

# Variation of the response of Electromagnetic Calorimeter in CMS experiment

## Study performed with electrons

**Presented by:** Gursharan singh

I-Ph.D student, 2nd year

**Supervised by:** Dr. Rajdeep Chatterjee

I would like to thank **Prof. Gobinda Majumder** and  
**Dr. Shilpi Jain** for their help with this project.



# Outline

- CMS
- ECAL
- Response of ECAL
- Intrinsic Factors
- In-situ Factors
- Algorithmic Corrections
- Study of Response with different parameters

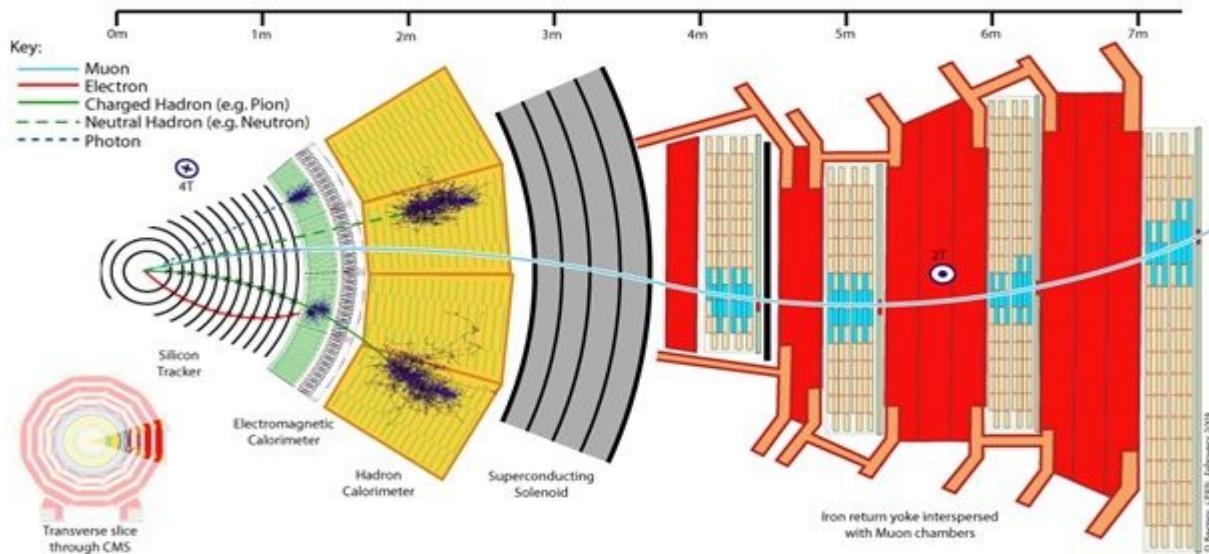
# COMPACT MUON SOLENOID DETECTOR

The [Compact Muon Solenoid](#) (CMS) is a general-purpose detector at the [Large Hadron Collider](#) (LHC) at CERN.



## CMS DETECTOR

Total weight	: 14,000 tonnes
Overall diameter	: 15.0 m
Overall length	: 28.7 m
Magnetic field	: 3.8 T



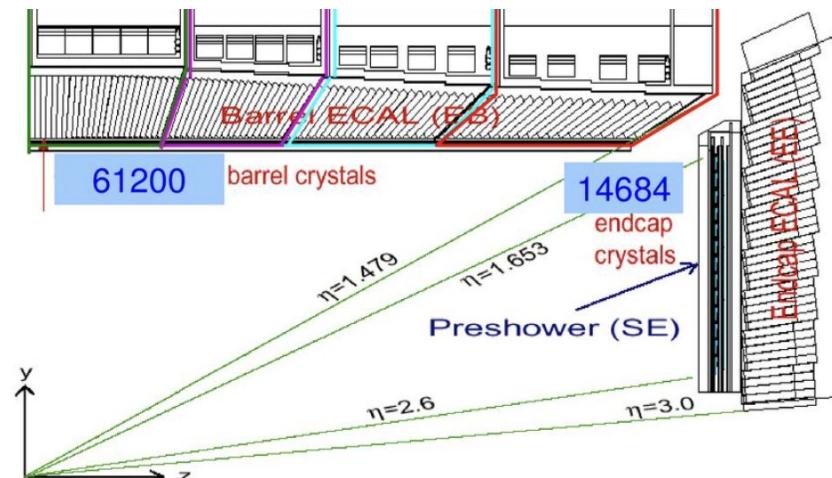
# Electromagnetic calorimeter (ECAL) in CMS

- Electromagnetic Calorimeter (ECAL) is used for measuring energies of electrons and photons
- The barrel portion ( $|\eta| < 1.47$ , called EB) consists of lead tungstate ( $\text{PbWO}_4$ ) crystals, which generate light proportional to the energy deposited by charged particles, enabling precise energy measurements.
- The end cap part ( $1.55 < |\eta| < 3.0$ ) also consists of lead tungstate ( $\text{PbWO}_4$ ) crystals along with the silicon preshower .
- **End Cap Crystals: Length:** 220 mm ( $24.7 X_0$ )
- **Barrel Crystals: Length:** 230 mm ( $25.8 X_0$ )

Energy Response of the calorimeter is defined as  
 $E_{\text{obs.}} / E_{\text{gen.}}$

For a good detector, **Response** should be **Linear and Stable**

I have studied the variation of the response of the ECAL with CMS Run2 simulation



<https://slideplayer.com/slide/17178528/>

# Electromagnetic Shower Development

## Initiation:

- High-energy photons or electrons enter the calorimeter crystal..

## Primary Interactions:

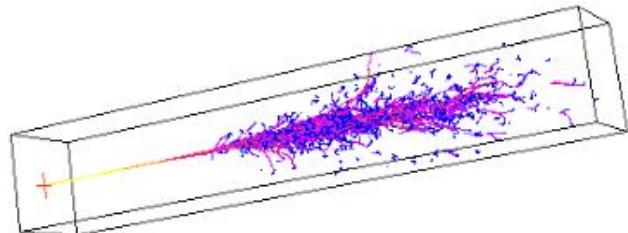
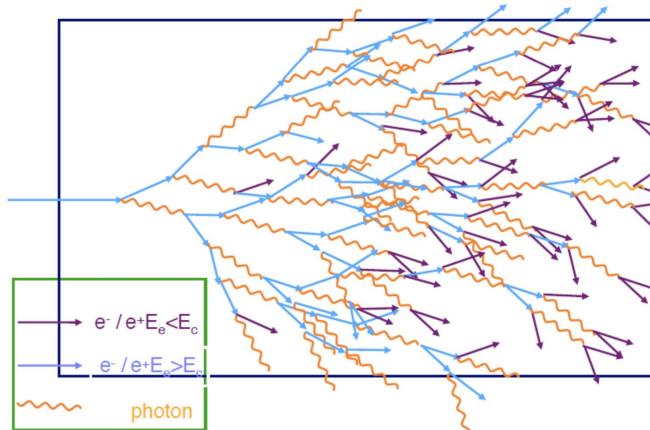
- **Photons** : Undergo pair production, creating an electron-positron pair.
- **Electrons**: Emit bremsstrahlung photons as they travel through the material.

## Secondary Interactions:

- The newly created electrons and positrons from pair production emit additional bremsstrahlung photons.
- These photons undergo further pair production, creating more electron-positron pairs.

## Termination:

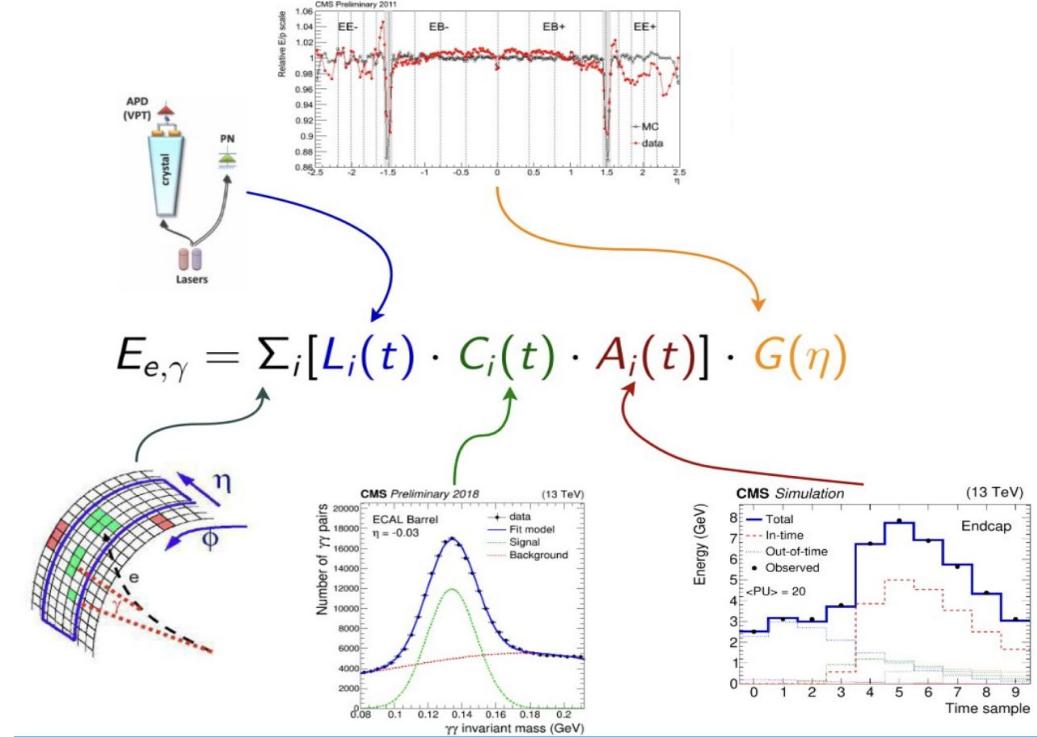
- The shower continues to develops until secondary particles reaches critical energy (EC) below which ionization energy loss is dominant.



L. Finco University of Nebraska – Lincoln

# Electron and Photon Energy Reconstruction

- L=correct for time response variation of each crystal.
- C=inter calibration coefficients
- A=amplitude in ADC counts
- G=ADC-to-GeV conversion factor



# Intrinsic ECAL energy relative response

The ECAL intrinsic energy resolution has been measured as follows:

Energy Response of the calorimeter is defined as  $E_{\text{obs.}}/E_{\text{gen.}}$

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

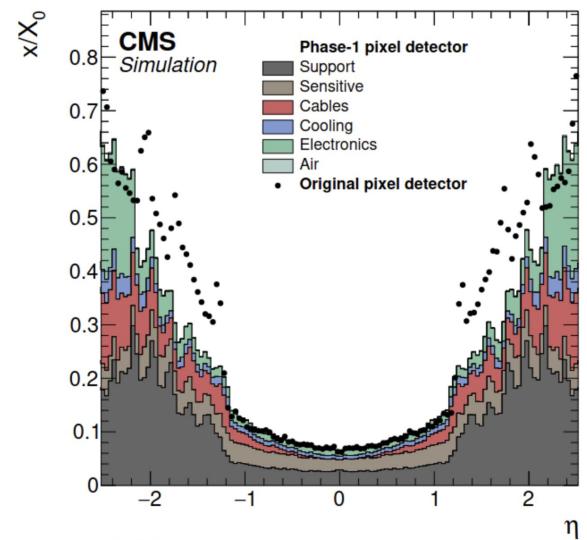
Where a, b and c are constants.

- First term a : stochastic term, intrinsic statistical shower fluctuations  $\Rightarrow$  dependency of  $1/\sqrt{E}$
- Second term b: arising due to Electronic noise,pile-up etc  $\Rightarrow$  dependency of  $1/E$
- Third term c: constant arises due the energy leakage from the back of the calorimeter crystals, energy deposit in the dead areas etc.

# Challenges in Electron and Photon Object Reconstruction

*Tracker material profile*

- Due to the material in the tracker region EM shower develops earlier (before entering the ECAL).
- Pile-up.
- Electronic noise in the crystals readout system.

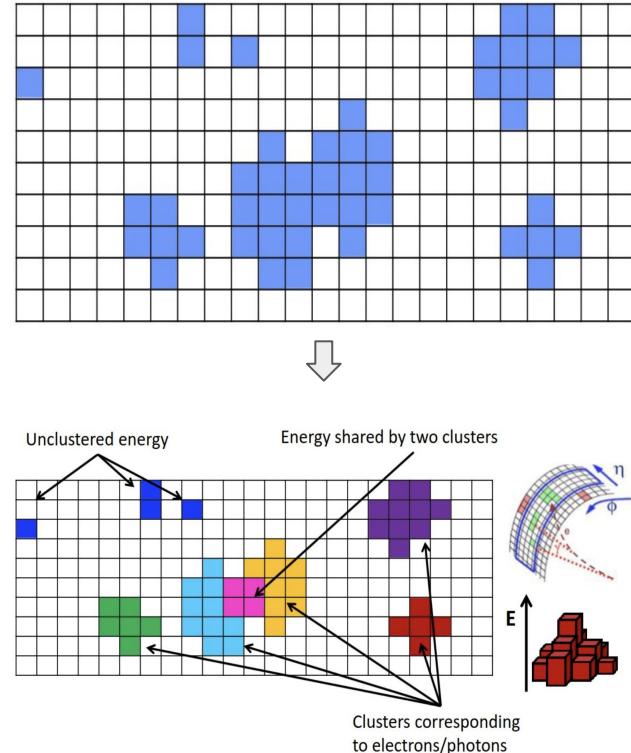


# Reconstruction algorithm

- Due to tracker material → Early showering occur
- Need an algorithm for collecting those hits in the ECAL → inorder to generate the energy of the primary  $e, \gamma$  .

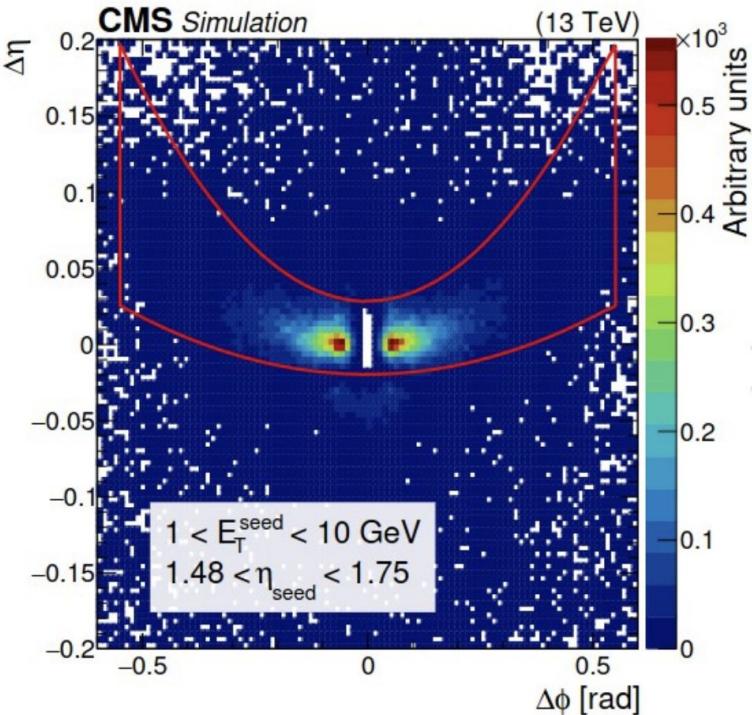
# Clustering Algorithm

- Identifying the crystals having energy deposition that exceeds a predefined threshold (typically  $\sim 80$  MeV for EB and  $\sim 300$  MeV in EE ), this threshold is designed to avoid noise so it's 2 to 3 times higher than the electronic noise expected in the crystals.
- Then, from these crystals, a seed with a minimum energy deposit of 1 GeV is defined. These are called seed crystals, and we start our clustering around them. Here we can see 5 cluster around the seed crystals.



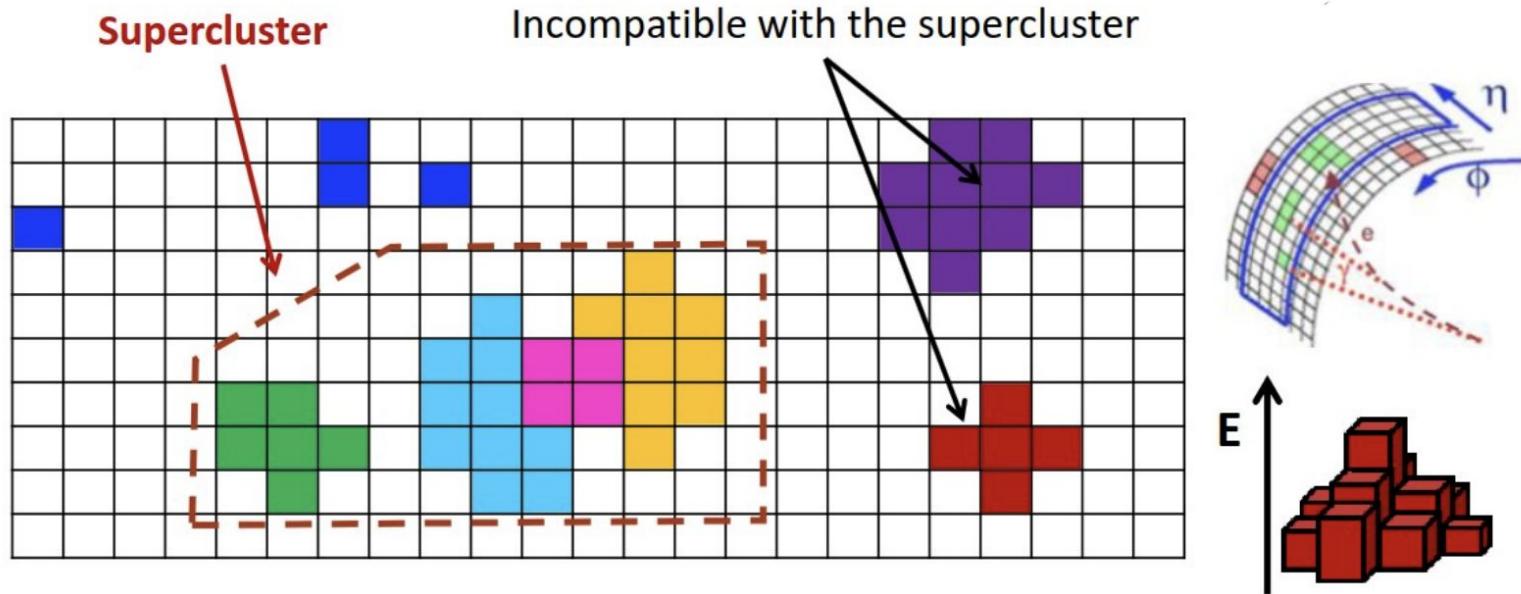
# Mustache algorithm

- It first identifies the cluster above a given threshold called seed cluster and adds other clusters that fall into the area that appears mustache in shape in the transverse plane to form a Supercluster.
- The mustache parabolas are parametrized by the  $\eta$  position of the seed, the energy ( $E$ ) and the transverse energy ( $E_T$ ) of the cluster.
- It is based on a purely geometrical algorithm, it has a high signal efficiency, but also it is highly contaminated by spurious clusters generated by the ECAL electronic noise and pile-up (PU).



arXiv:2012.06888 [hep-ex]

# Super clustering



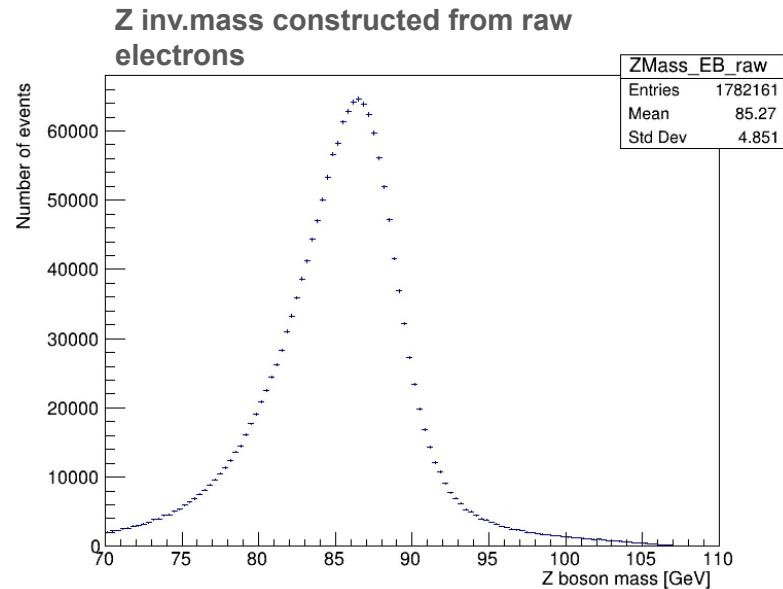
- Supercluster formed via mustache algorithm

## RAW ENERGY

- The energy obtained from the Supercluster is called the Raw energy.

## Still there is Mismatch of Energy

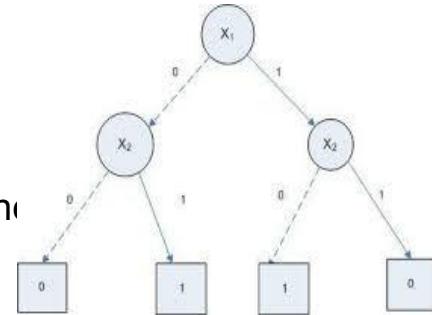
- Energy lost in gaps.
- Large amount of upstream material.
- Pile up .
- Energy lost due to leakage at the back of the crystals.
- Algorithm Inefficiency.



## Correction Procedure

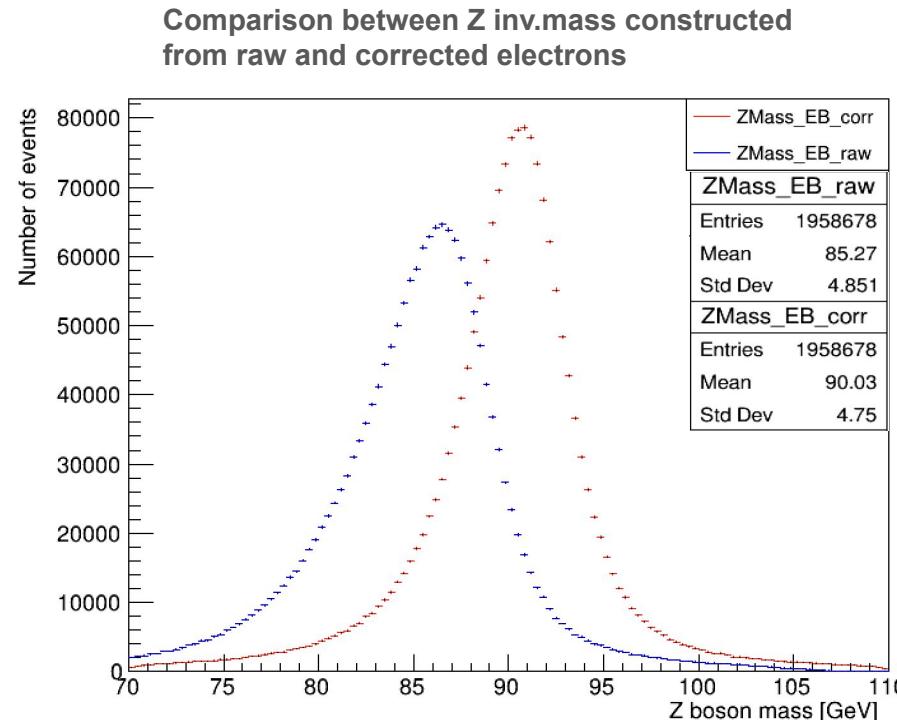
Using a regression model of Boosted Decision Trees (BDT), the raw energy is corrected.

- The electron and photon regressions are trained on samples of simulated events with two electrons or photons in each event, generated with a flat transverse momentum spectrum, where the true value of the e/ $\gamma$  energy is known and the geometric condition  $\Delta R < 0.1$  is used to find a match of the reconstructed e/ $\gamma$  to the true ones.
- The target variable which is the ratio of generator level e/ $\gamma$  energy to the raw SC energy.
- The input variables are the SC position and shower shape variables [which tell about the degree of showering in the material]



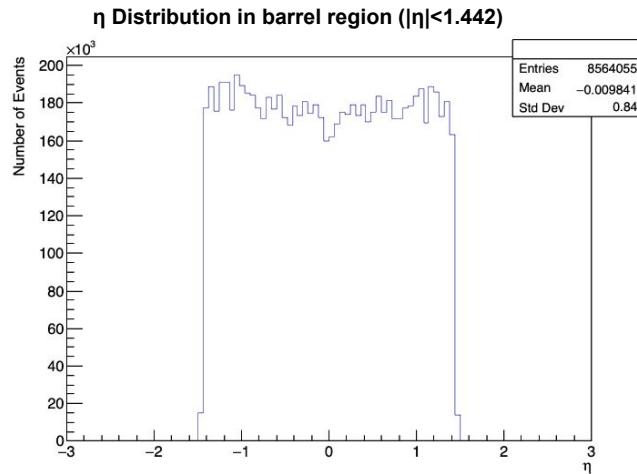
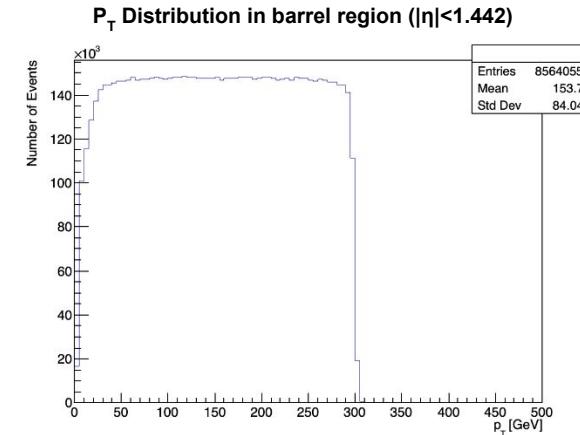
## Corrected Energy

This is the energy obtained after applying the corrections to the raw energy.

