templates**

Many closures are used and the bias is defined as:

$$\text{bias} = < c^{meas} - c^{input} >$$

SMEARING

Different sources of statistical fluctuations are needed to simulate:

- the constant term resolution on pseudo-data
- the multiple resolutions on MC templates that will be used to determine the pseudo-data resolution

Reminder: random smearing is done using $E^{corr} = E^{true}(1 + \mathcal{N}(0,1) * c)$ with $\mathcal{N}(0,1)$ a Gaussian distributed random number. 2 types of smearing are done:

- Pseudo-data are smeared with c^{input}
- MC templates are similarly smeared with different test values of c

Bias due to a given set of events

- **To remove the bias coming from the selection of a given set of events**, each closure uses a different set of events selected with the **bootstrap** method:
  - Each event selected with a Poissonian probability
  - 1 event can be chosen several times
- **3 different samples** are generated to look at the influence of statistics:
  - 100k, 1M with bootstrap
  - 2.7M without

⇒ 3 different randomizations are done: 2 on pseudo-data, 1 on MC template

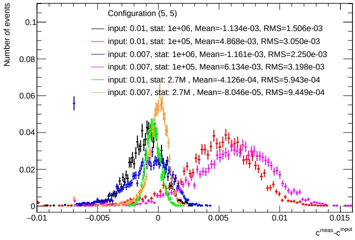
Bias due to the 2.7M statistics

If the number of pseudo-data events = 2.7M and no bootstrap is performed, the bias found corresponds to:

- The limited statistics of 2.7M in MC
- The fact that different events are used as pseudo-data and MC templates

Bias(100k, 1M) estimated as bias(100k, 1M) - bias (2.7M)

Bias of (i, j) configurations (1)

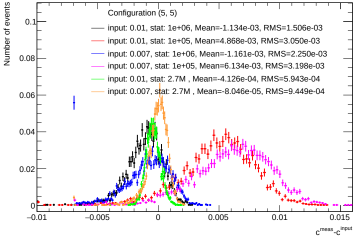


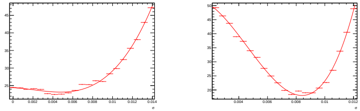
Endcap-encap configuration

- 2 values of c^{input}: 0.7% and 1%
- Mean and RMS obtained from the histogram
- Left points bias = $-c^{input}$ correspond to $c^{meas} = 0$ $\rightarrow$ lots of them because of the $c_{ij} > 0$ constraint

BIAS OF (i, j) CONFIGURATIONS (2)

Endcap - encap configuration

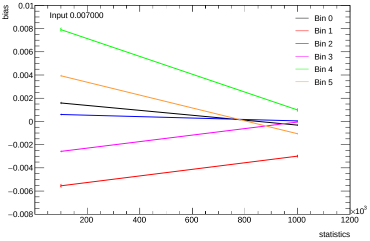




Barrel - barrel configuration for $c^{input} = 0.7\%$ for 100k events (left) and 1M (right)

- Mean and RMS ↘ when stat ↗
  → more parabolic shape around the minimum of the χ^2 distribution for larger stat i.e. easier to fit (still some features to be understood)
- For a higher input, minimum is farther away from 0
  → χ^2 distribution more parabolic and bias should have a better behaviour

Bias of η bins



- c_i (electrons) obtained from c_{ij} (Z) after the inversion procedure
- Bias from the 2.7M events is substracted
- Worst results are in the crack regions (bins 1, 4)

At large statistics, $|\textbf{Bias}| \leq \textbf{0.1\%}$ (similar to Run I and II)

CONCLUSION

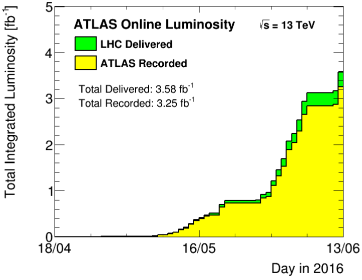
- Understanding the bias arising from the template method is essential to assess the uncertainty on c which is important in the H(125 GeV) and X(750 GeV) (?) studies
- Bias has a value of about 0.1%
- Study still ongoing!

CONCLUSION

- Understanding the bias arising from the template method is essential to assess the uncertainty on c which is important in the H(125 GeV) and X(750 GeV) (?) studies
- Bias has a value of about 0.1%
- Study still ongoing!

Prospects:

- More detailed study of the influence of the differents steps of the template method
- Switch from 6 to 24 bins
- Check the bias at high statistics (6M $Z \to ee$ events in 2012, more in 2016)

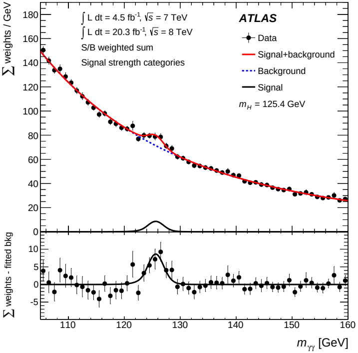


BACK-UP

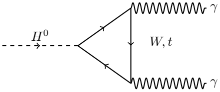
Higgs boson: mass

Combined ATLAS+CMS value:

$m_H = 125.09 \pm 0.21$ (stat) $\pm\ 0.11$ (syst) GeV



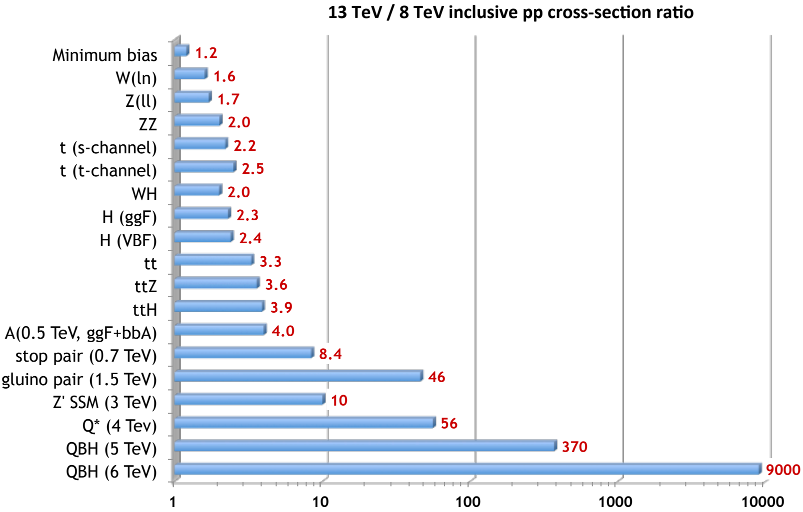
H produced/decays through a loop → sensitive to BSM



- 2018: number of events x9 ⇒ (stat)/3

⇒ systematic uncertainties will become dominant, calibration (energy scale) being the dominant systematics in $H \to \gamma\gamma$

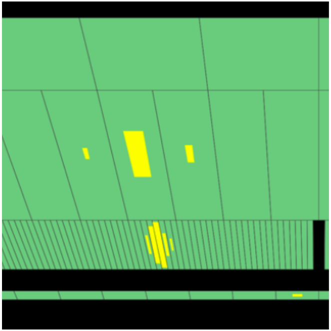
CROSS-SECTION RATIO 13 TEV / 8 TEV



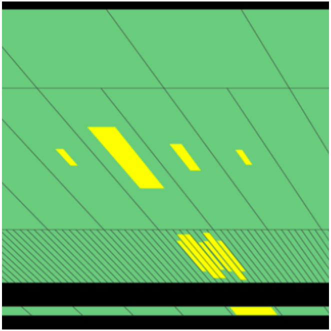
Systematic uncertainties on the measured constant term (2012)

	Exp. sources of uncertainty ($\times 10^{-4}$)									
η bin	PileUp	Reco eff.	Trig. eff.	ID	Clos.	Window	Fbrem	Meth.	EW	QCD
[0;0.2]	15.8	0.71	2.85	8.40	4.80	6.86	1.21	2.74	5.11	10.9
[0.2;0.4]	7.17	1.85	0.066	8.09	2.16	3.84	14.8	10.4	0.76	10.9
[0.4;0.6]	2.99	0.03	1.64	17.5	13.6	10.8	11.4	3.67	2.89	10.9
[0.6;0.8]	1.14	0.42	0.42	11.6	0.15	1.46	6.71	8.21	1.21	7.72
[0.8;1]	4.03	0.92	0.063	2.63	11.3	2.28	1.72	3.64	5.42	7.72
[1;1.2]	9.37	0.19	0.44	9.02	12.4	9.91	25.8	9.48	12.2	5.68
[1.2;1.37]	0.43	0.34	0.11	0.22	27.1	4.33	19.2	18.3	1.77	5.68
[1.37;1.55]	17.5	1.37	2.51	28.3	4.30	40.5	75.2	19.6	7.04	41.7
[1.55;1.82]	0.45	0.09	0.52	14.9	5.50	16.7	5.46	16.4	0.35	12.2
[1.82;2]	6.84	11.3	15.8	6.68	17.6	1.08	19.7	42.1	4.43	15.4
[2;2.3]	2.55	3.32	0.53	19.2	5.68	25.2	23.8	0.14	3.49	15.4
[2.3;2.47]	2.87	0.9	0.62	28.2	15.8	10.6	36.2	19.4	0.91	15.4

γ / π^0 SEPARATION



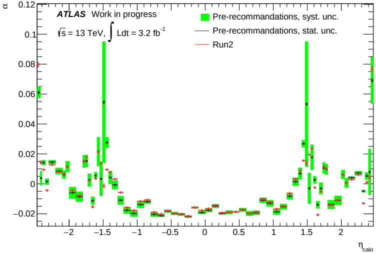
γ display

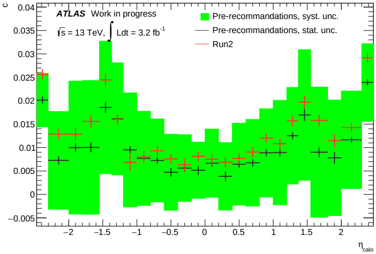


π^0 display

(https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20100721/display-photons/index.html)

NOMINAL SCALE FACTORS





Configurations: bias distribution