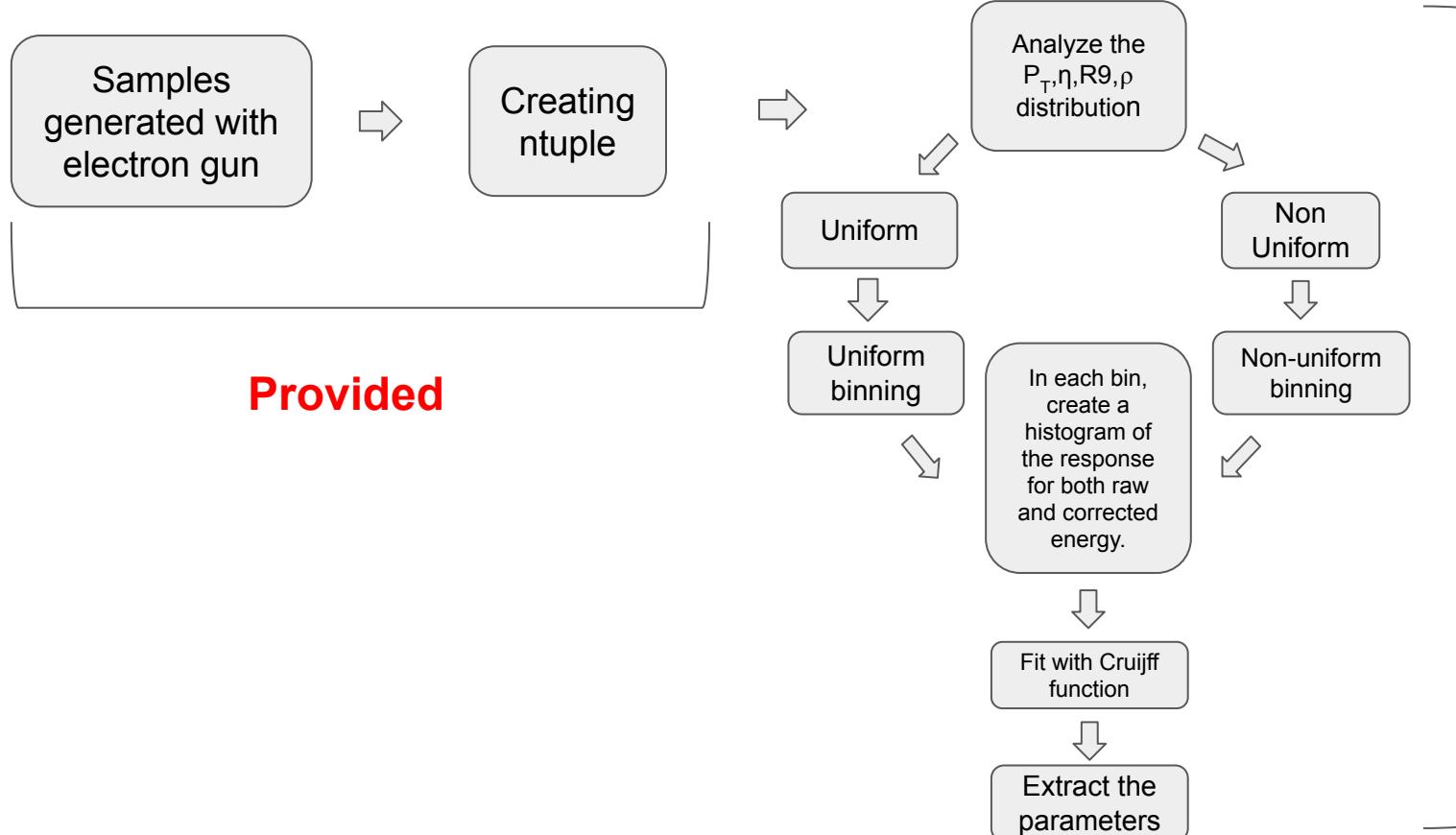


# SAMPLES

- These are the electron samples.
- They are flat in transverse momentum  $P_T$ .
- And uniform in Pseudorapidity  $\eta$ .



# Flow chart of framework



Build by  
me

<https://drive.google.com/drive/folders/1KB0gpQGR118KvqkJRwQn8ogDR9ISKLM?usp=sharing>

# Cruijff function

x: The variable of interest, typically the mass or another observable in the experiment.

m0: The central value of the distribution, often representing the peak position.

$\sigma_L$ : The width (standard deviation) of the left side of the peak (for  $x < m_0$ ).

$\sigma_R$ : The width (standard deviation) of the right side of the peak (for  $x > m_0$ ).

$\alpha_L$ : The parameter that describes the asymmetry on the left side of the distribution. It affects the shape and steepness of the left tail.

$\alpha_R$ : The parameter that describes the asymmetry on the right side of the distribution. It affects the shape and steepness of the right tail.

<https://probfit.readthedocs.io/en/latest/api.html>

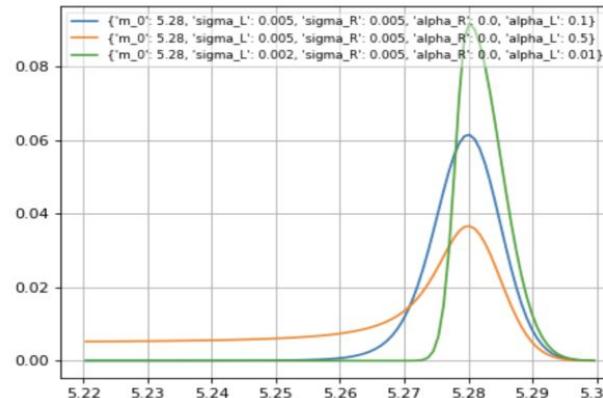
## cruijff

```
probfit.pdf.cruijff(double x, double m_0, double sigma_L, double sigma_R, double alpha_L, double alpha_R) →  
double
```

Unnormalized cruijff function

$$f(x; m_0, \sigma_L, \sigma_R, \alpha_L, \alpha_R) = \begin{cases} \exp\left(-\frac{(x-m_0)^2}{2(\sigma_L^2 + \alpha_L(x-m_0)^2)}\right) & \text{if } x < m_0 \\ \exp\left(-\frac{(x-m_0)^2}{2(\sigma_R^2 + \alpha_R(x-m_0)^2)}\right) & \text{if } x > m_0 \end{cases}$$

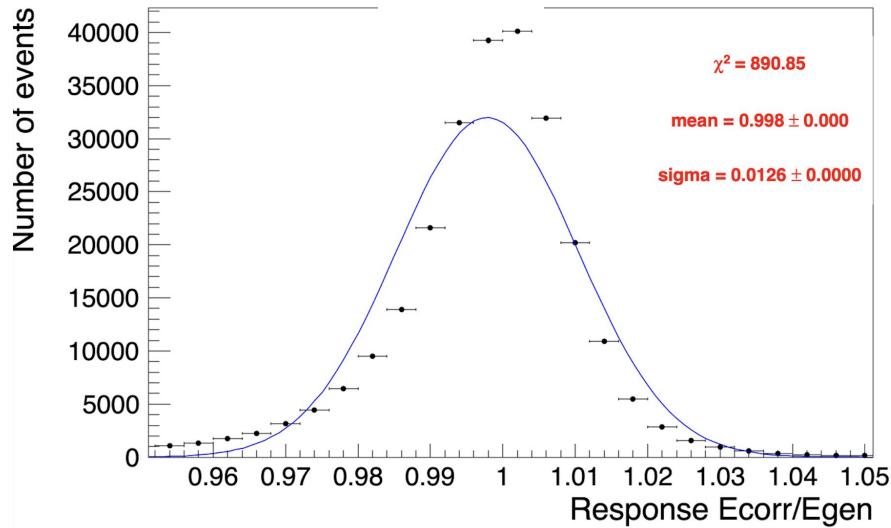
([Source code](#), [png](#), [hires.png](#), [pdf](#))



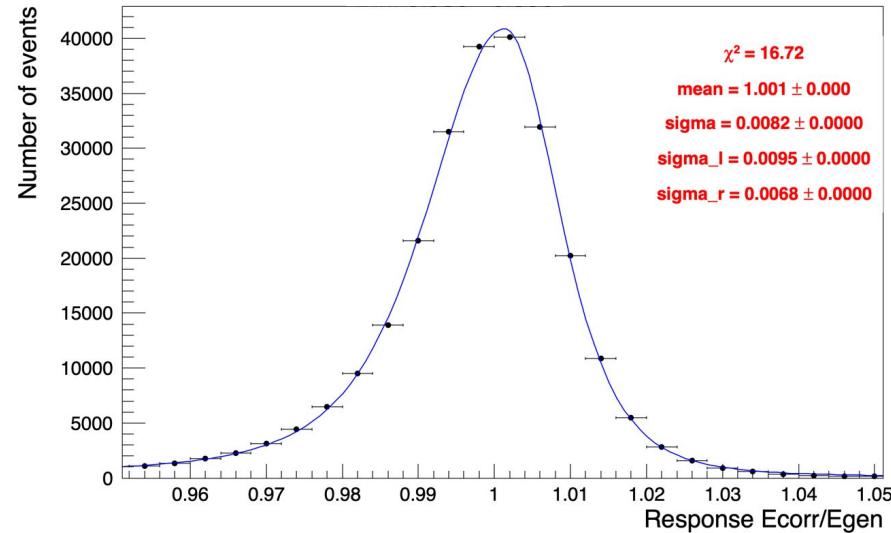
# Why Cruijff Function is used for fitting?

Corrected R9 0.98-0.99

## Fitted with Gaussian function



## Fitted with Cruijff function

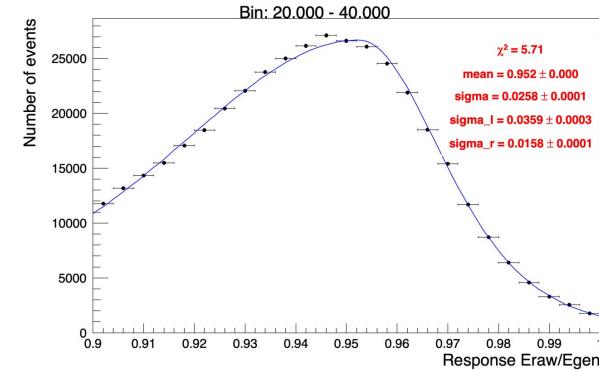
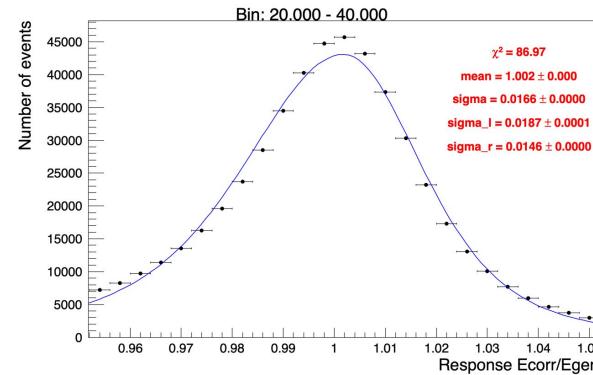


So, We selected Cruijff function to fit our response histograms.

# Variation of relative energy response with Transverse momentum $P_T$

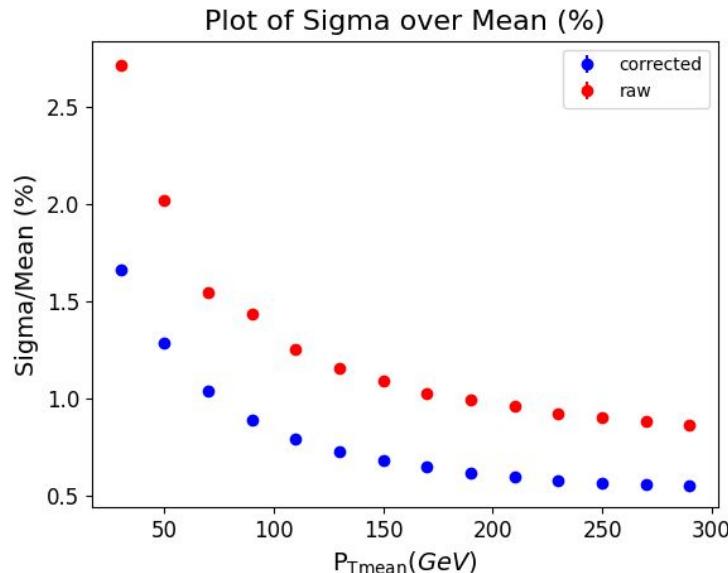
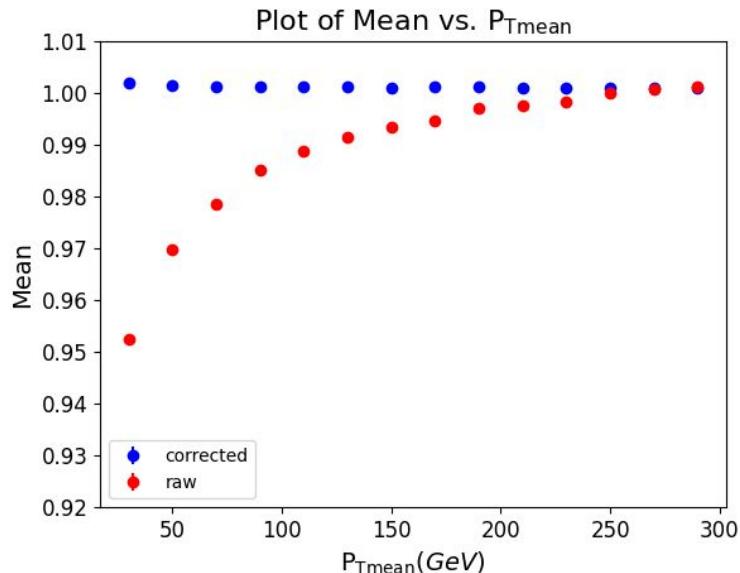
- In Order to derived the relative response, we divide the  $P_T$  distributing into bins  $\rightarrow$  20 GeV to 300 GeV in binning of 20 GeV  $\rightarrow$  equal statistics.
- There in each bin we plotted the response, for both raw and corrected, as an example shown here.
- We fitted it with cruijff function with reference taken as mean and obtained the energy resolution.

Response fitted in 20-40 GeV bin of  $P_T$  distribution



## Relative response and mean of response with Mean $P_T$

- Also low  $P_T$  means, electron will have more bend in tracker → more energy loss.
- Algorithmic inefficiency.
- Stability of corrected response  $P_T$  Mean is 0.0953 %.



# Variation of relative energy response with Pseudorapidity $\eta$

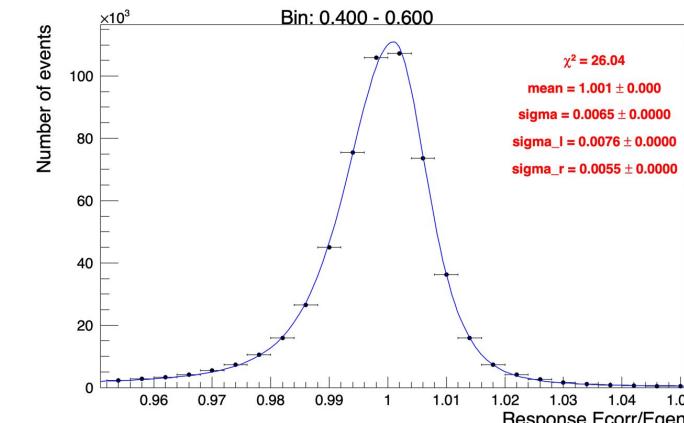
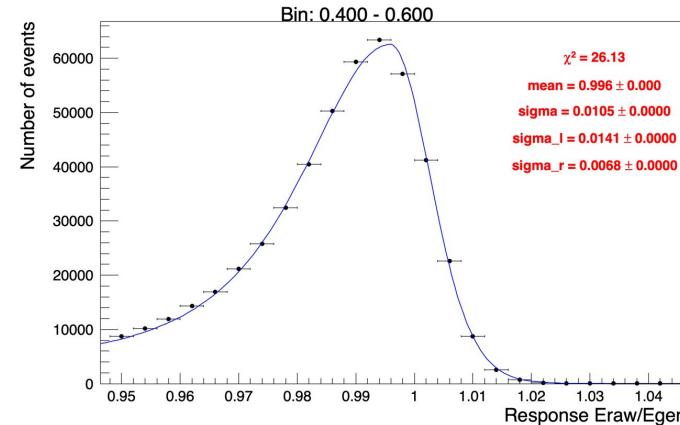
- It is a spatial coordinate used to describe the angle of a particle relative to the beam axis.
- Defined as:

$$\eta = -\ln \left( \tan \left( \frac{\theta}{2} \right) \right)$$

Where  $\theta$  is the polar angle

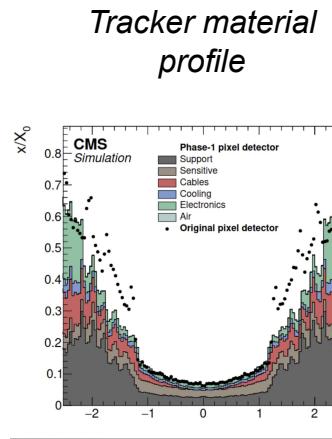
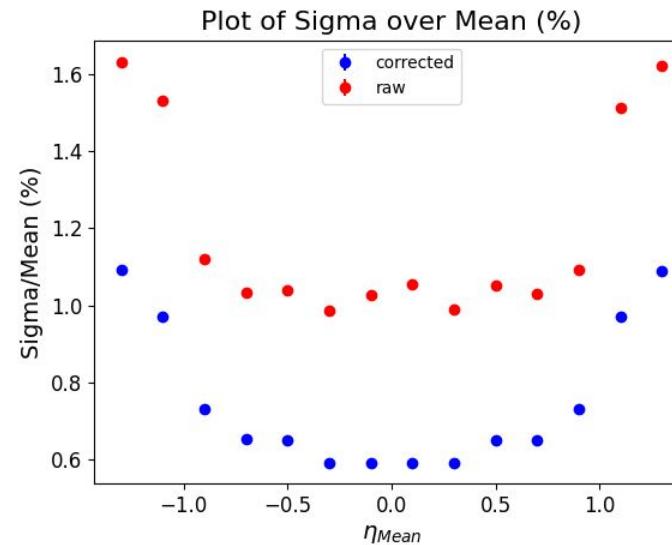
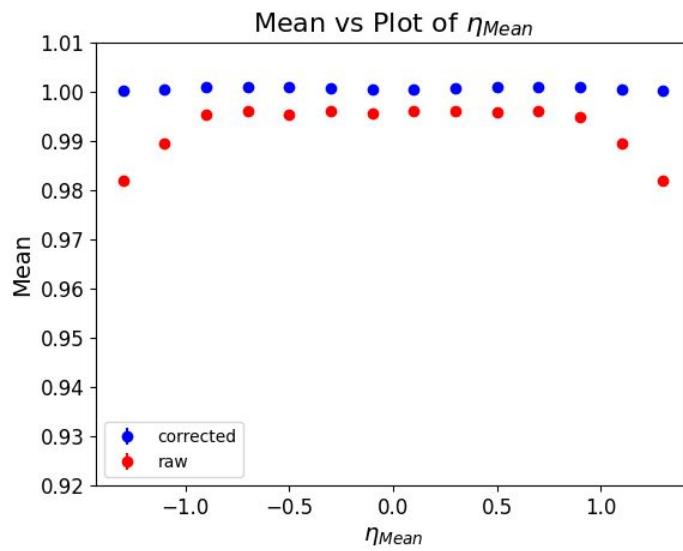
relative to the beam axis.

Response fitted in 0.4-0.6 bin of  $\eta$  distribution



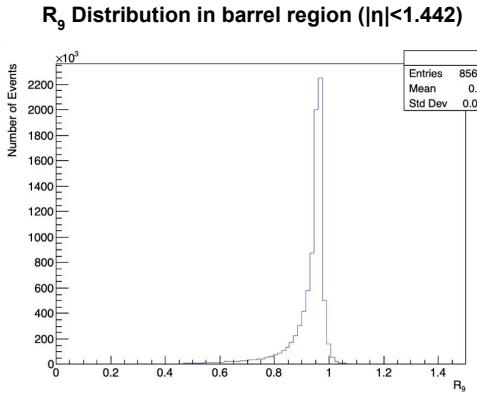
# Relative response and mean of response with Mean $\eta$ bins

- As we move to either side of  $\eta = 0$ , the amount of tracker material symmetrically increases, causing electrons to lose more energy in the tracker medium.
- Stability of corrected response  $\eta$  Mean is 0.0817 %.

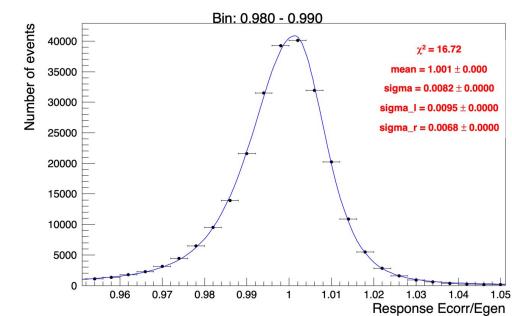
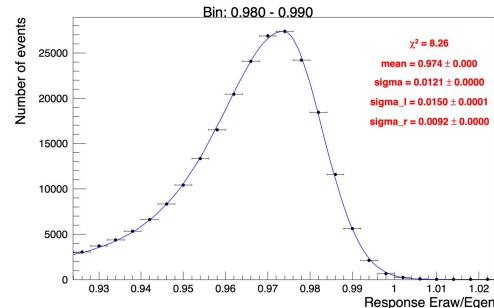


# Variation of relative energy response with $R_9$

- It is defined as  $R_9 = E_{3\times 3}/E_{\text{Supercluster}}$ .
- Its a shower shape variable, higher its value indicate less spreading in the shower and vice versa
- Here , as the distribution is asymmetrical , we did variable binning.

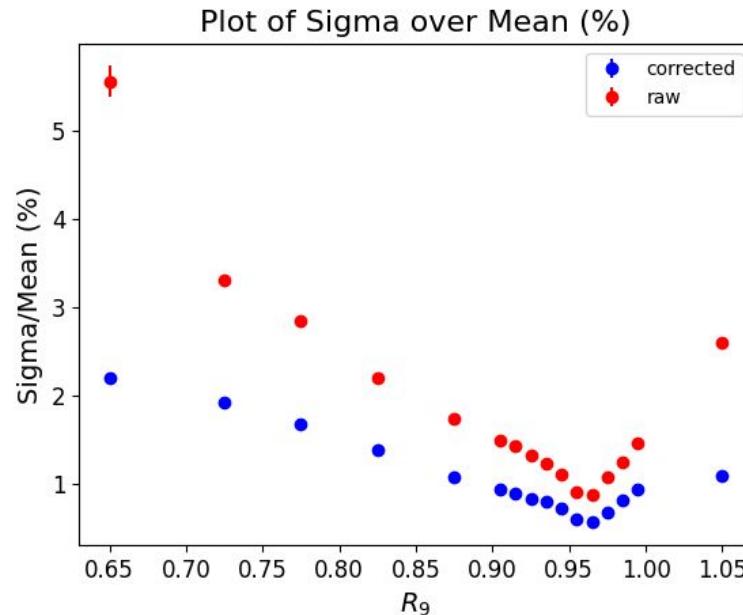
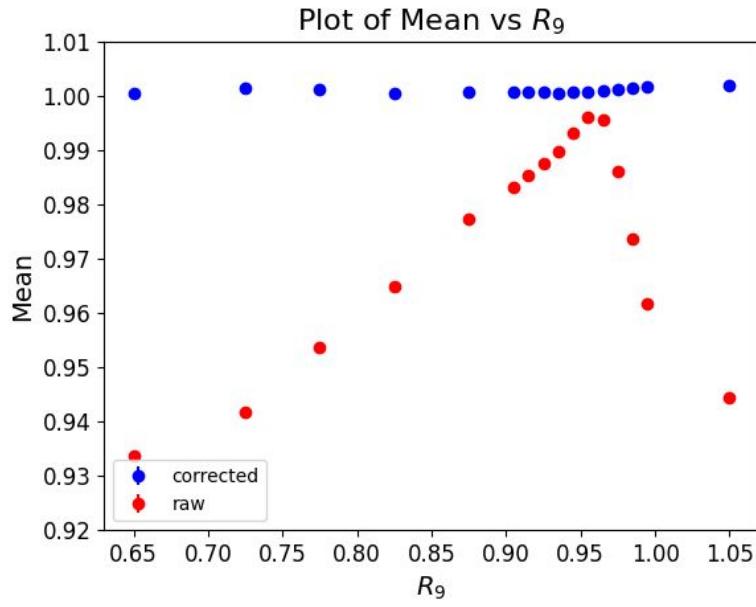


Response fitted in 0.98-0.99 bin of  $R_9$  distribution



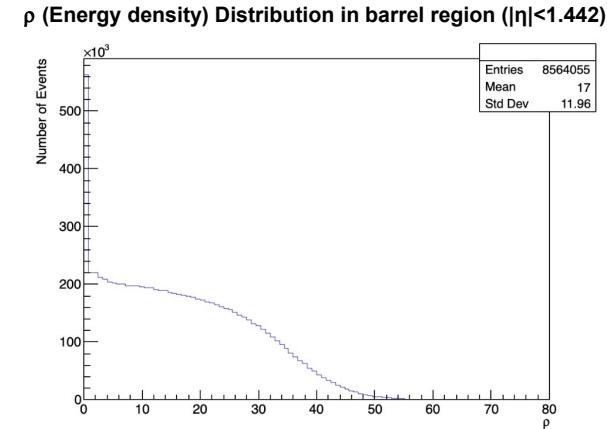
# Relative response and mean of response with Mean $R_9$ bins

- This is because as the  $R_9$  increases , there is less showering due to tracker →less loss of energy .
- Stability of corrected response  $R_9$  Mean is 0.1337 %.

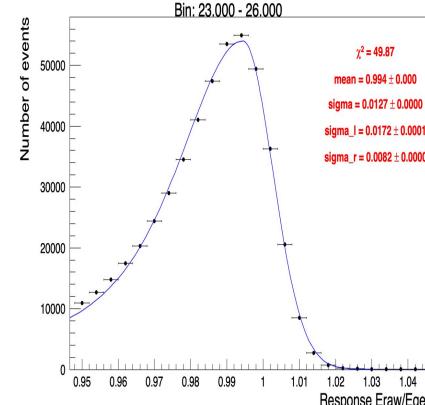
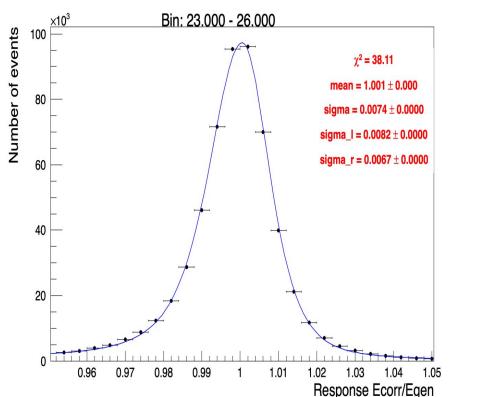


# Variation of relative energy response with energy density $\rho$

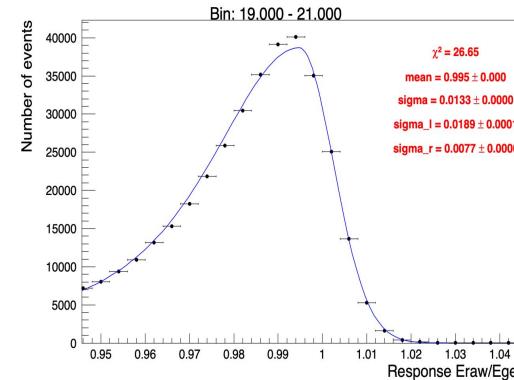
- It is defined as the energy density due to soft interaction .
- Higher  $\rho$  mean higher pile-up.
- Here also variable binning .



Response fitted in 23-26 bin of  $\rho$  distribution

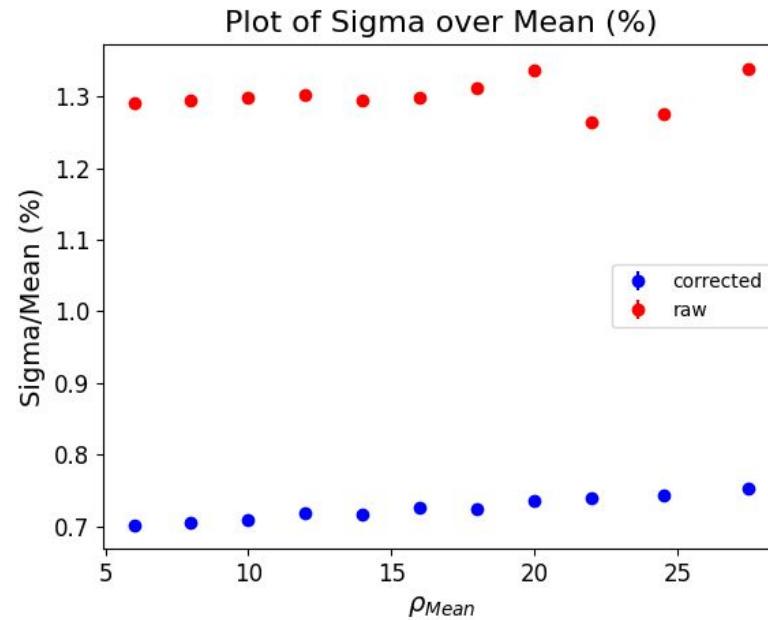
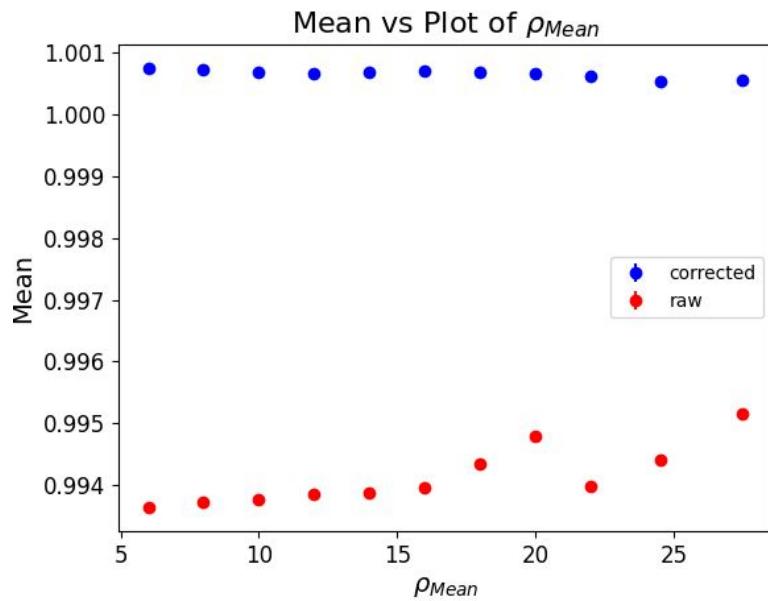


Response fitted in 19-21 bin of  $\rho$  distribution



## Relative response and mean of response with Mean $\rho$ bins

- Higher  $\rho$  means higher pile up , hence relative energy resolution increases
- Stability of corrected response  $\rho$  Mean is 0.0214 %.



# Conclusion

We have observed that the response of  $E_{\text{Raw}}$  vs  $E_{\text{Corrected}}$  varies across different parameters, with  $E_{\text{Corrected}}$  exhibiting comparatively lower relative resolution of response of ECAL. Therefore, all the corrections applied to  $E_{\text{Corrected}}$  effectively enhance our detector's response.

# References

- Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC- [arXiv:2012.06888 \[hep-ex\]](https://arxiv.org/abs/1206.8888)
- CMS Physics Technical Design Report Volume I: Detector Performance and Software : <https://cds.cern.ch/record/922757/files/lhcc-2006-001.pdf>
- [https://iopscience.iop.org/article/10.1088/1742-6596/293/1/012030/pdf#:~:text=The%20energy%20deposited%20in%20the,crystals%20around%20the%20impact%20point.](https://iopscience.iop.org/article/10.1088/1742-6596/293/1/012030/pdf#:~:text=The%20energy%20deposited%20in%20the,crystals%20around%20the%20impact%20point)

# THANK YOU

# BACKUP SLIDES

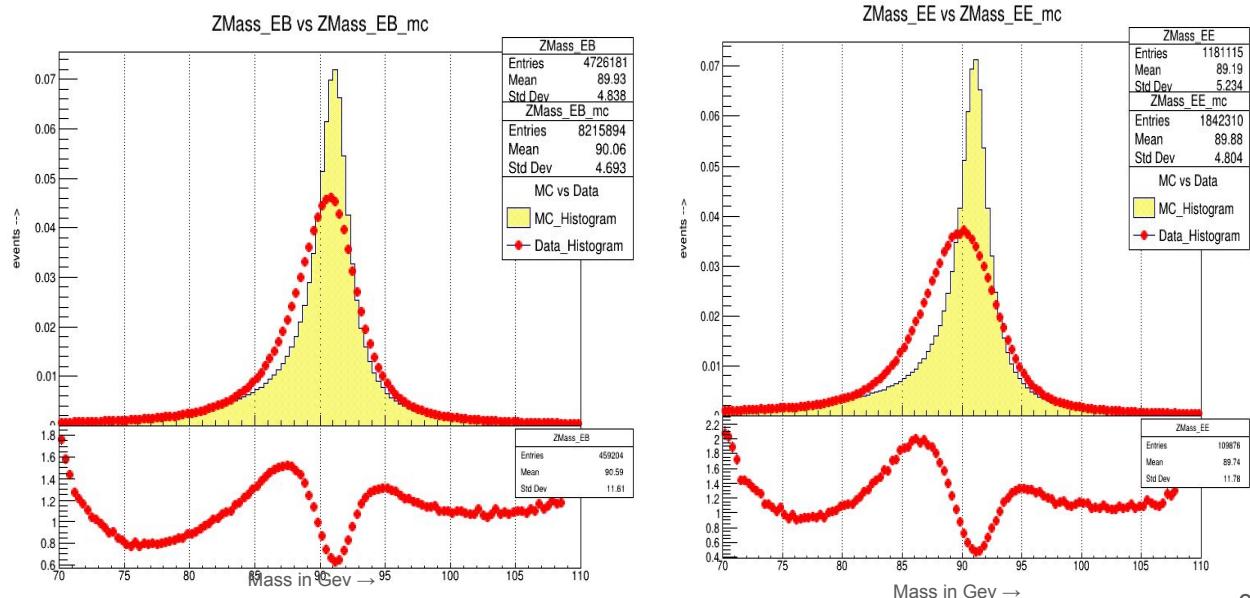
# More work to show

# SCALE AND SMEARINGS

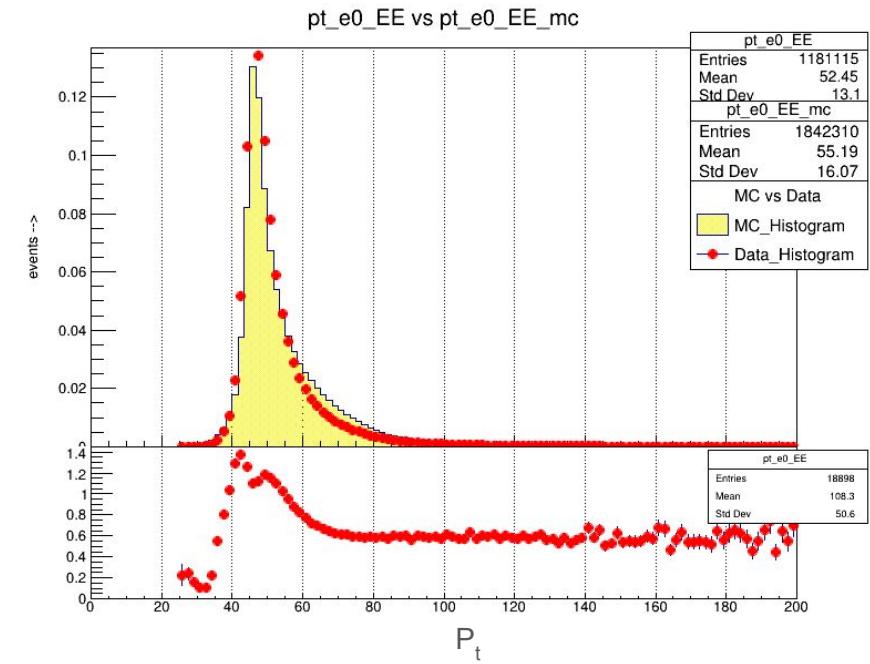
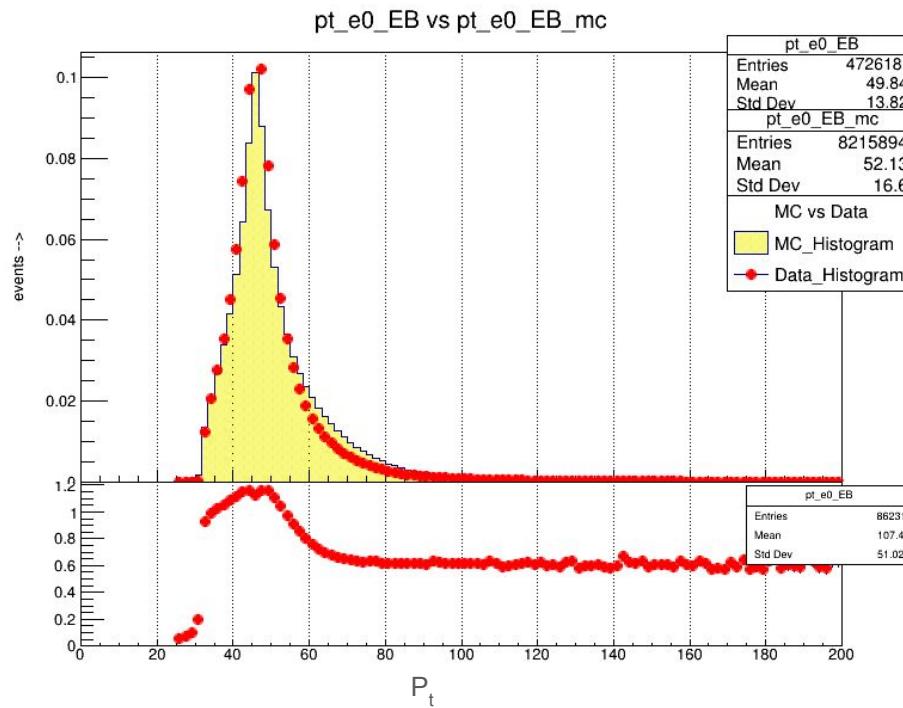
# The need for residual scales and additional smearings

While the regression corrects for energy loss in material in front of ECAL, variation of cluster response, effects of pileup, etc there are differences w.r.t data that is beyond the scope of the regression -

- The energy scale is correction based on training with an idealized IC sample.
- Non-uniformity in light collection
- Modelling of upstream material



# Pt distribution of the data



# Scale stabilization

This function derives the scales known as 'step1' scales. These scales correct data to the pdg Z mass in bins of Run and Eta to stabilize the scale as a function of time and location in the detector.

## Scale stabilization Dataset

Z-> ee data and MC from UL 2018 datasets

**Data:**

2018A :

```
/eos/cms/store/group/dpg_ecal/alca_ecalcalib/ecalelf/ntuples/13TeV/MINIAODNTUPLE//106X_dataRun2_UL18/EGamma-106X-2018UL-A/315410-315430/2018_314472-325175_SS/2018UL_SS/EGamma-106X-2018UL-A-315410-315430.root
```

2018B :

```
/eos/cms/store/group/dpg_ecal/alca_ecalcalib/ecalelf/ntuples/13TeV/MINIAODNTUPLE//106X_dataRun2_UL18/EGamma-106X-2018UL-B/317180-317650/2018_314472-325175_SS/2018UL_SS/EGamma-106X-2018UL-B-317180-317650.root
```

2018C :

```
/eos/cms/store/group/dpg_ecal/alca_ecalcalib/ecalelf/ntuples/13TeV/MINIAODNTUPLE//106X_dataRun2_UL18/EGamma-106X-2018UL-C/320060-320070/2018_314472-325175_SS/2018UL_SS/EGamma-106X-2018UL-C-320060-320070.root
```

2018D :

```
/eos/cms/store/group/dpg_ecal/alca_ecalcalib/ecalelf/ntuples/13TeV/MINIAODNTUPLE//106X_dataRun2_UL18/EGamma-106X-2018UL-D/324020-325200/2018_314472-325175_SS/2018UL_SS/EGamma-106X-2018UL-D-324020-325200.root
```

**MC:**

```
/eos/cms/store/group/dpg_ecal/alca_ecalcalib/ecalelf/ntuples/13TeV/MINIAODNTUPLE/106X_MCRun2_UL18/DYJets_M50-CP5-amcatnlo-Run2018UL-MC-Scale/allRange/2016ULMCpostVFP/DYJets_M50-CP5-amcatnlo-Run2018UL-MC-Scale-allRange.root
```

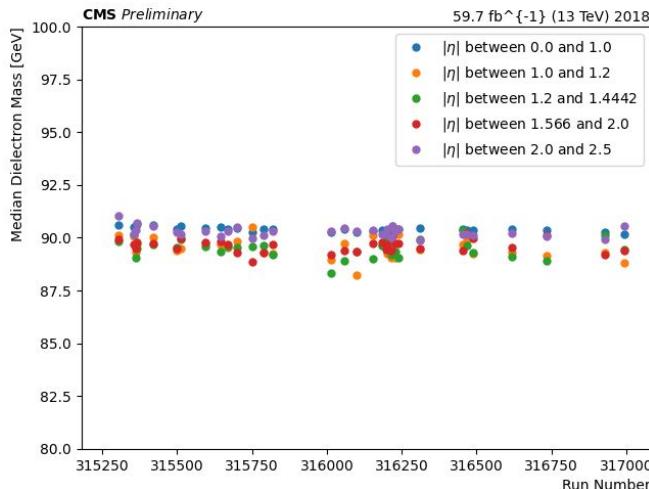
# Median method of corrections

On plotting the median energies → they are varying as eta and run-number .

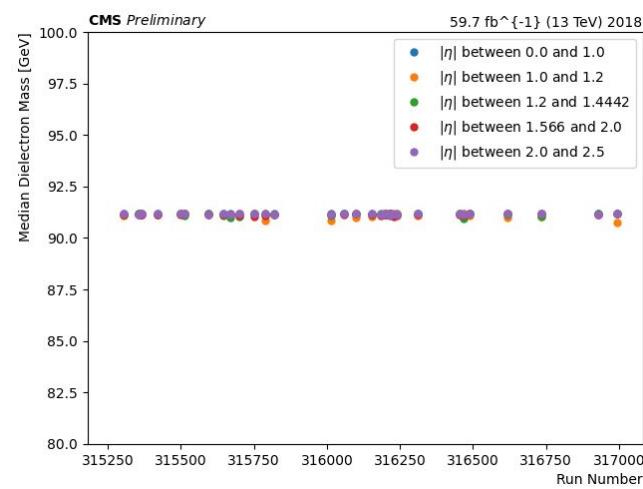
To correct this → Making it linear and stabilize it→multiplied each median by pdg Z mass/median

so that each median now have a value equal to pdg z.

Before correction



After correction



## R9 explanation

Relative response decreases with increasing value of r9 , because it shows less spreading due to early shower hence it would be easy for mustache algorithm to recover the hits due to early shower, after certain point its again increases because here due to corrections/thresholds that are only applied to Esupercluster only due to which these are r9 which more spread but are there because of the more correction to ESupercluster .

# Pile up

During multiple proton proton bunch collision , there are so many particles that are produced that are not of our interest but still deposited energy on the detector , this effect is called pile up.

## Bunch info

A proton bunch in the Large Hadron Collider (LHC) at CERN contains around  $1.15 \times 10^{11}$  protons. The bunches are typically 7–10 cm long and travel in the beam pipe almost parallel to the outgoing beams.

# Choice of electromagnetic calorimetric material

Crystal		NaI(Tl)	CsI(Tl)	CsI	BaF <sub>2</sub>	BGO	CeF <sub>3</sub>	PbWO <sub>4</sub>	LAr	Plastic	Pb	W	Cu	Fe	U
Density	gm/cm <sup>3</sup>	3.67	4.51	4.51	4.89	7.13	6.16	8.28	1.4	1.03	11.4	19.3	8.96	7.87	19.0
Rad. Length	cm	2.59	1.85	1.85	2.06	1.12	1.68	0.89	13.5	42.4	0.56	0.35	1.43	1.7	0.32
Moliére rad	cm	4.5	3.8	3.8	3.4	2.4	2.6	2.2							
Inter. length	cm	41.4	36.5	36.5	29.9	22.0	25.9	22.4	65.0	78.9	17.6	9.6	15.1	16.7	11.0
Decay time	ns	250	1000	35	630	300	10	15		1 - 5					
				6	0.9		30	5							
Peak emission	nm	410	565	420	300	480	310	420		370-430					
				310	220		340	440							
Rel. light yield	%	100	45	5.6	21	9	10	0.7		28-34					
				2.3	2.7										
D(LY)/dT	%/°C	≈0	0.3	-0.6	-2.0	-1.6	0.15	-1.9	-						
r.i. (n)		1.85	1.80	1.80	1.56	2.20	1.68	2.16	1.6	1.58					

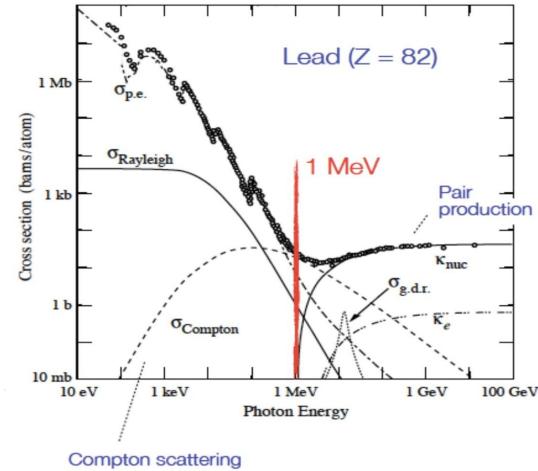
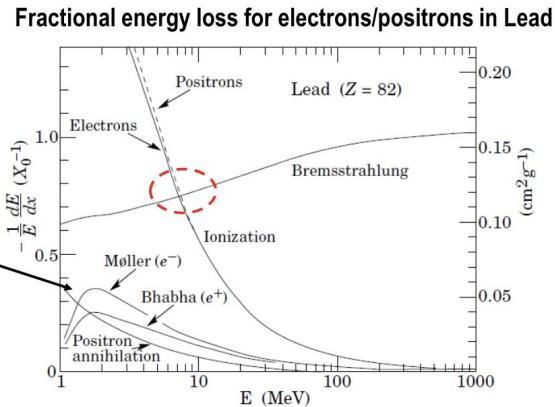
NaI(Tl) : Light output = 7%

LAr : dE/dx=2.2 MeV/cm, mobility ~ 5 mm/μs at 1 KV/mm

Radiation hardness

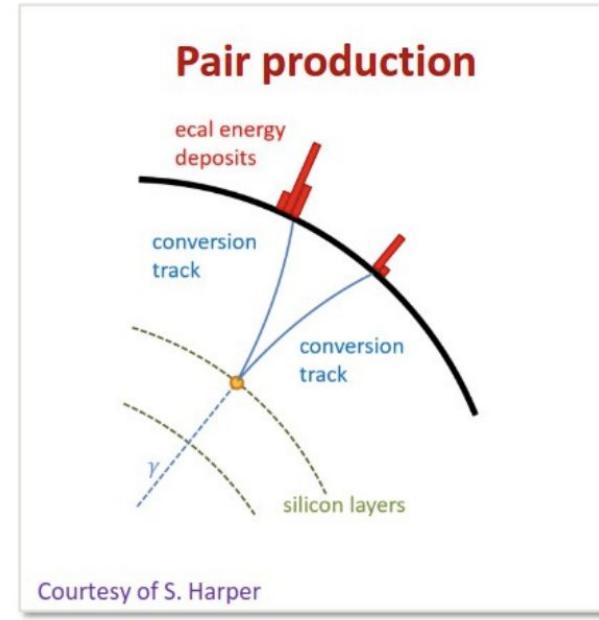
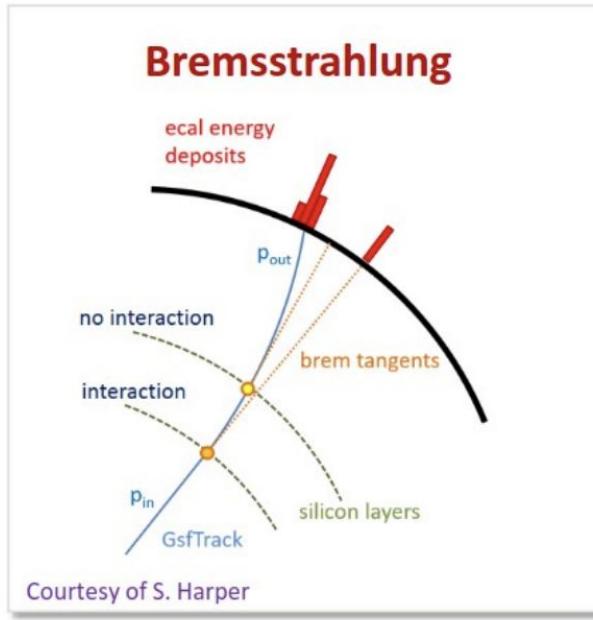
# Matter interaction

Other processes  
(Bhabha, meller, ...) neglected in HEP  
(most of the time)



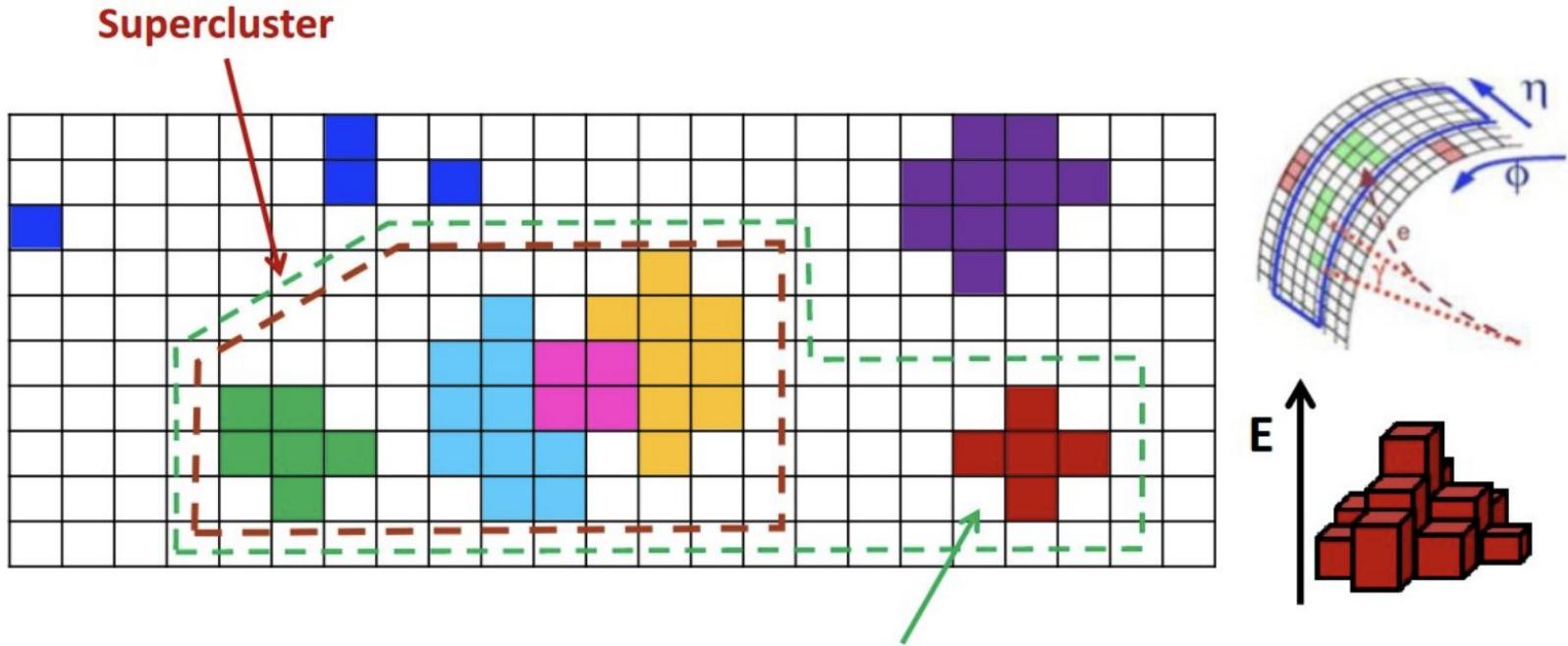
- For high energy electron, Bremsstrahlung is the major cause of energy loss.
- For high energy photon, pair production is the dominant process.

# Bremsstrahlung and photon conversion recovery



- Tangents to the tracks are extrapolated to ecal to recover bremsstrahlung photons

# Refined Supercluster



# Why we use cruijff function

Because of tails , because of low pt we have more event to event fluctuation hence our Eraw/Egen will have more left tail or longer left tail.

Equation (2.26) is commonly known as the *Bethe-Bloch formula* and is the basic expression used for energy loss calculations. In practice, however, two corrections are normally added: the *density effect* correction  $\delta$ , and the *shell* correction  $C$ , so that

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[ \ln \left( \frac{2m_e \gamma^2 v^2 W_{\max}}{I^2} \right) - 2\beta^2 - \delta - 2 \frac{C}{Z} \right], \quad (2.27)$$

with

$$2\pi N_a r_e^2 m_e c^2 = 0.1535 \text{ MeV cm}^2/\text{g}$$

$r_e$ : classical electron radius = $2.817 \times 10^{-13}$ cm	$\rho$ : density of absorbing material
$m_e$ : electron mass	$z$ : charge of incident particle in units of $e$
$N_a$ : Avogadro's number = $6.022 \times 10^{23}$ mol $^{-1}$	$\beta$ = $v/c$ of the incident particle
$I$ : mean excitation potential	$\gamma$ = $1/\sqrt{1-\beta^2}$
$Z$ : atomic number of absorbing material	$\delta$ : density correction
$A$ : atomic weight of absorbing material	$C$ : shell correction
	$W_{\max}$ : maximum energy transfer in a single collision.

The maximum energy transfer is that produced by a head-on or *knock-on* collision. For an incident particle of mass  $M$ , kinematics gives

$$W_{\max} = \frac{2m_e c^2 \eta^2}{1 + 2s\sqrt{1 + \eta^2 + s^2}}, \quad (2.28)$$

where  $s = m_e/M$  and  $\eta = \beta\gamma$ . Moreover, if  $M \gg m_e$ , then

$$W_{\max} \approx 2m_e c^2 \eta^2.$$

**The Mean Excitation Potential.** The mean excitation potential,  $I$ , is the main parameter of the Bethe-Bloch formula and is essentially the average orbital frequency  $\bar{\nu}$  from Bohr's formula times Planck's constant,  $\hbar\bar{\nu}$ . It is theoretically a logarithmic average of

$\nu$  weighted by the so-called oscillator strengths of the atomic levels. In practice, this is a very difficult quantity to calculate since the oscillator strengths are unknown for most materials. Instead, values of  $I$  for several materials have been deduced from actual measurements of  $dE/dx$  and a semi-empirical formula for  $I$  vs  $Z$  fitted to the points. One such formula is

$$\begin{aligned} \frac{I}{Z} &= 12 + \frac{7}{Z} \text{ eV} & Z < 13 \\ \frac{I}{Z} &= 9.76 + 58.8 Z^{-1.19} \text{ eV} & Z \geq 13. \end{aligned} \quad (2.29)$$

It has been shown, however, that  $I$  actually varies with  $Z$  in a more complicated manner [2.2]. In particular, there are local irregularities or *wiggles* due to the closing of certain atomic shells. Improved values of  $I$  are given in Table 2.1 for several materials. A more extensive list may be found in the articles by Sternheimer et al. [2.2 – 3].

**The Shell and Density Corrections.** The quantities  $\delta$  and  $C$  are corrections to the Bethe-Bloch formula which are important at high and low energies respectively.

The *density effect* arises from the fact that the electric field of the particle also tends to polarize the atoms along its path. Because of this polarization, electrons far from the path of the particle will be shielded from the full electric field intensity. Collisions with these outer lying electrons will therefore contribute less to the total energy loss than predicted by the Bethe-Bloch formula. This effect becomes more important as the particle energy increases, as can be seen from the expression for  $b_{\max}$  in (2.24). Clearly as the velocity increases, the radius of the cylinder over which our integration is performed also increases, so that distant collisions contribute more and more to the total energy loss. Moreover, it is clear that this effect depends on the density of the material (hence the term "density" effect), since the induced polarization will be greater in condensed materials than in lighter substances such as gases. A comparison of the Bethe-Bloch formula with and without corrections is shown in Fig. 2.3.

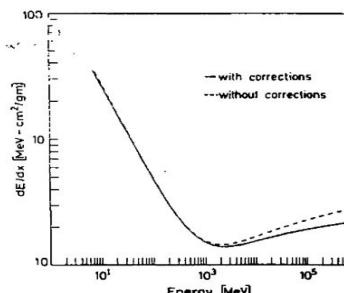
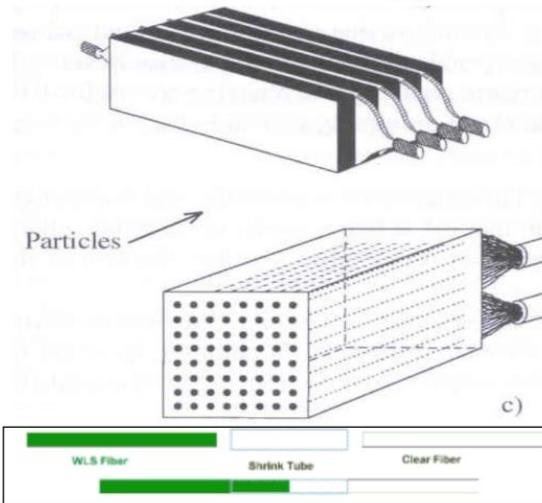


Fig. 2.3. Comparison of the Bethe-Bloch formula with and without the shell and density corrections. The calculation shown here is for copper

# Energy resolution of EM calorimeter



- (c1) Electronic Noise
- (c2) Particle other the one in interest, e.g., pile-up
- (c3) Analog to digital : loss of information
- (c4) Shift in pedestal level
- (b1) Fluctuation in cascading, charges/neutral ratio, sampling etc.
- (b2) Photon/p.e. statistics

Overall resolution : 
$$\frac{\sigma_E}{E} = a \oplus \frac{b}{\sqrt{E}} \oplus \frac{c}{E}$$

- (a1) Non-uniformity in signal generation, e.g., thickness of scintillator, uniformity of scintillator properties, position of shower
- (a2) Collection of photon : crystal shape, fraction of crystal surface covered by the PMT, reflectivity at surface, self attenuation
- (a3) Propagation of photons, attenuation, surface loss, bending of fibre
- (a4) Loss in splice, connectors
- (a5) Cell-to-cell Inter calibration error
- (a6) Non containment of shower, energy leakage in rear/side ( $E^{-(1/4)}$ ), albedo
- (a7) Energy deposit in dead areas in front or inside the calorimeter
- (a8) Fluctuation in timing measurement( TDC)
- (a9) Position dependent QE of photon-transducer (and/or cell-to-cell variation), e.g. PMT/SiPM
- (a10) Gain of PMT/APD/SiPM + HV stability
- (a11)  $dL/dT$ , variation with temperature
- (a12) Gas composition, contamination of electronegative substance(in particular, oxygen), temperature, pressure

# PILE UP

Pile-up in the context of the CMS detector at CERN refers to multiple proton-proton collisions occurring simultaneously or nearly simultaneously during a single beam crossing in the Large Hadron Collider (LHC). These overlapping collisions can complicate data analysis by creating additional, unwanted signals that can interfere with the detection and measurement of particles from the primary collision of interest.

## Types of Pile-up:

- **In-time Pile-up:** Multiple interactions occurring within the same bunch crossing.
- **Out-of-time Pile-up:** Interactions from adjacent bunch crossings that overlap due to the finite resolution of the detector's timing.

# STOCHASTIC TERM

## 1. Lateral Shower Containment

- **Lateral Shower Shape:** When a high-energy particle enters a detector, it creates an electromagnetic shower, depositing energy across several crystals around the impact point. The shape of this lateral shower can vary from one event to another.
- **Containment Fluctuations:** The energy deposited in the active medium of each crystal fluctuates due to differences in the lateral development of the shower. This variability affects the total energy measured because the crystals might capture different portions of the shower in different events.

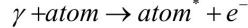
## 2. Photo-electron Statistics

**Photo-electron Production:** When the deposited energy in the crystals is converted into light and subsequently detected by photodiodes, the number of photo-electrons produced varies due to statistical fluctuations.

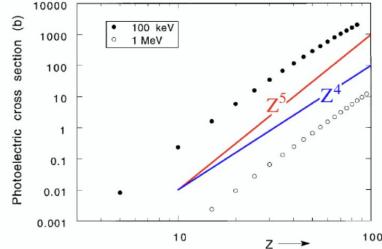
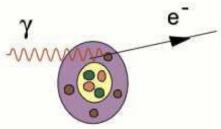
- **Statistical Nature:** The number of photo-electrons generated follows a Poisson distribution, introducing a degree of randomness in the signal. This contributes to the overall uncertainty in the energy measurement.

## Photons: Photo-Electric effect

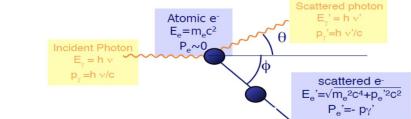
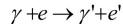
- Photon extract an electron from the atom



$$\sigma_{pe} \approx Z^5 \alpha^4 \left( \frac{m_e c^2}{E_\gamma} \right)^{7/2}$$



## Photons: Compton scattering



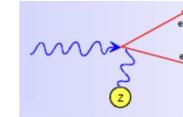
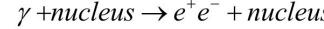
$$\sigma_{Compton} \approx Z \frac{\ln E_\gamma}{E_\gamma}$$

### Remarks:

- $\sigma_{Compton} \propto Z, E^{-1}$
- Electrons are emitted (more or less) isotropically

## Photons: Pair production

- Can only occurs in the Coulomb field of a nucleus (or an electron) if  $E_\gamma > 2m_e c^2$



$$\sigma_{pair} \approx 4\alpha r_e^2 Z^2 \left( \frac{7}{9} \ln \frac{183}{Z^{1/3}} \right) \approx \frac{7}{9} \frac{A}{N_A} \frac{1}{X_0}$$

- Mean free path of photon before it creates a pair

$$\lambda_{pair} \approx \frac{9}{7} X_0$$

### Remarks:

- $\sigma_{pair} \propto Z(Z+1)$
- Photons have a high penetrating power than electrons
- Pair creation is independent of incident energy (for  $E_\gamma > 1$  GeV)
- $e^+ e^-$  is emitted in photon direction

[https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204\\_calorimetry\\_AppendixEMshowers.pdf](https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204_calorimetry_AppendixEMshowers.pdf)

## Bremsstrahlung

- Radiation of real photons in the Coulomb field of the atomic nuclei
- Dominant process at high energy

$$\left( -\frac{dE}{dx} \right)_{rad} = 4\alpha N_A \frac{Z^2}{A} z^2 \left( \frac{1}{4\pi\epsilon_0} \frac{e^2}{mc^2} \right)^2 E \ln \frac{183}{Z^{1/3}} \propto \frac{E}{m^2}$$

- Important for electrons, much less for muons (apart from ultra-relativistic)

$$\left( -\frac{dE}{dx} \right)_{rad} = 4\alpha N_A \frac{Z^2}{A} r_e^2 E \ln \frac{183}{Z^{1/3}} \quad (\text{for electrons})$$

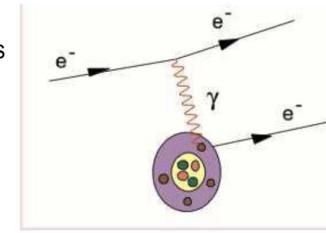
Radiation length

- Conveniently re-written as:  $\left( \frac{dE}{dx} \right)_{rad} = \frac{E}{X_0}$

$$X_0 = \frac{A}{4\alpha N_A Z^2 r_e^2 \ln \frac{183}{Z^{1/3}}}$$

## Ionization

- Interaction of charged particles with electron cloud of atoms (loss of electrons, atoms → ions)
- Dominant process at low energy



- Bethe-Bloch formula (general)

$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] (\text{MeV.g}^{-1}.\text{cm}^2)$$

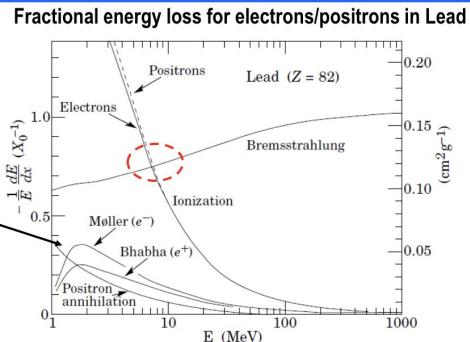
Energy loss depends:

- quadratic ally on the charge and velocity of the incident particle (but not on its mass)
- Linearly on the material (through electron density)
- Logarithmically on the material (through mean ionization I)

[https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204\\_calorimetry\\_AppendixEMshowers.pdf](https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204_calorimetry_AppendixEMshowers.pdf)

## Critical Energy

Other processes  
(Bhabha, meller, ...) neglected in HEP  
(most of the time)



- Radiation (ionization) dominant at high (low) energies
- Crossing point:  $\left(\frac{dE}{dx}\right)_{rad} (E_c) = \left(\frac{dE}{dx}\right)_{ionz} (E_c)$  **E<sub>c</sub>: critical energy** *Strongly material dependent (scales as 1/Z)*
- Examples:

Material	W	Pb	(liquid) Ar	Cu
Z	74	82	29	13
E <sub>c</sub> (MeV)	8,4	7,1	37	20,2

$$E_c(solid) = \frac{610 \text{ MeV}}{Z + 1.24}$$

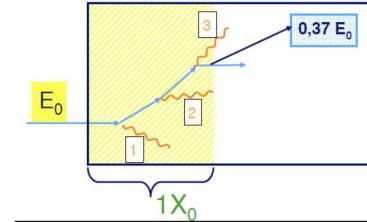
$$E_c(liquid) = \frac{710 \text{ MeV}}{Z + 0.92}$$

## Radiation Length

➤ **Definition:** mean distance over which the incident electron loses all BUT 1/e ≈ 37% of its incident energy via radiation (ie, it radiated ≈63% of its incident energy)

$$\left(\frac{dE}{dx}\right)_{rad} = \frac{E}{X_0} \implies \frac{dE}{E} = \frac{dx}{X_0}$$

$$E = E_0 e^{-x/X_0}$$



➤ Useful approximation:

$$X_0 \approx \frac{180A}{Z^2} \text{ (g.cm}^{-2}\text{)} \quad \text{Also in cm (taking into account density)}$$

### Examples:

Material	W	Pb	Cu	Al	Stainless Steel	PbWO4	(dry) Air	(liquid) Water
Z	74	82	29	13	-	-	-	-
X <sub>0</sub> (cm)	0,35	0,56	1,4	8,9	1,76	0,89	30390	36,08

[https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204\\_calorimetry\\_AppendixEMshowers.pdf](https://indico.cern.ch/event/782305/contributions/3256087/attachments/1789793/2916004/20190204_calorimetry_AppendixEMshowers.pdf)

## EM shower: Moliere Radius

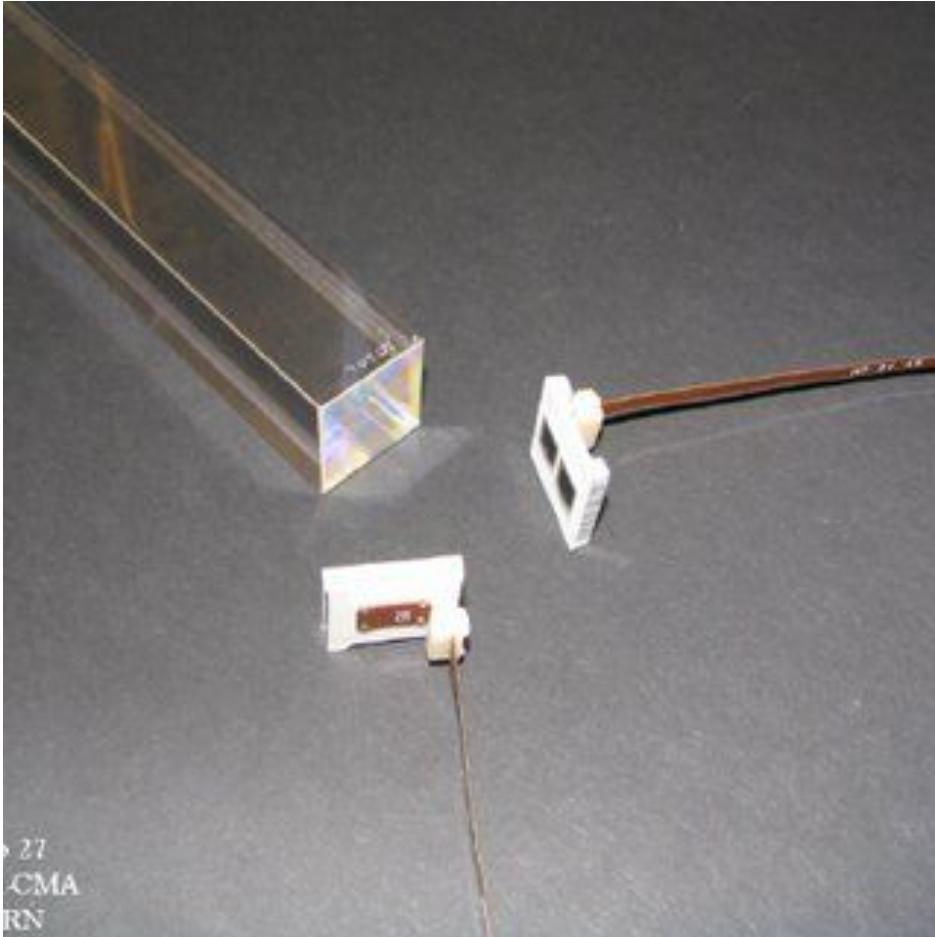
---

- **Moliere radius:** characteristic of a material giving the scale of the transverse dimension of an EM shower

$$R_M = \frac{21MeV}{E_C} X_0 \quad (\text{g.cm}^{-2})$$

Scales as A/Z, while X<sub>0</sub> scales as A/Z<sup>2</sup>. much less dependent on material than X<sub>0</sub> !

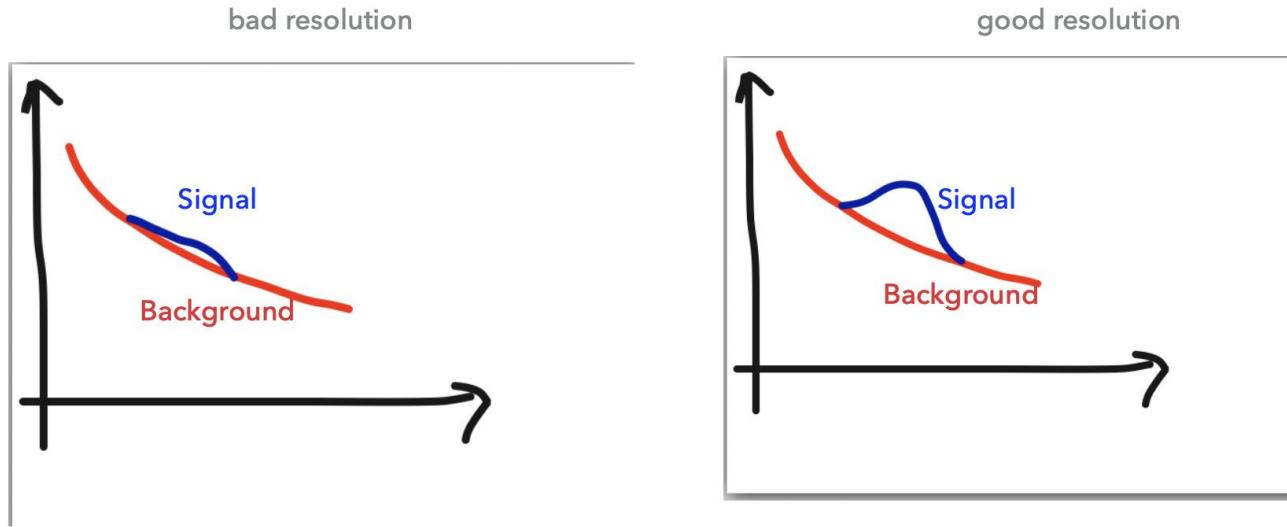
- 90% of shower energy contained in a cylinder of 1R<sub>m</sub>
- 95% of shower energy contained in a cylinder of 2R<sub>m</sub>
- 99% of shower energy contained in a cylinder of 3.5R<sub>m</sub>



27  
CMA  
RN

# WHY IS RESOLUTION IMPORTANT

4



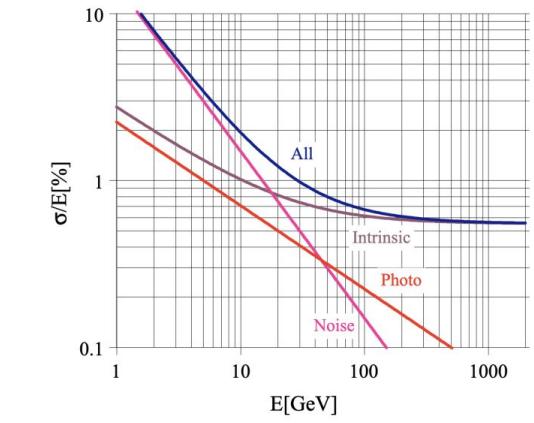
- ▶ In which scenario is it easy to discover signal?

Dr. Shilpi Jain  
ehep\_2024 school  
slides onwards

# RESOLUTION

Reference: [CMS ECAL TDR](#) 6

Energy resolution of a homogeneous calorimeter



$$(\sigma/E)^2 = (a/\sqrt{E})^2 + (\sigma_n/E)^2 + c^2 \quad (\text{E in GeV})$$

Photo statistics  
Noise  
Constant

$a = 2.8\%$   
 $\sigma_n = 12\%$   
 $c = 0.3\%$

"+" means addition in quadrature

## ▶ Photo-statistics:

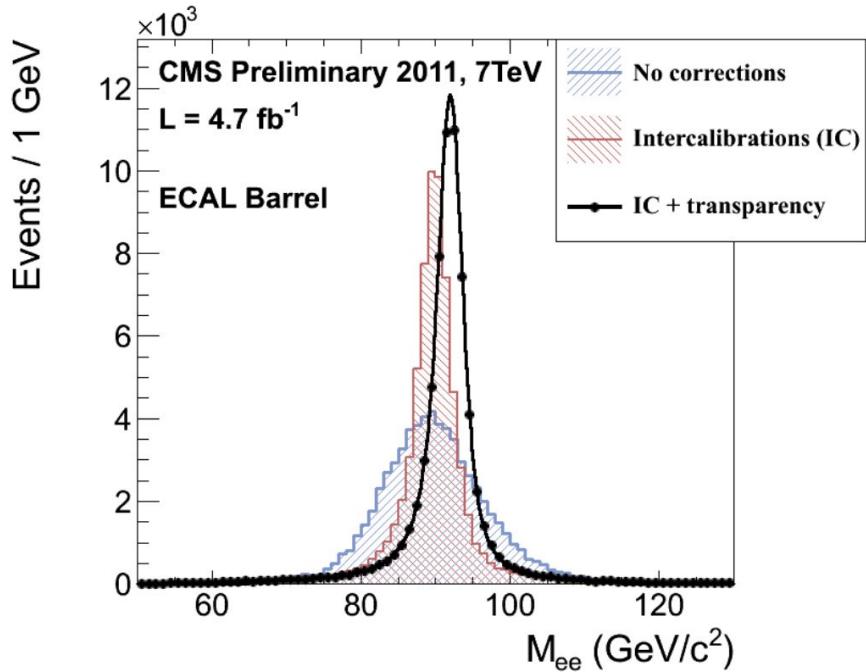
- ▶ Due to number fluctuations in the number of shower particles
- ▶ Noise: Electronic noise and the pileup energy
- ▶ Corresponds to 40 MeV in the EB and 60 MeV in the EE
- ▶ Constant term: Due to limited shower containment, longitudinal light collection, inter-calibration
- ▶ Dominates for high energy electrons and photons

## INTERCALIBRATION CONSTANTS (IC) - THE C TERM

14

- ▶ All the crystals in ECAL do not have same response.
- ▶ What this means is that a 10 GeV electron can give a final reconstructed signal of 5 GeV in 1 crystal and full 10 GeV in another crystal.
- ▶ So essentially the intercalibration constant for the 1st crystal = 2.
- ▶ Resolution is written as: 
$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$
  - ▶ where a: statistical term; b is the noise term and c is the constant term.
  - ▶ For high enough energies (typically > 100 GeV), c is the dominating term
- ▶ ICs have the dominant contribution in the constant term of the resolution
  - ▶ Important to have the precise estimation of ICs
- ▶ To equilize the response, dedicated methods are performed to get the ICs:
  - ▶ phi symmetry method: Azimuthal symmetry of the energy distribution
  - ▶ pi0 and eta mesons: Exploits the peak position in the invariant mass distribution of pi0 or eta mesons
  - ▶ W/Z electrons: same as above but the mass is that of Z. In case of electrons coming from W, it is the E/p ratio which is used.
- ▶ Reference: <https://arxiv.org/pdf/1306.2016.pdf>

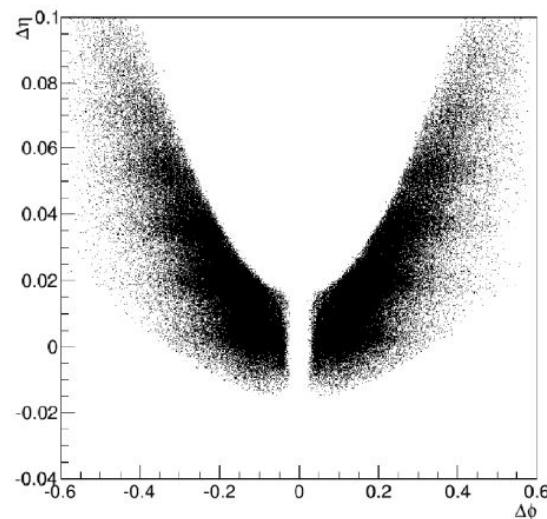
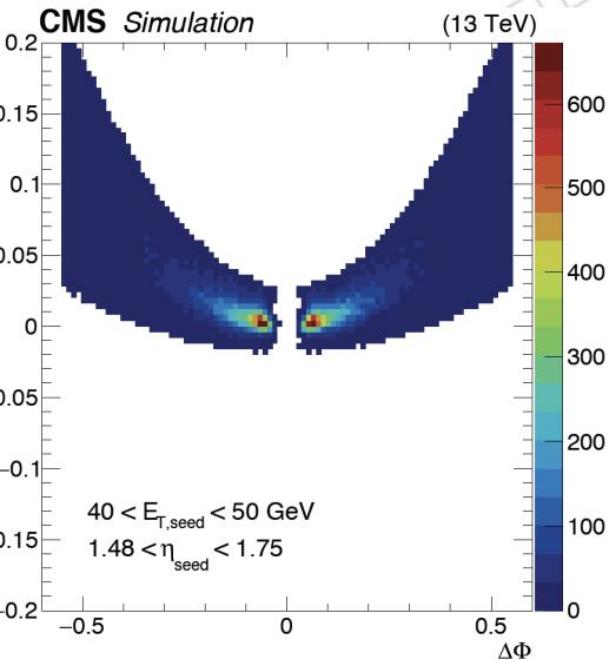




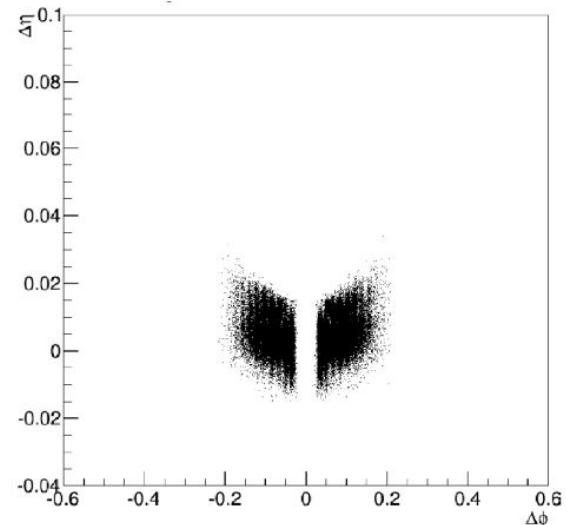
- ▶ Summary: transparency corrections essentially correct for the energy scale (and hence resolution). ICs correct for resolution mostly

# MUSTACHE SUPERCLUSTER

28



(a)  $0.5 < E_T(\text{subcluster}) < 1.0 \text{ GeV}$



(b)  $4 < E_T(\text{subcluster}) < 6 \text{ GeV}$

EGM-17-001

# ECAL READOUT CHAIN

54

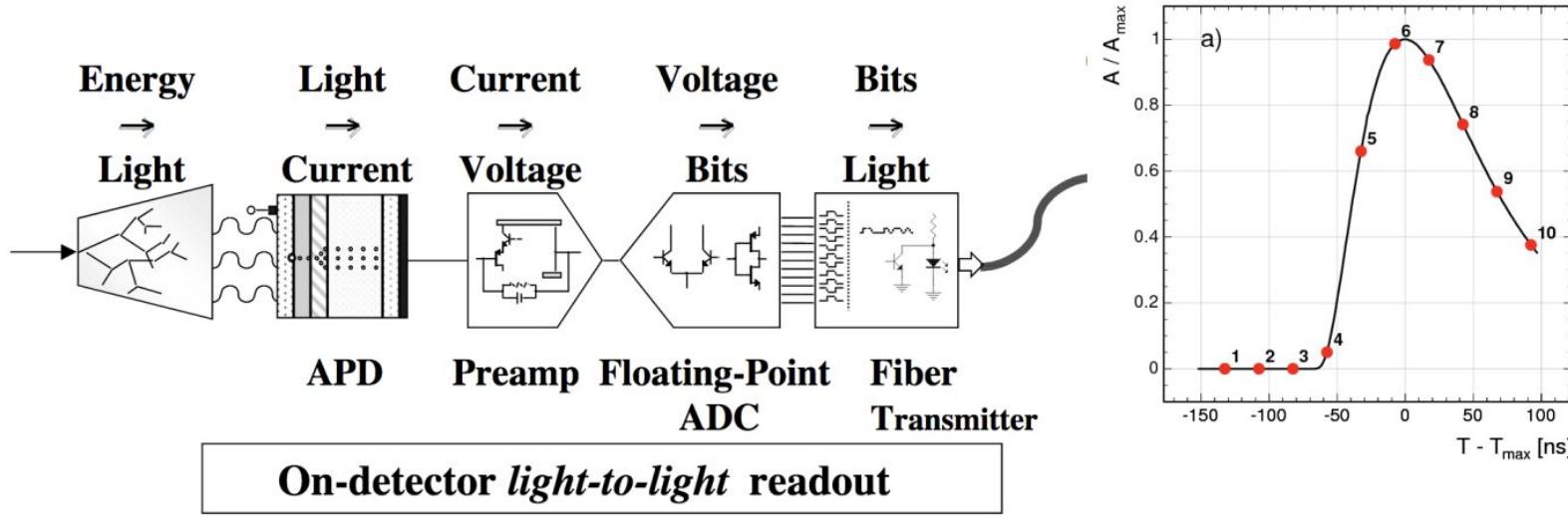
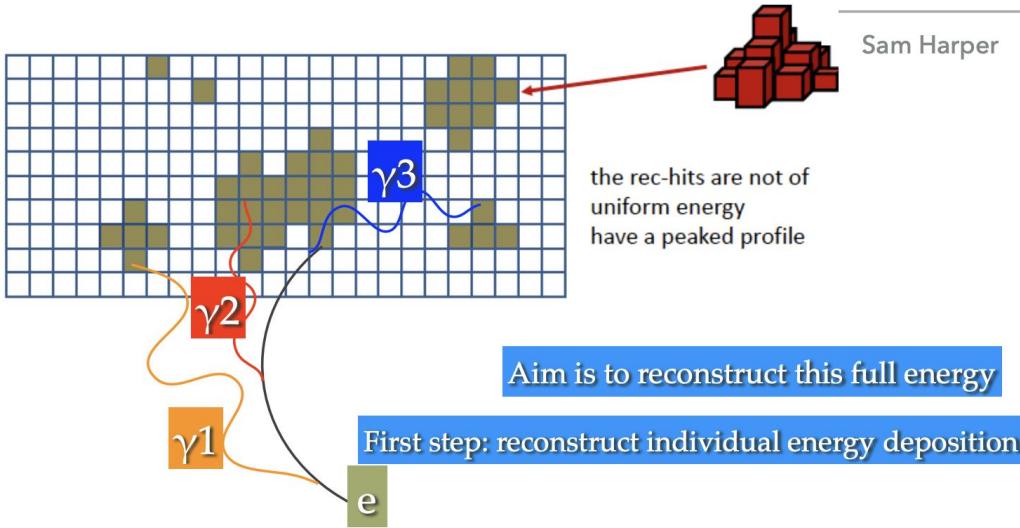
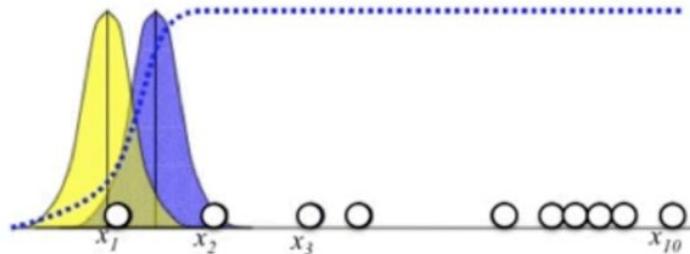


Fig. 5.4: ECAL readout chain.



- An electron or a photon deposits energy in several crystals
- Aim: reconstruct the full energy of electron or a photon
- Once we have the crystal energies reconstructed, we need to reconstruct individual particles from deposit of a single photon/electron
  - i.e. in this picture reconstructing individual energy deposits from  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$  and the final electron
  - this is the PF (particle flow) clustering step - i.e. making small clusters
  - These small clusters are then combined to form super-clusters

## EM: 1-d example



$$P(x_i | b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(x_i - \mu_b)^2}{2\sigma_b^2}\right\}$$

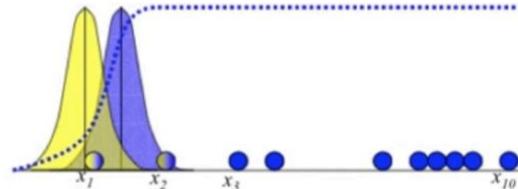
$$b_i = P(b | x_i) = \frac{P(x_i | b)P(b)}{P(x_i | b)P(b) + P(x_i | a)P(a)}$$

- ▶ Start with some initial values of means ( $\mu$ ) and sigma ( $\sigma$ ) of two gaussians 'a' and 'b'
- ▶ Calculate the probability that given Gaussian 'a' with mean  $\mu_a$  and  $\sigma_a$ , what is the probability of getting point  $x_i$  for Gaussian 'a'
- ▶ Same for Gaussian 'b'

## STEP 2: MAXIMIZATION STEP

24

### EM: 1-d example



$$P(x_i | b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(x_i - \mu_b)^2}{2\sigma_b^2}\right\}$$

$$b_i = P(b | x_i) = \frac{P(x_i | b)P(b)}{P(x_i | b)P(b) + P(x_i | a)P(a)}$$

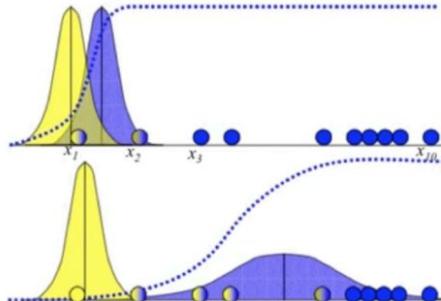
$$a_i = P(a | x_i) = 1 - b_i$$

$$\mu_b = \frac{b_1 x_1 + b_2 x_2 + \dots + b_n x_n}{b_1 + b_2 + \dots + b_n}$$

$$\sigma_b^2 = \frac{b_1(x_1 - \mu_b)^2 + \dots + b_n(x_n - \mu_b)^2}{b_1 + b_2 + \dots + b_n}$$

- ▶ For Gaussian 'a', calculate  $\mu_a$  and  $\sigma_a$  again using these estimated probabilities by summing over all the hits
- ▶ Same for Gaussian 'b'
- ▶ Now you have updated  $\mu_a$  and  $\sigma_a$  for Gaussian 'a' and  $\mu_b$  and  $\sigma_b$  for Gaussian 'b'

## EM: 1-d example



$$P(x_i | b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{(x_i - \mu_b)^2}{2\sigma_b^2}\right\}$$

$$b_i = P(b | x_i) = \frac{P(x_i | b)P(b)}{P(x_i | b)P(b) + P(x_i | a)P(a)}$$

$$a_i = P(a | x_i) = 1 - b_i$$

$$\mu_b = \frac{b_1 x_1 + b_2 x_2 + \dots + b_n x_n}{b_1 + b_2 + \dots + b_n}$$

$$\sigma_b^2 = \frac{b_1(x_1 - \mu_b)^2 + \dots + b_n(x_n - \mu_b)^2}{b_1 + b_2 + \dots + b_n}$$

$$\mu_a = \frac{a_1 x_1 + a_2 x_2 + \dots + a_n x_n}{a_1 + a_2 + \dots + a_n}$$

$$\sigma_a^2 = \frac{a_1(x_1 - \mu_a)^2 + \dots + a_n(x_n - \mu_a)^2}{a_1 + a_2 + \dots + a_n}$$

- ▶ With these updated  $\mu_a$  and  $\sigma_a$ , go to step 1
- ▶ Repeat both the steps for let's say 100 iterations

Illustration plotted using an adapted code from Satyaki Bhattacharya : [here](#)

**Table 4.1:** Properties of the APDs at gain 50 and 18°C.

Sensitive area	$5 \times 5 \text{ mm}^2$
Operating voltage	340–430 V
Breakdown voltage - operating voltage	$45 \pm 5 \text{ V}$
Quantum efficiency (430 nm)	$75 \pm 2\%$
Capacitance	$80 \pm 2 \text{ pF}$
Excess noise factor	$2.1 \pm 0.2$
Effective thickness	$6 \pm 0.5 \mu\text{m}$
Series resistance	$< 10 \Omega$
Voltage sensitivity of the gain ( $1/M \cdot dM/dV$ )	$3.1 \pm 0.1\%/\text{V}$
Temperature sensitivity of the gain ( $1/M \cdot dM/dT$ )	$-2.4 \pm 0.2\%/\text{°C}$
Rise time	$< 2 \text{ ns}$
Dark current	$< 50 \text{ nA}$
Typical dark current	3 nA
Dark current after $2 \times 10^{13} \text{ n/cm}^2$	5 $\mu\text{A}$

# Z-boson

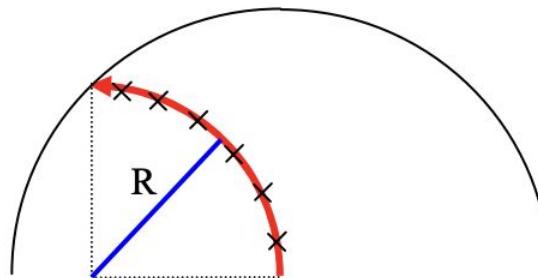
The Z boson is a subatomic particle with a mass of 91.19 (GeV) and a width of 2.5 GeV. It's an electrically neutral vector boson with a spin of one. The Z boson is very short-lived, with a half-life of about  $3 \times 10^{-25}$  seconds.

## 4.2.2 Momentum Measurement

The basic idea is that a charged particle bends in a magnetic field, and when it traverses the silicon sensors it creates an electrical signal that can be detected. Dividing the sensors into strips or pixels allows an estimation of the incidence position of the charged particles. Combining the information from many layers enables a “track” to be reconstructed. Once the track path is reconstructed, measuring the radius of curvature of the track gives an estimation of the particle momentum, according to

$$p = R \times 0.3B$$

where “p” is the particle momentum in GeV/c, R is the radius of curvature in metres and B is the magnetic field strength in Tesla.



*Schematic showing how the particle momentum is measured*

Radius of Curvature

$$R = \frac{P_t}{0.3 \times B} = \frac{P_t}{0.3 \times 3.8} = \frac{P_t}{1.14}$$

$P_t \rightarrow$  in GeV       $R \rightarrow$  meters.

if  $P_t = 10$  GeV

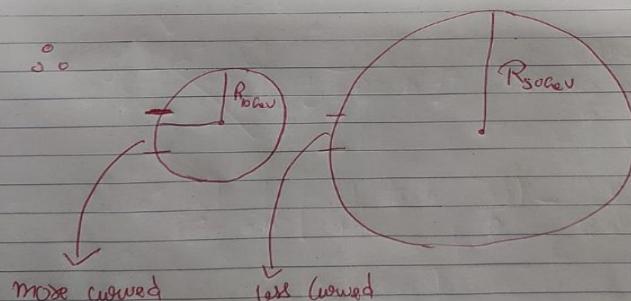
$$\text{Then } R = \frac{10}{1.14} \text{ m}$$

& If  $P_t = 50$  GeV

$$R = \frac{50}{1.14} \text{ m}$$

So Clearly  $R_{50\text{GeV}} > R_{10\text{GeV}}$

We have tracker  $\approx 1$  meter



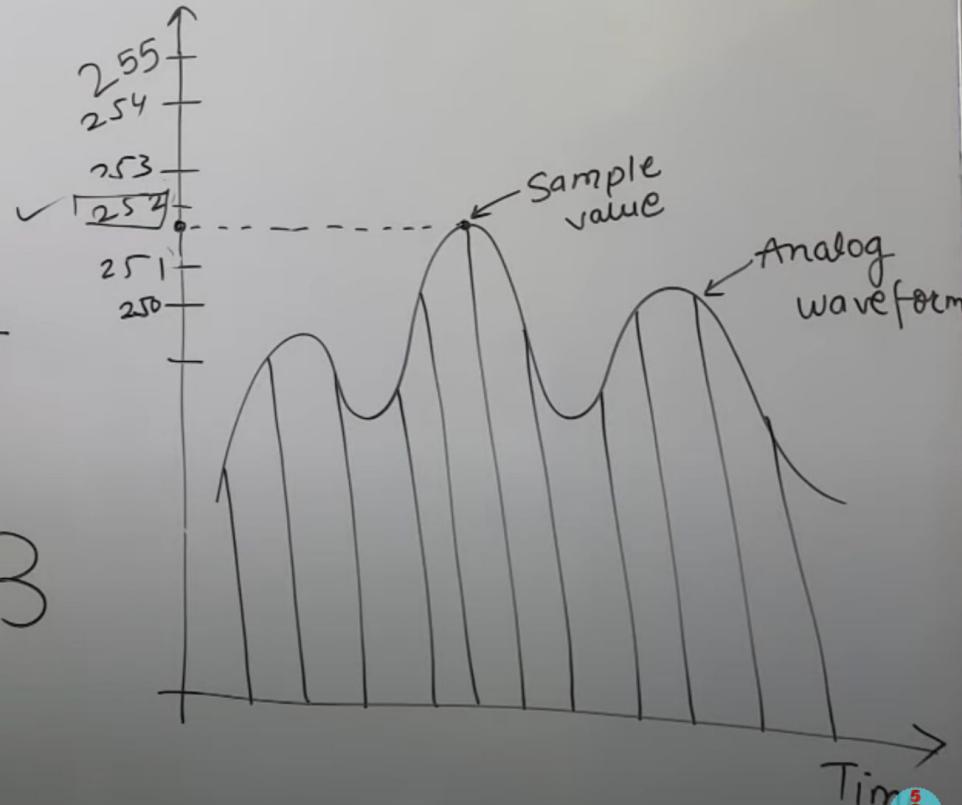
Hence for low  $p_t$   $e^-$  will curve more, so have more inefficiency in Algorithm.

# ADC

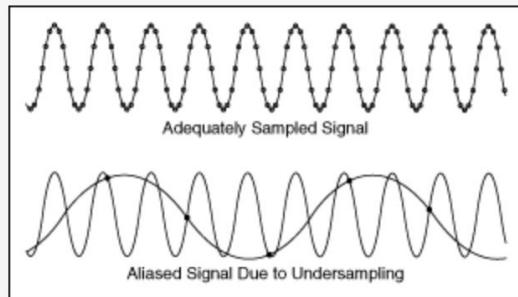
- Sampling
- quantization
- Nyquist Theorem
- Aliasing

256 — 8 bits

$$\begin{array}{ccc} 4 \text{ kHz} & & B \text{ Hz} \\ \downarrow & & \downarrow \\ 8000 & & 2B \end{array}$$



An aliased signal provides a poor representation of the analog signal. Aliasing causes a false lower frequency component to appear in the sampled data of a signal. The following figure shows an adequately sampled signal and an inadequately sampled signal.



In the previous figure, the inadequately sampled signal appears to have a lower frequency than the actual signal—two cycles instead of ten cycles.

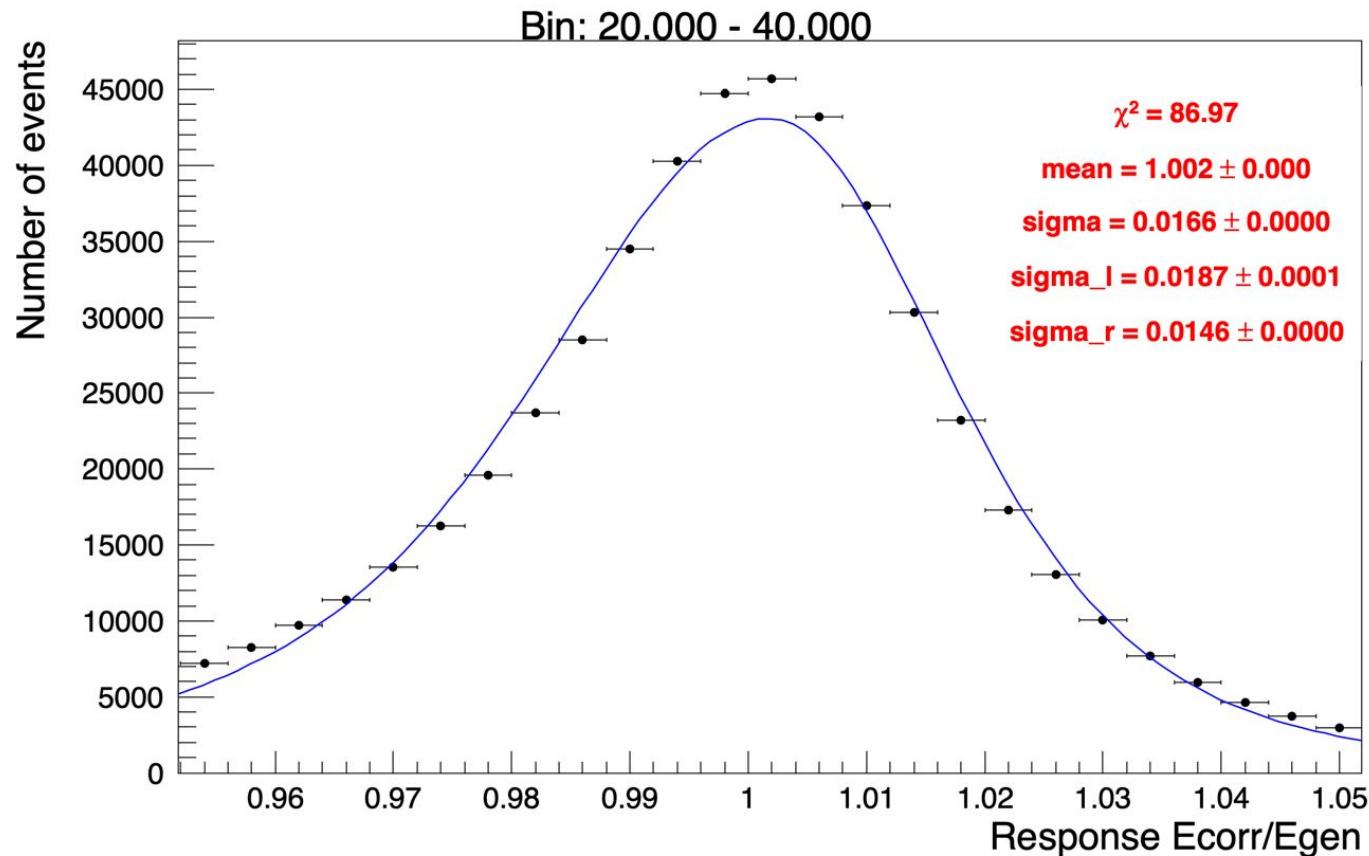
Increasing the sampling frequency increases the number of data points acquired in a given time period. Often, a fast sampling frequency provides a better representation of the original signal than a slower sampling frequency.

For a given sampling frequency, the maximum frequency you can accurately represent without aliasing is the Nyquist frequency. The Nyquist frequency equals one-half the sampling frequency, as shown by the following equation.

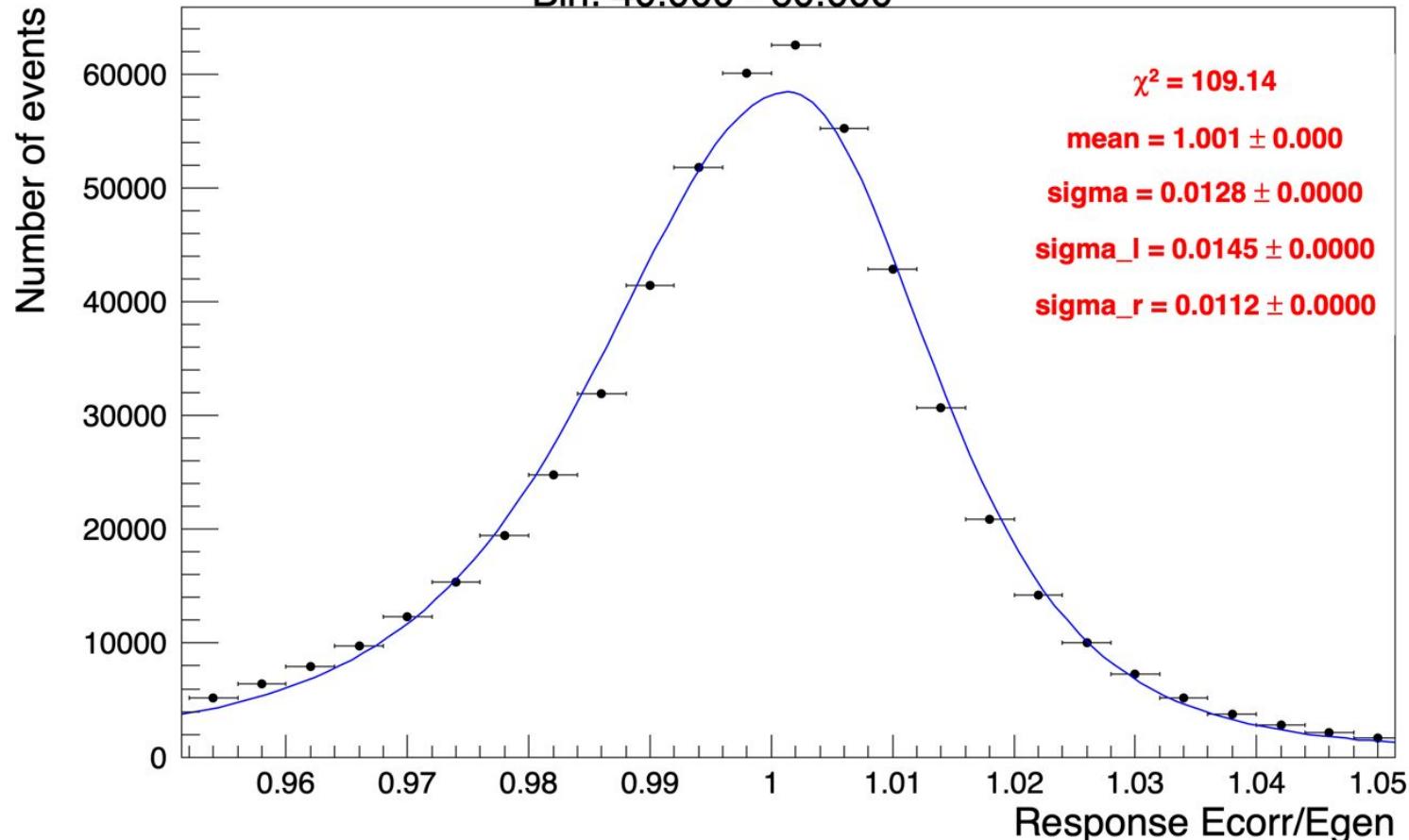
$$f_N = \frac{f_s}{2}$$

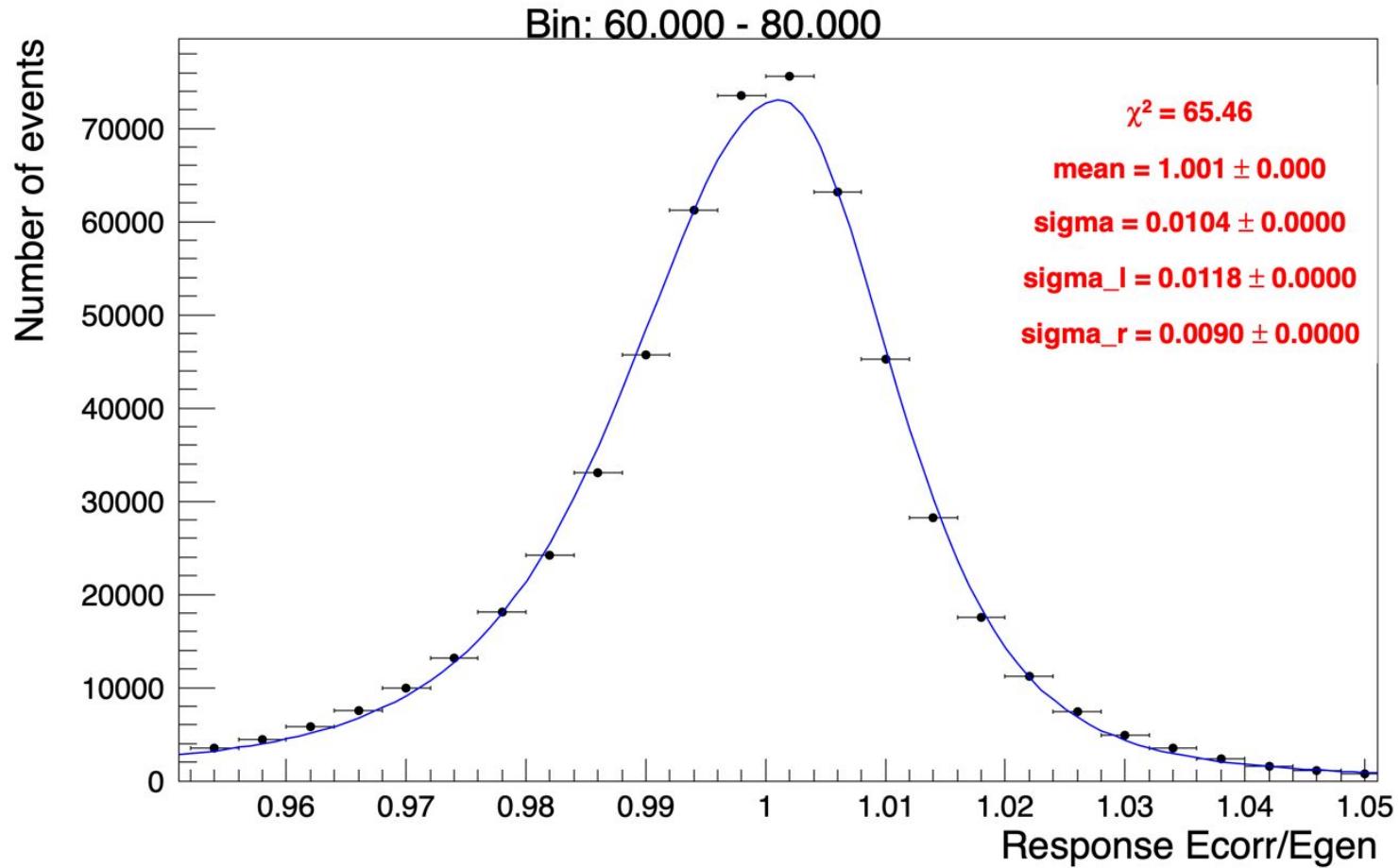
where  $f_N$  is the Nyquist frequency and  $f_s$  is the sampling frequency.

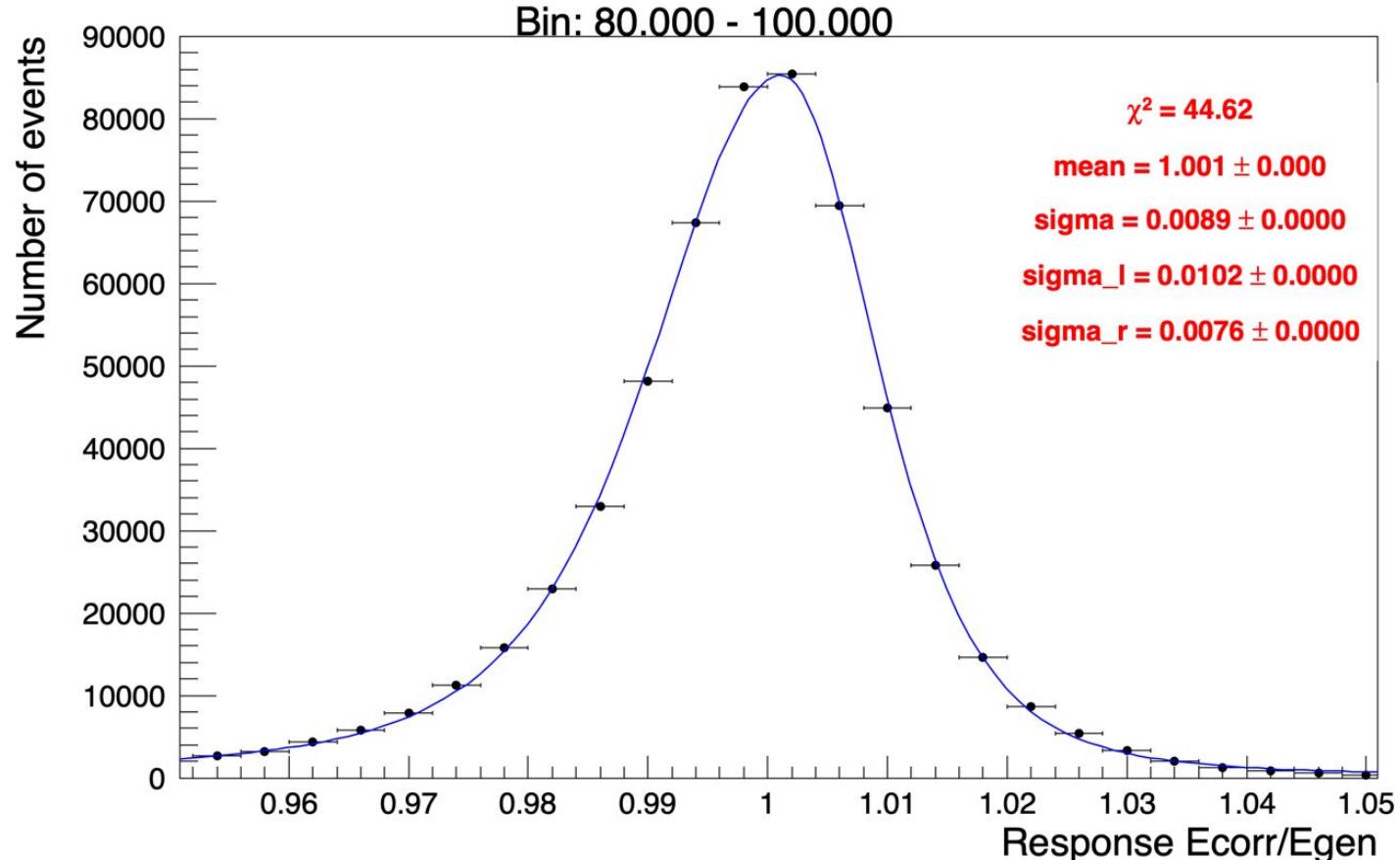
# Pt corrected

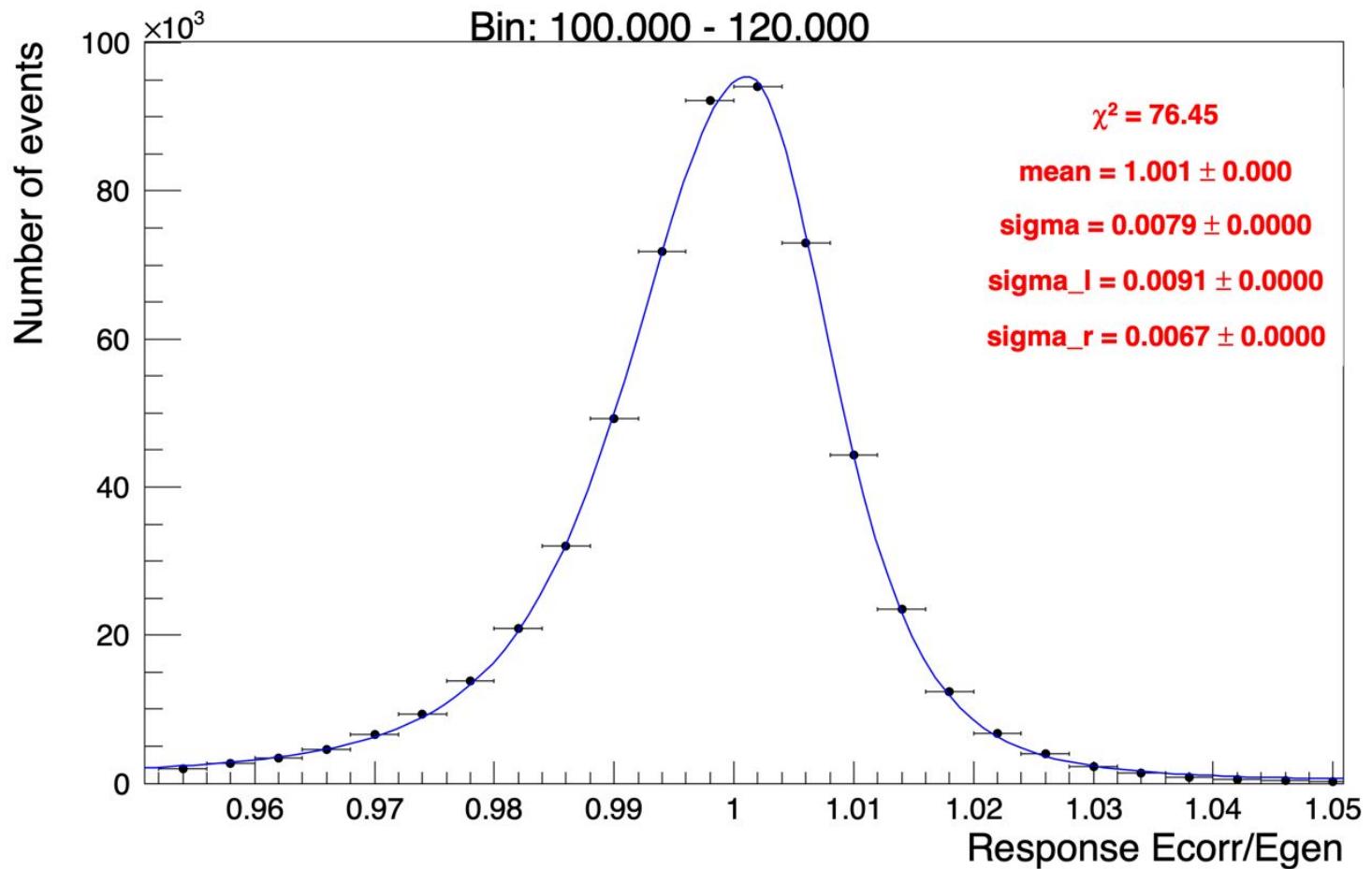


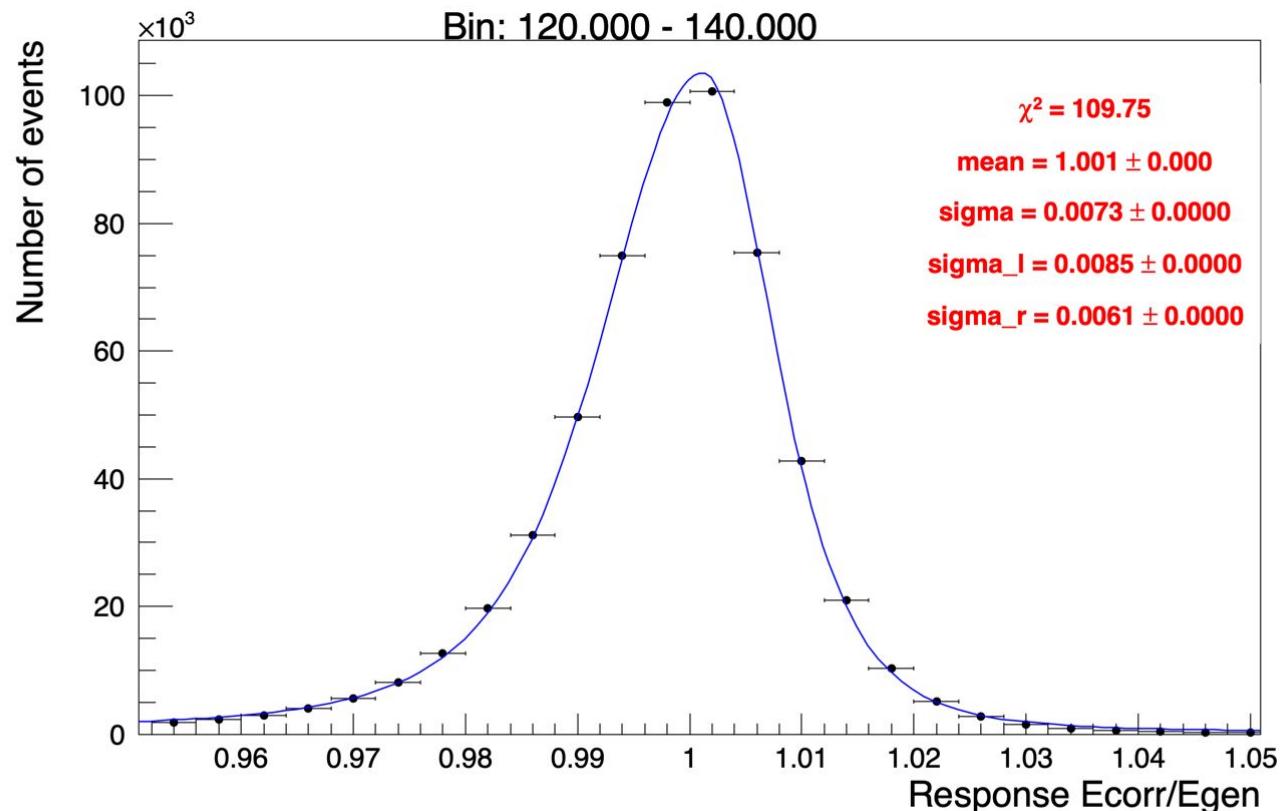
Bin: 40.000 - 60.000

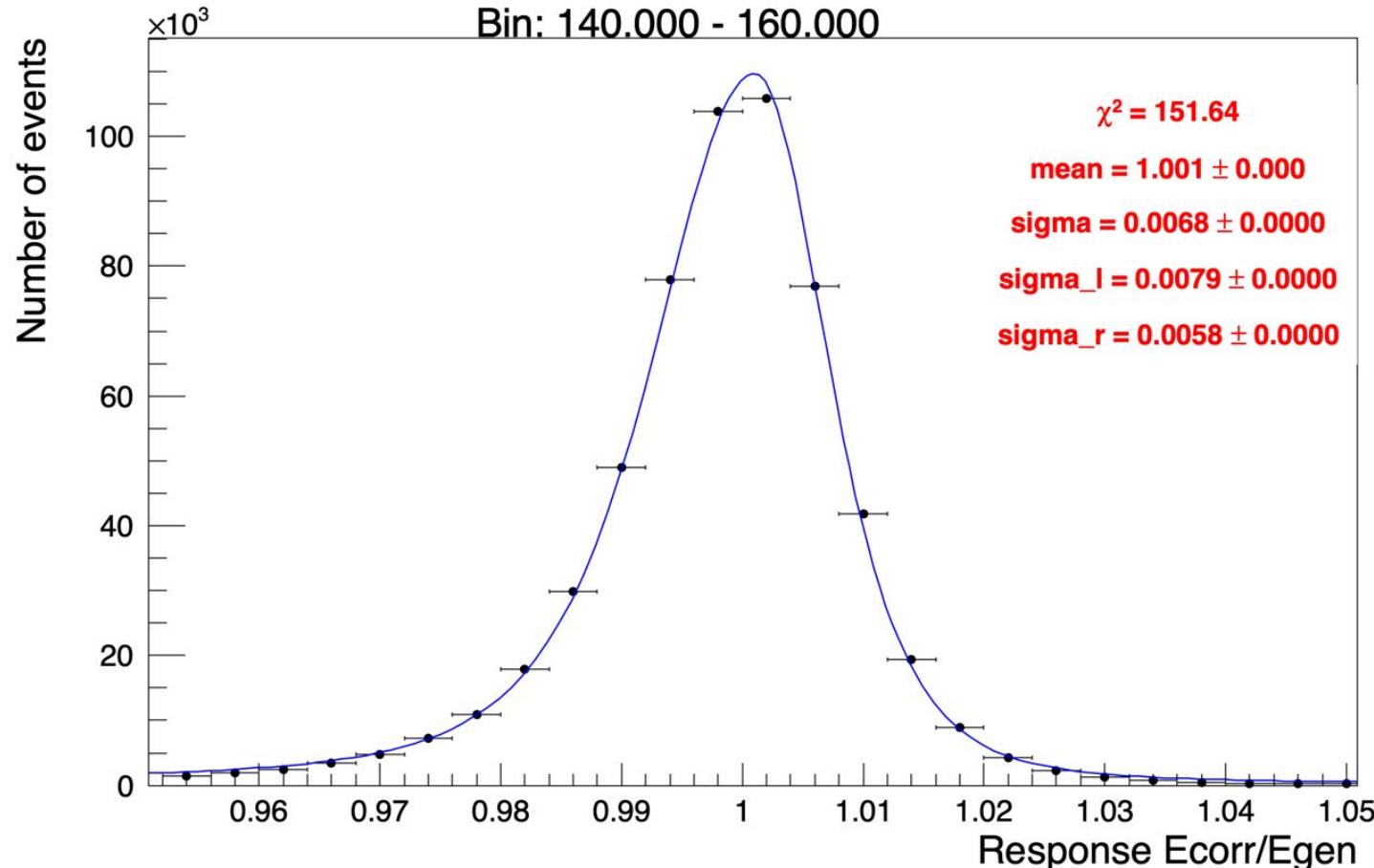


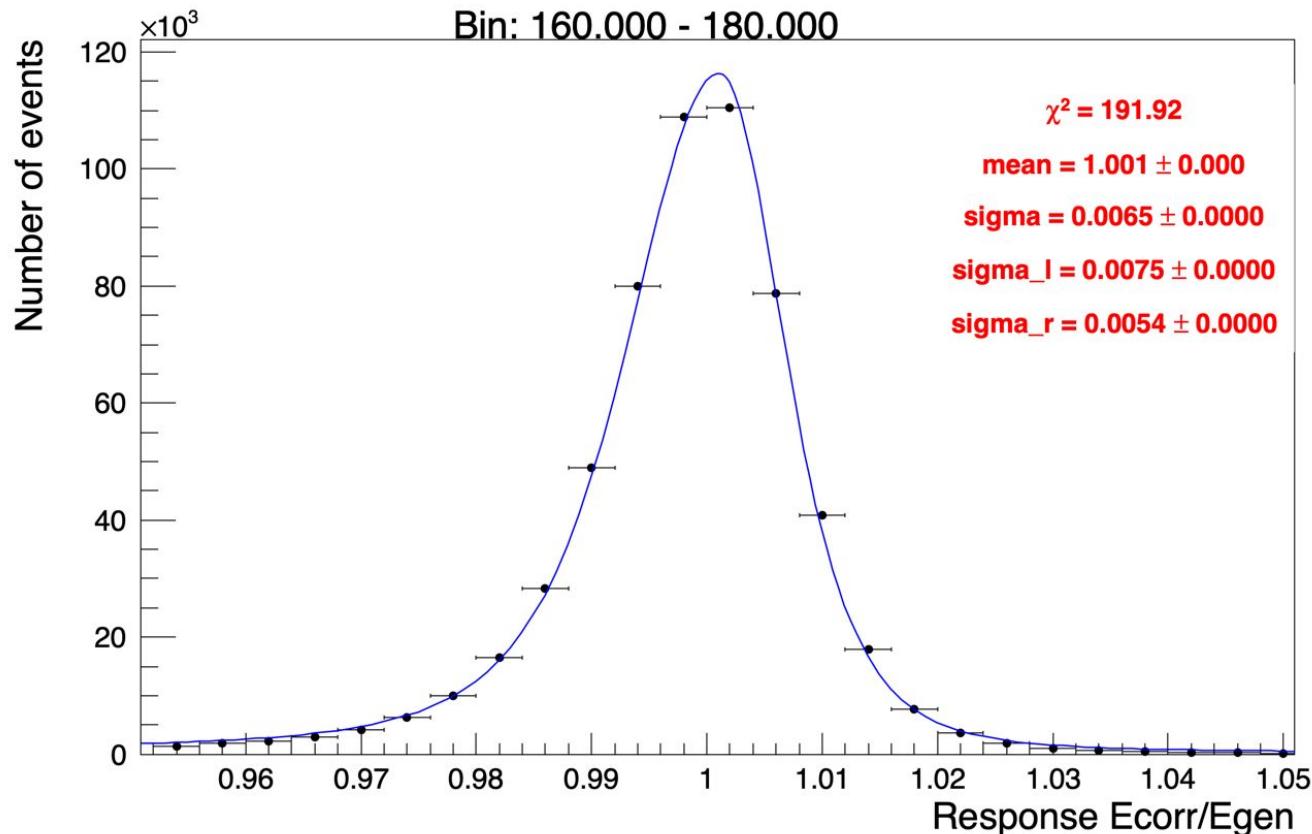


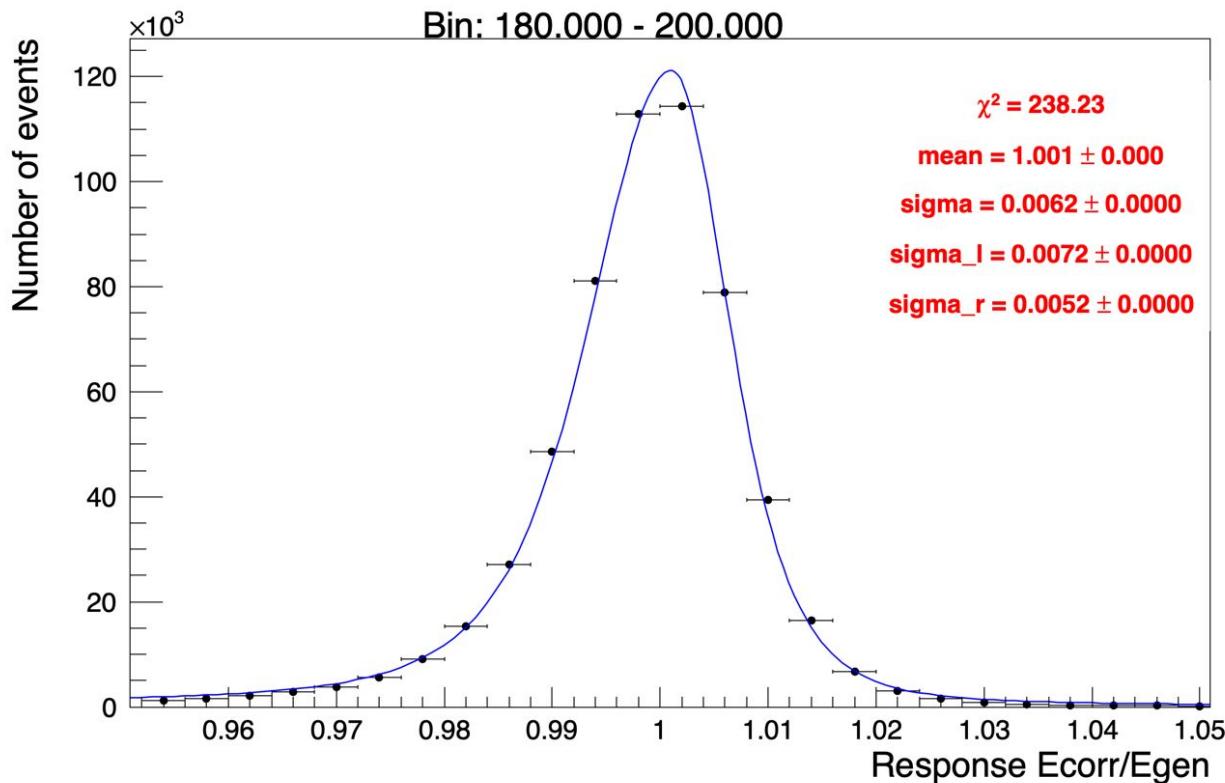


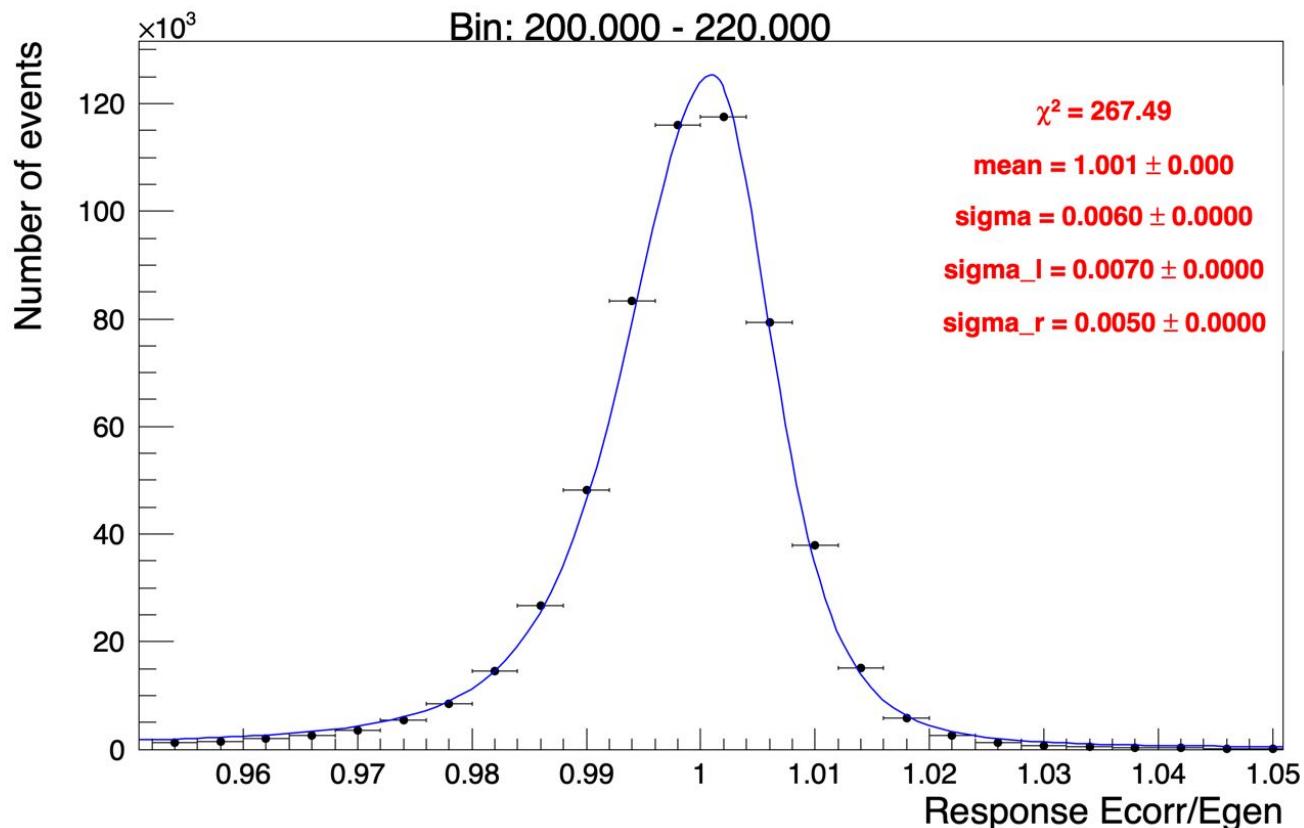


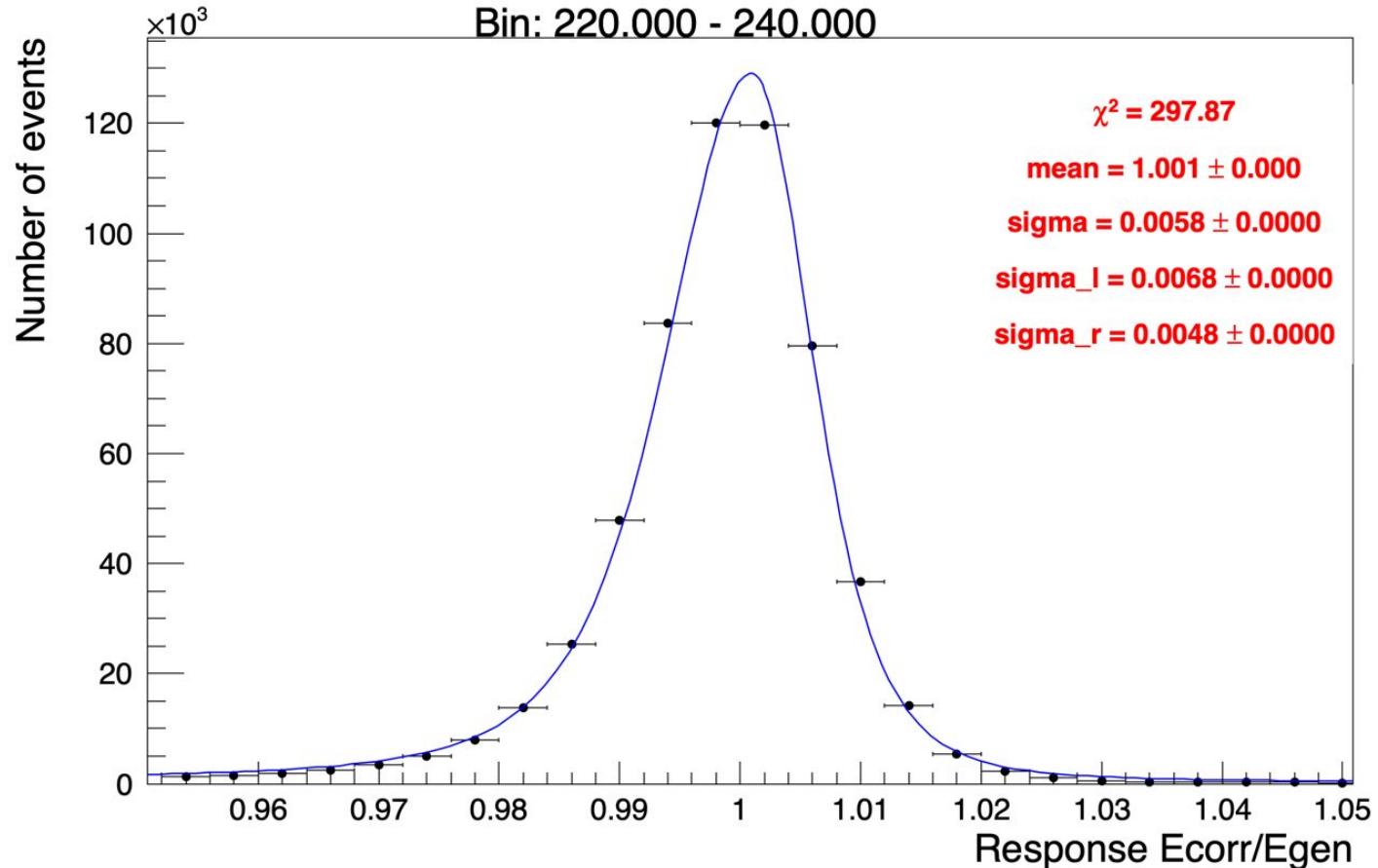


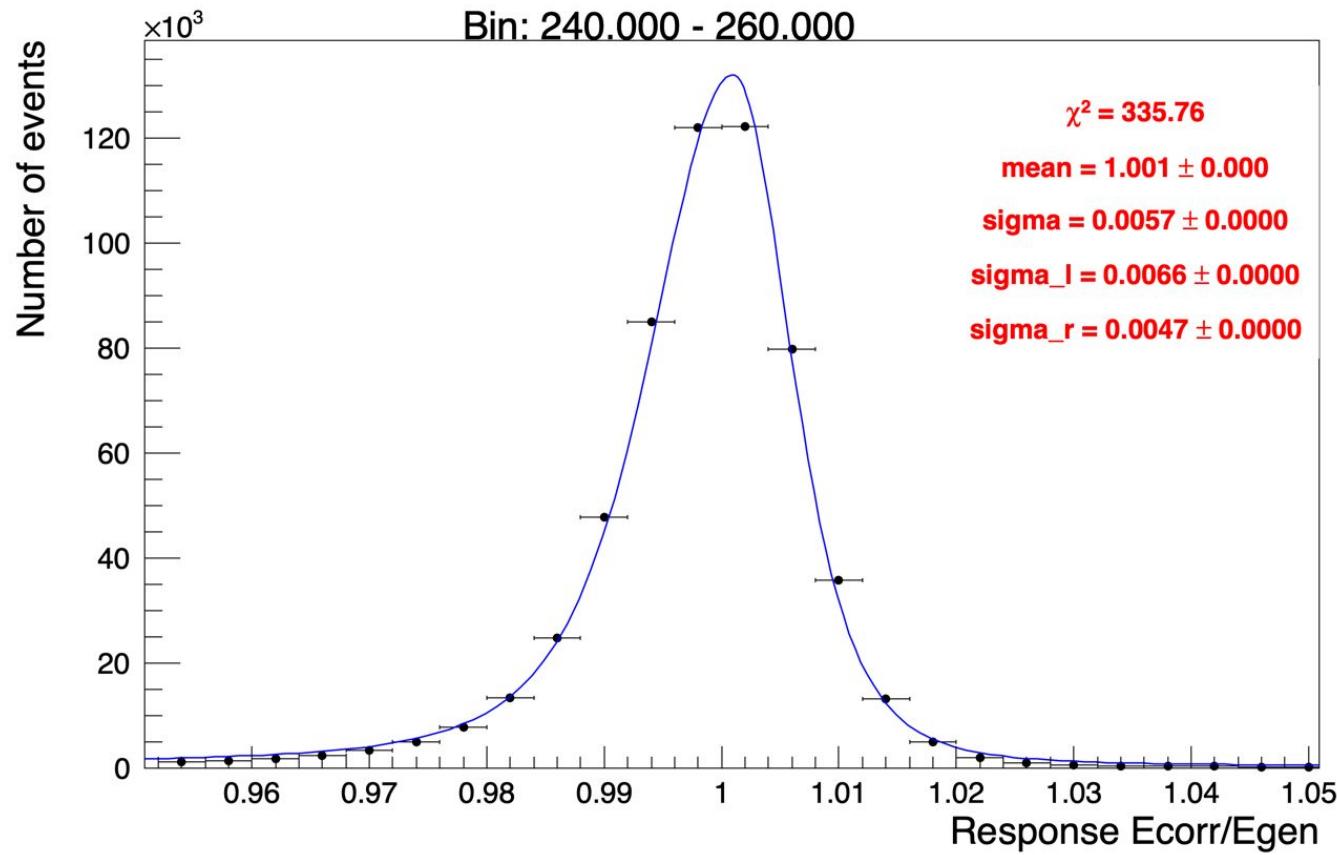


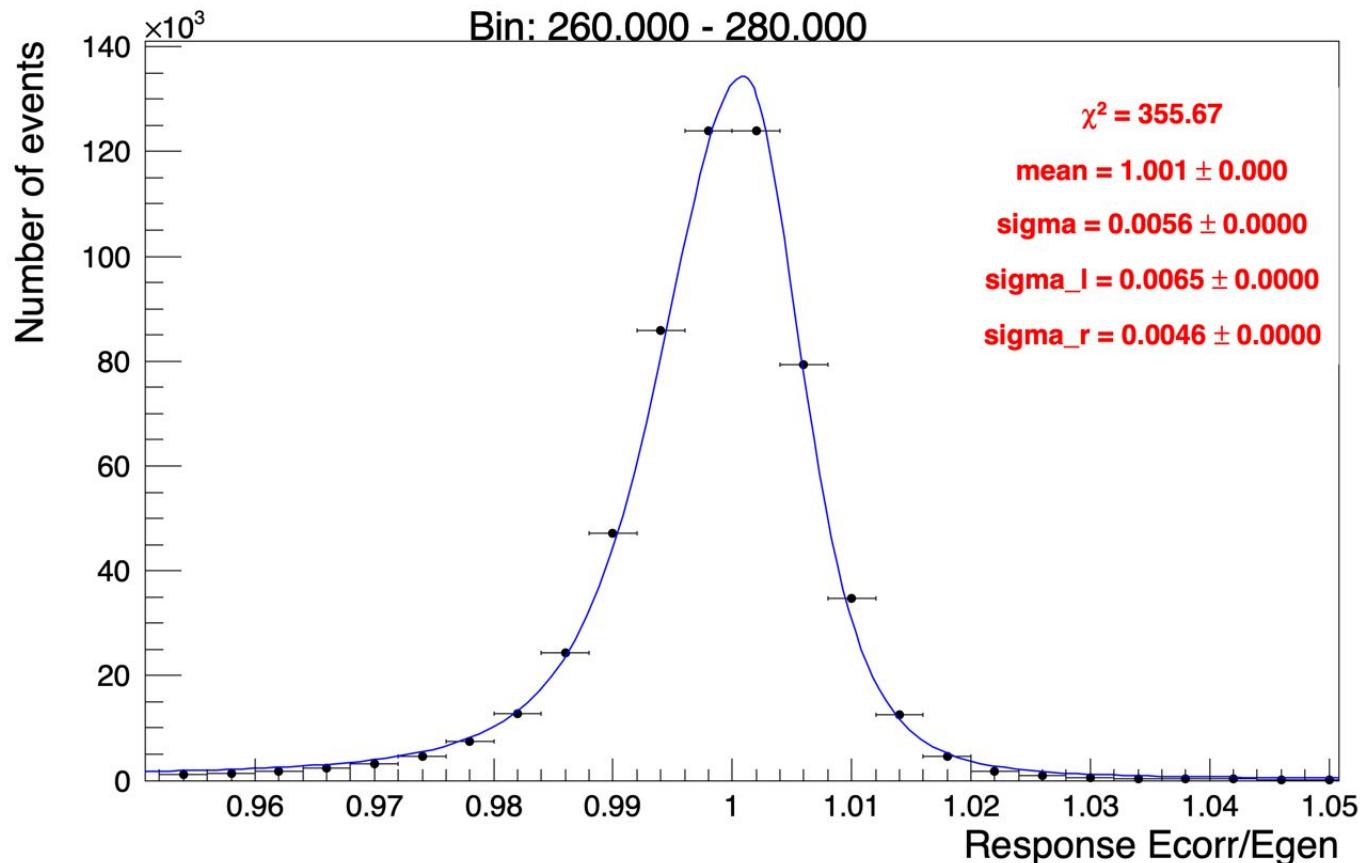


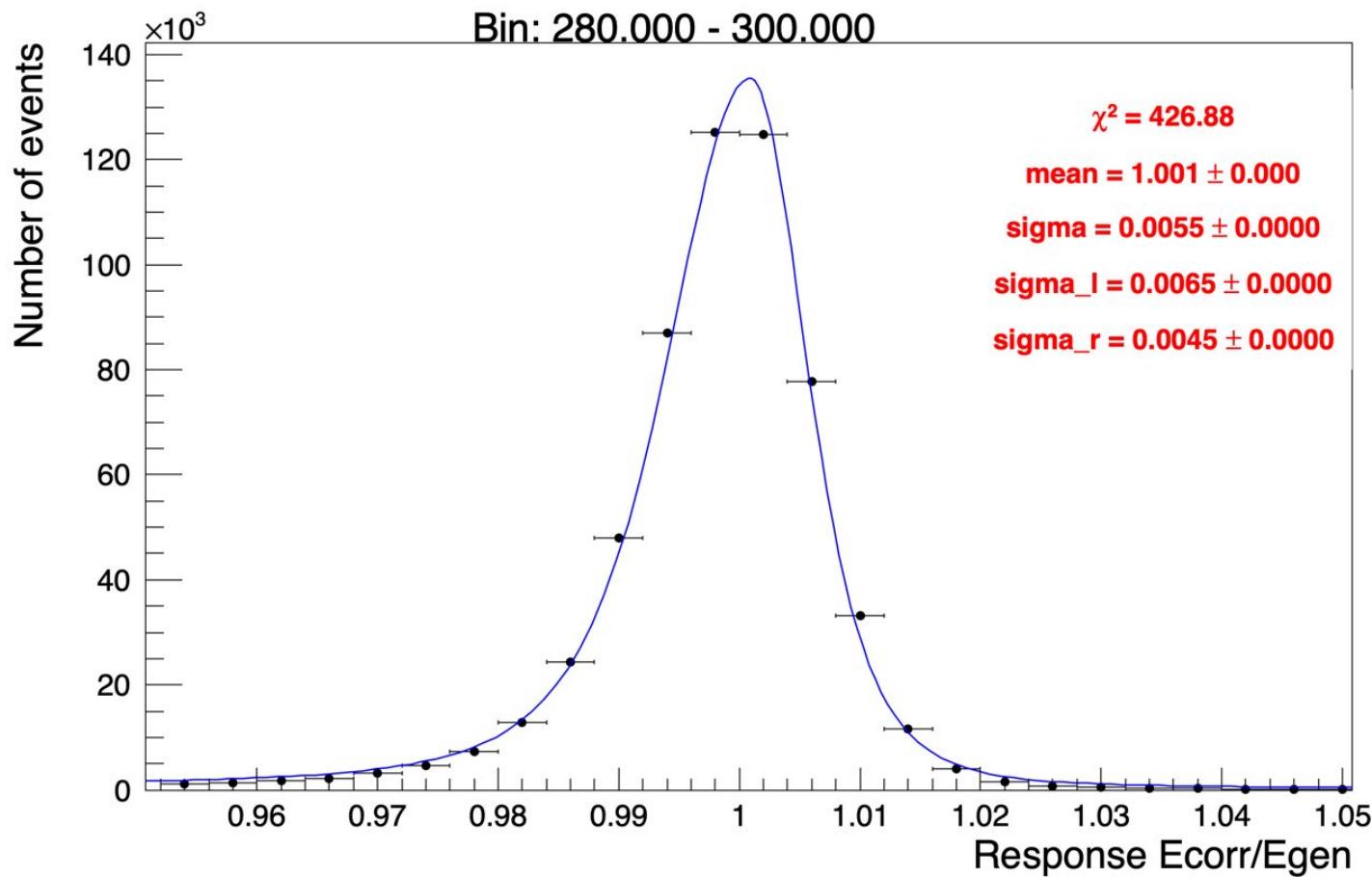




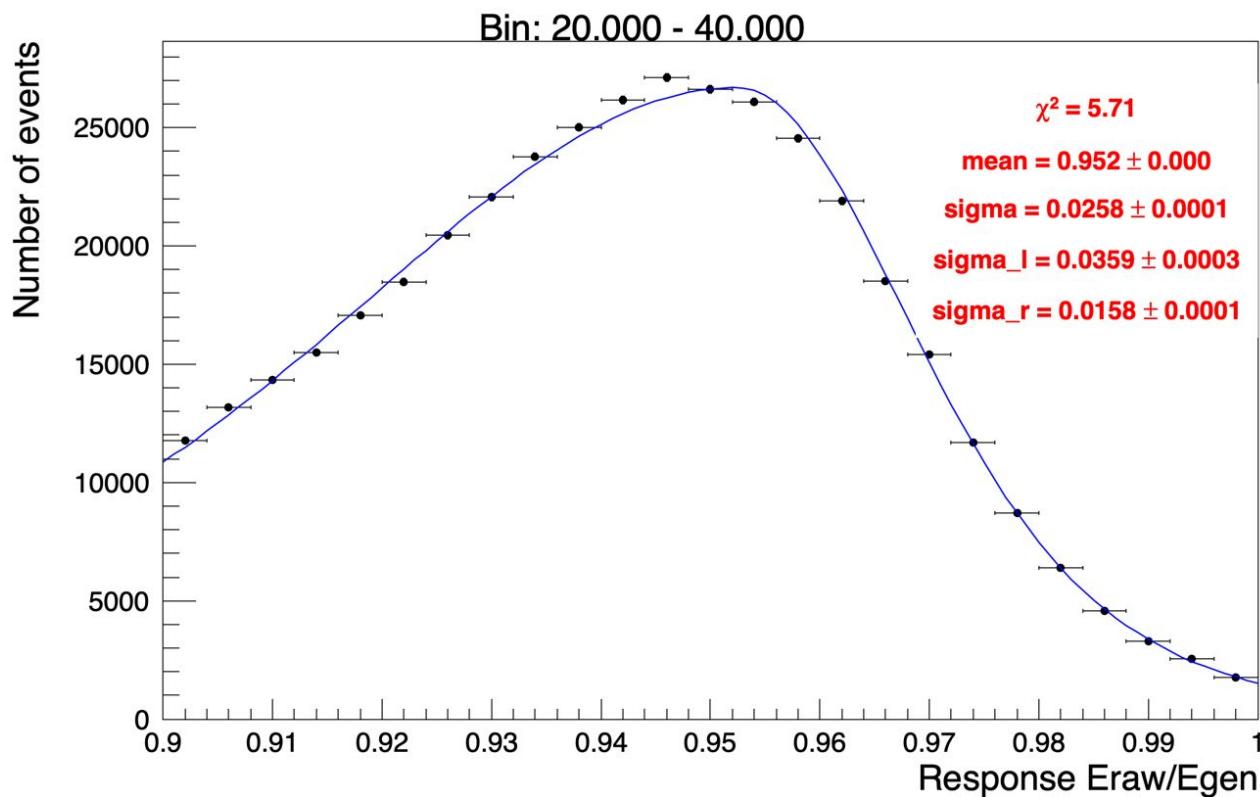


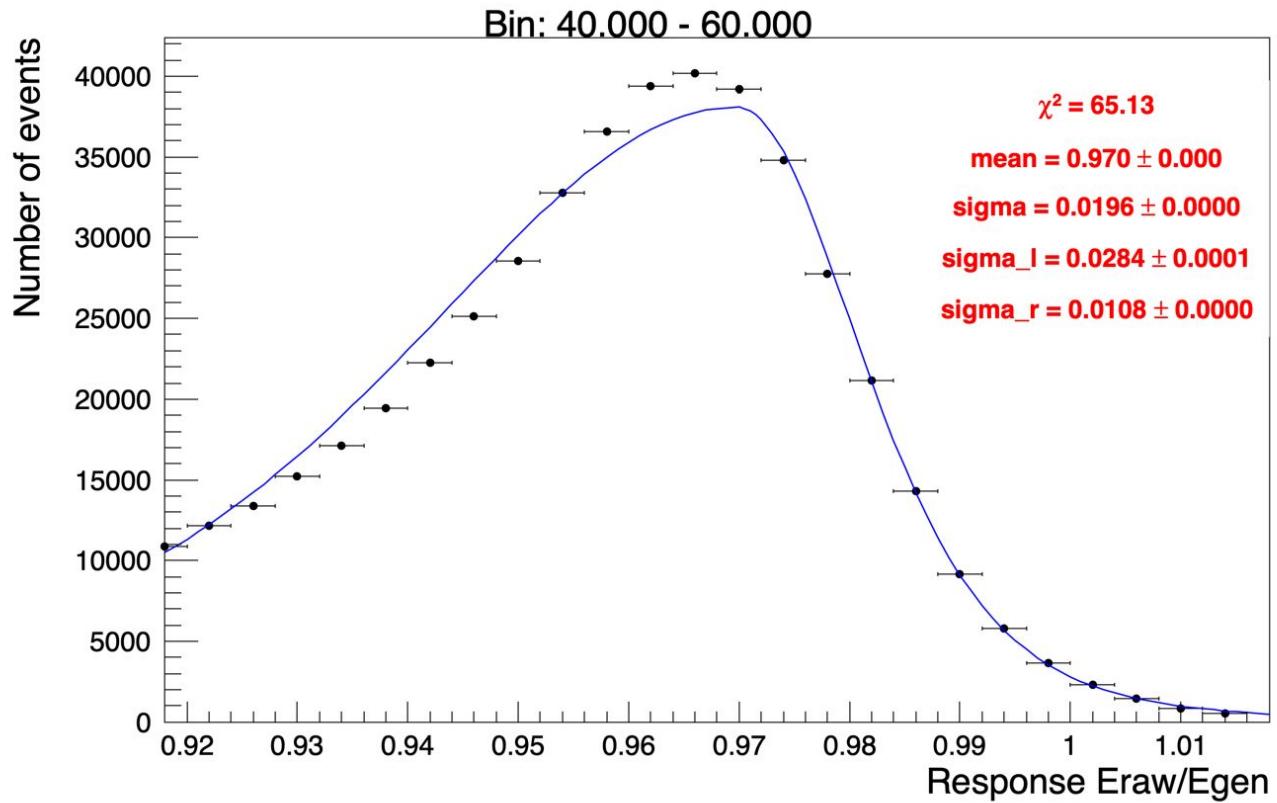


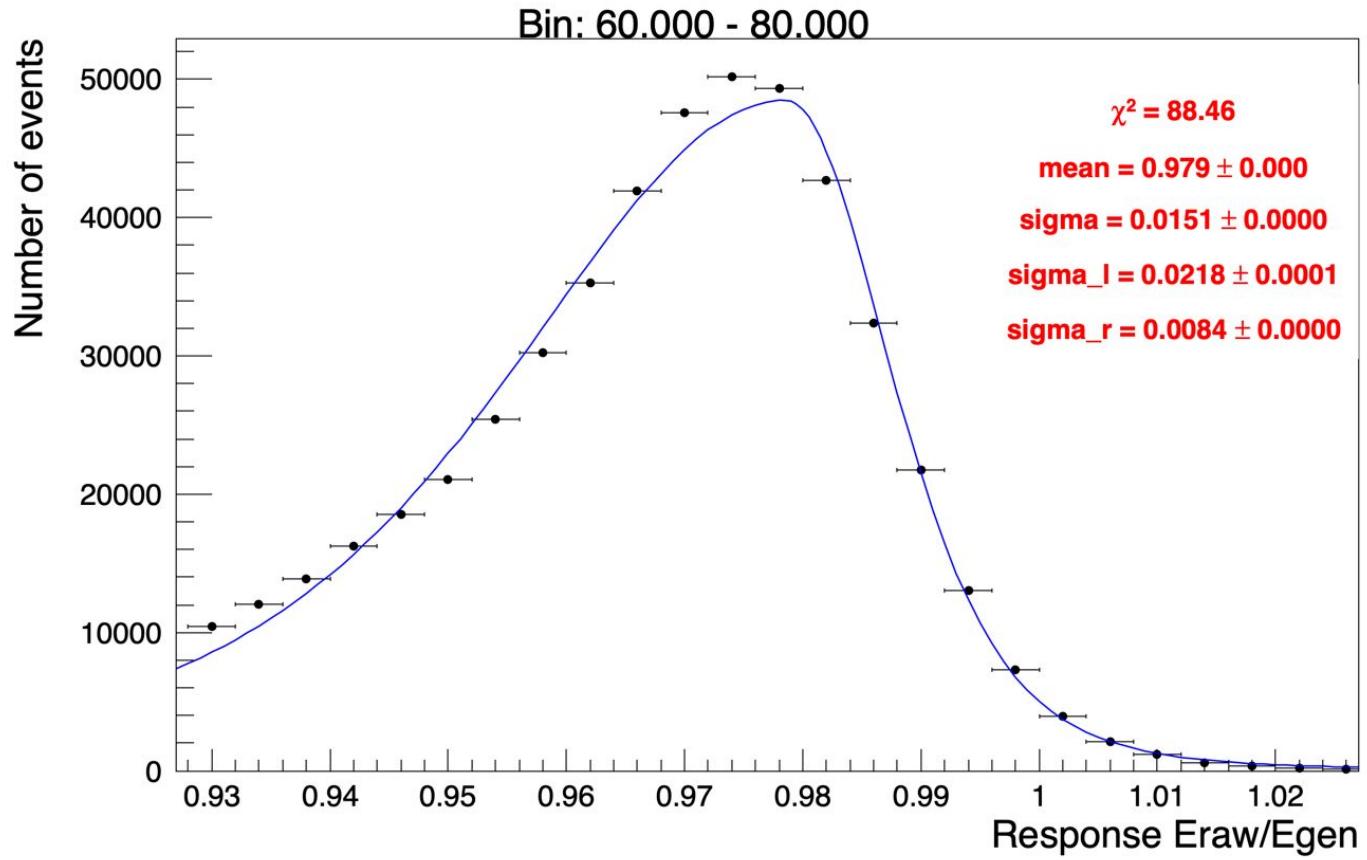


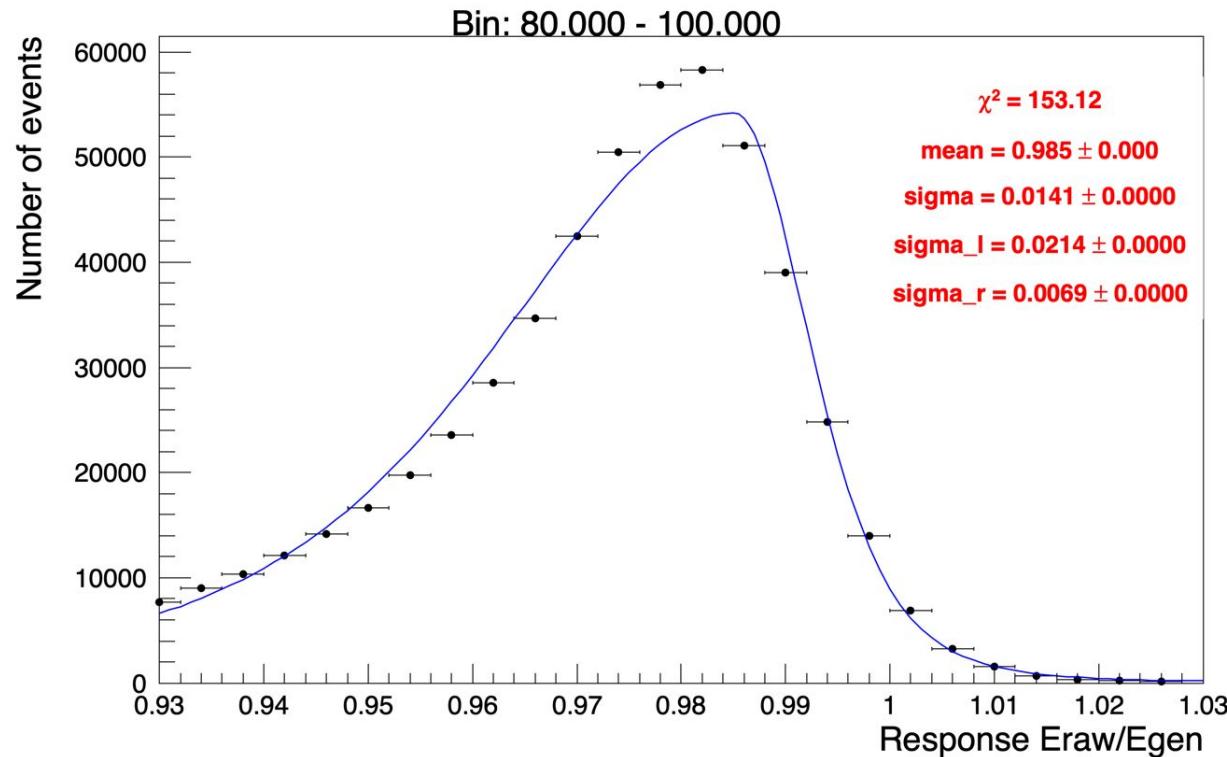


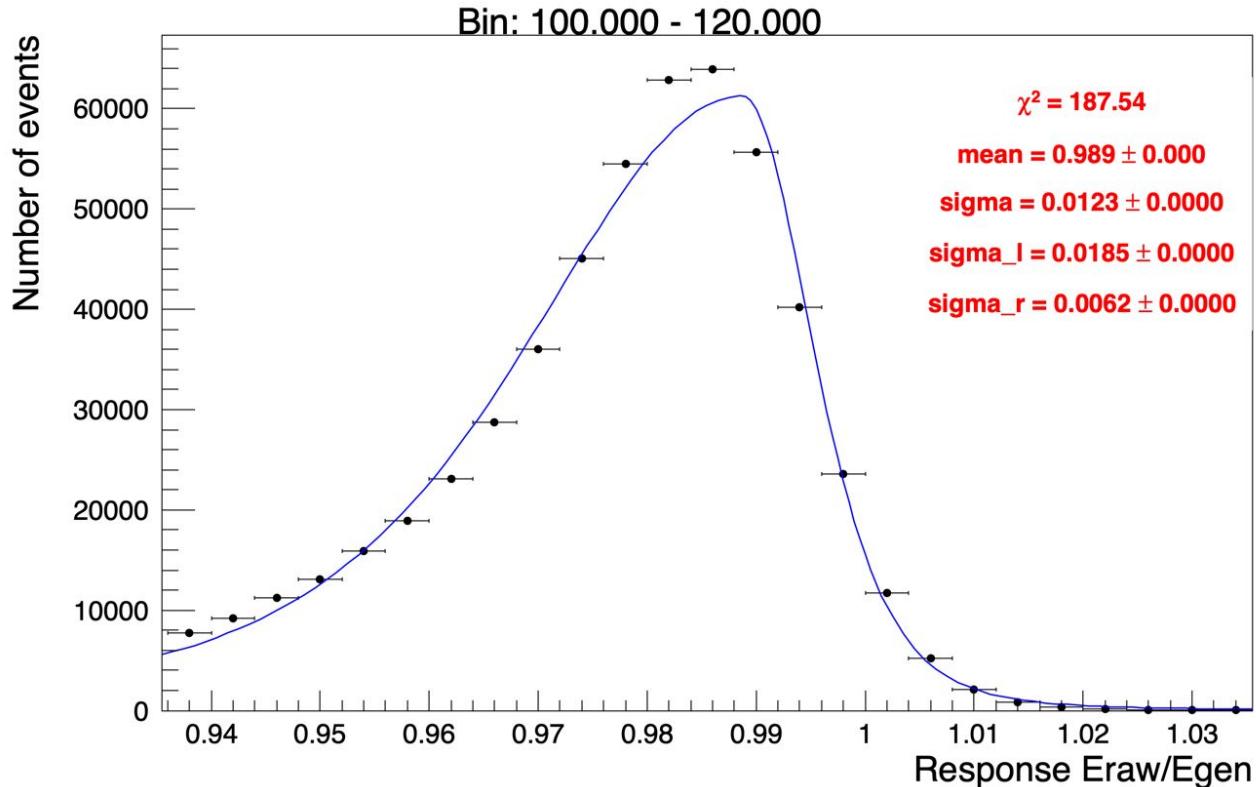
# Pt raw

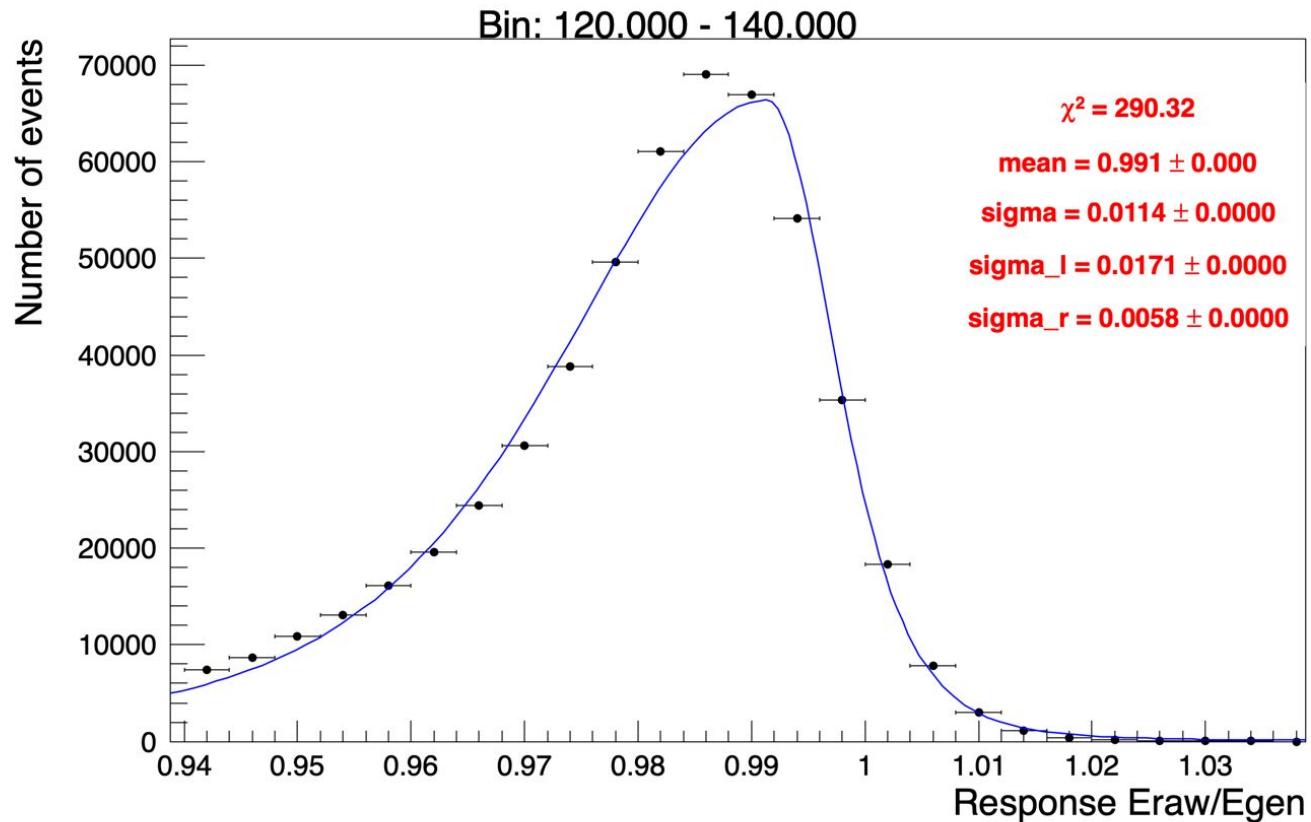


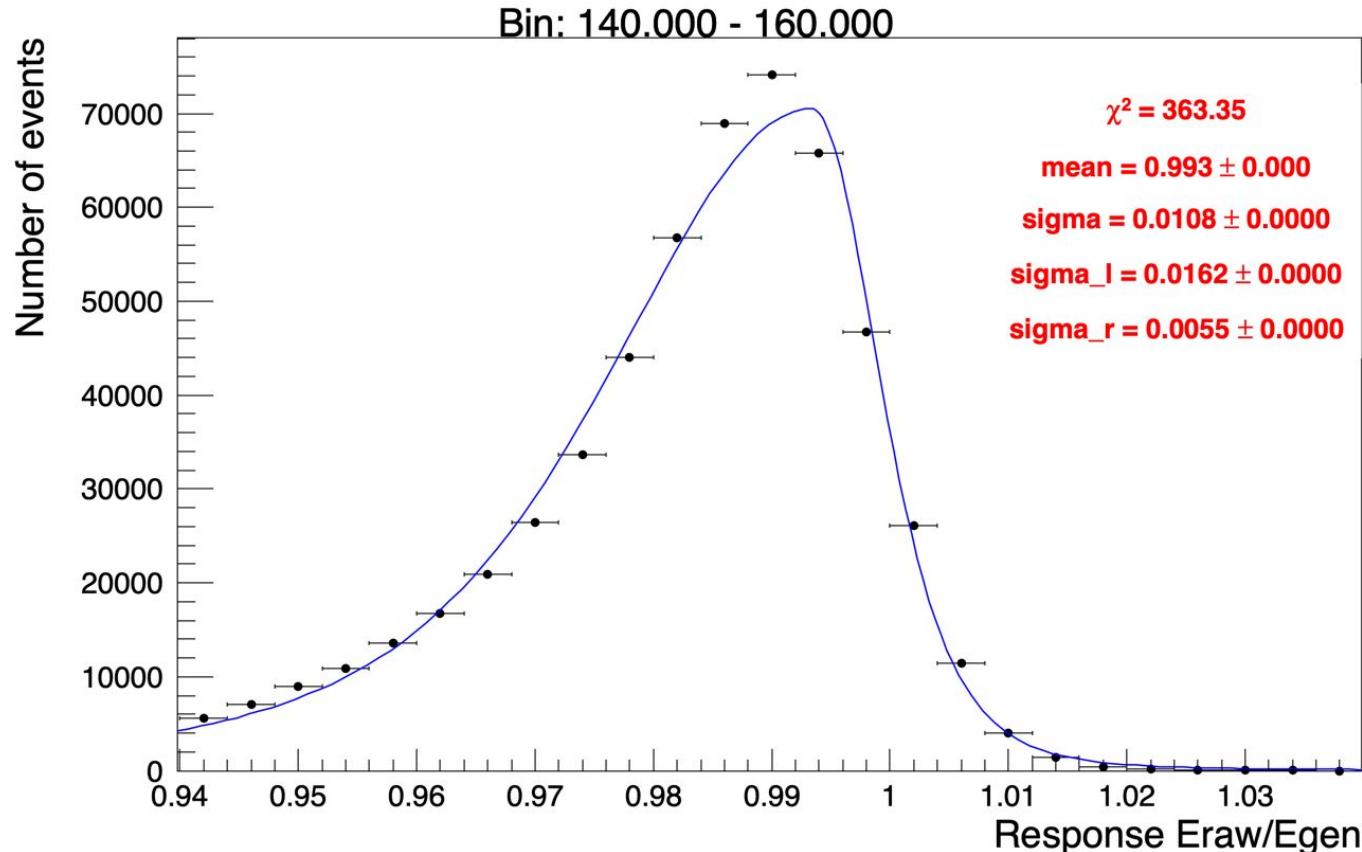


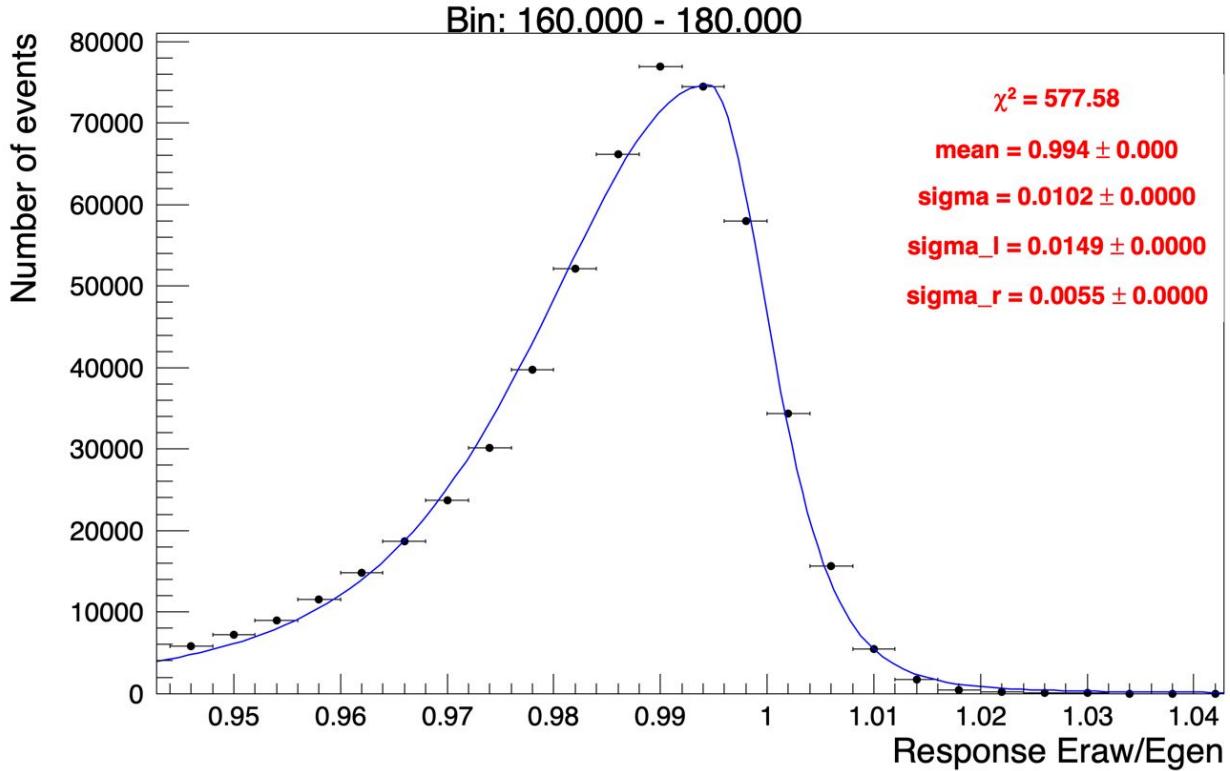


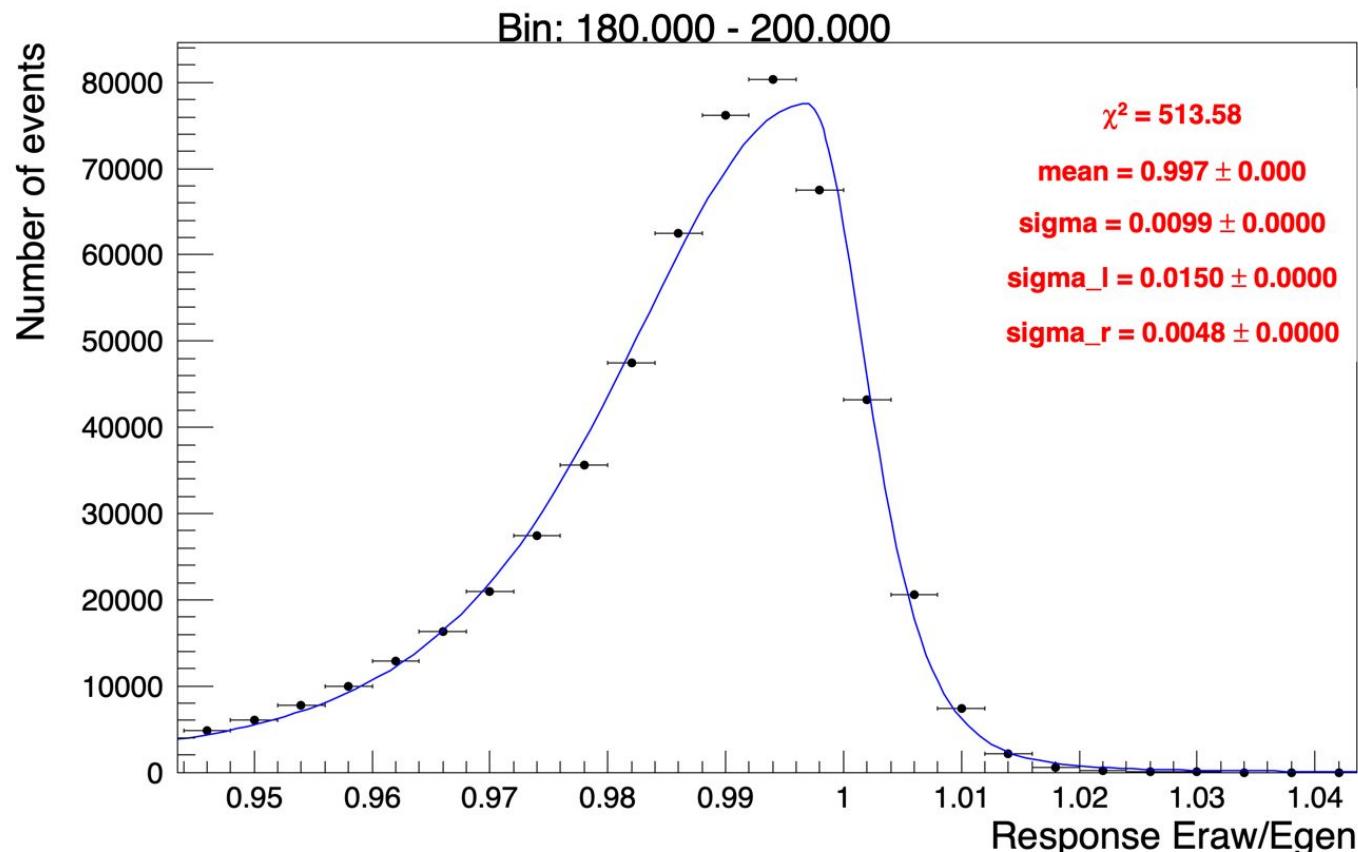


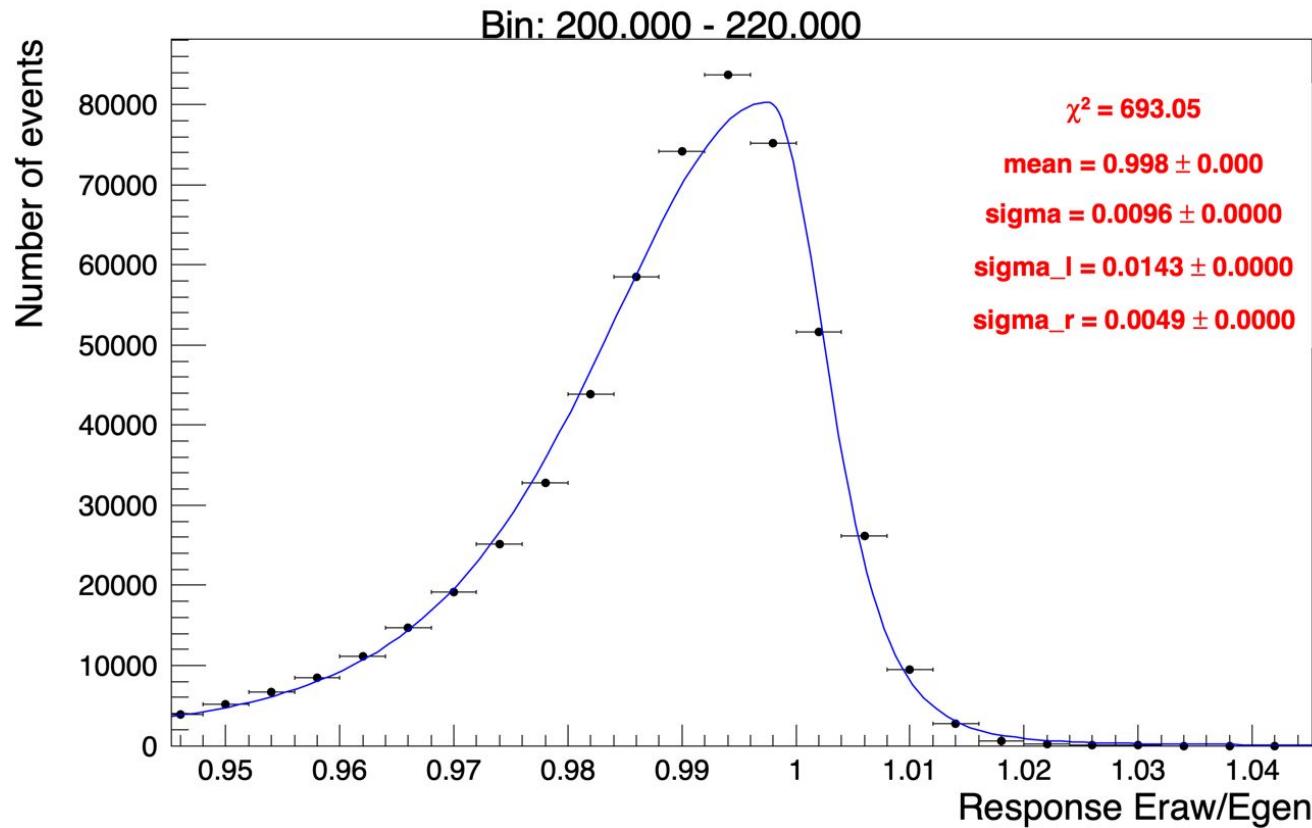


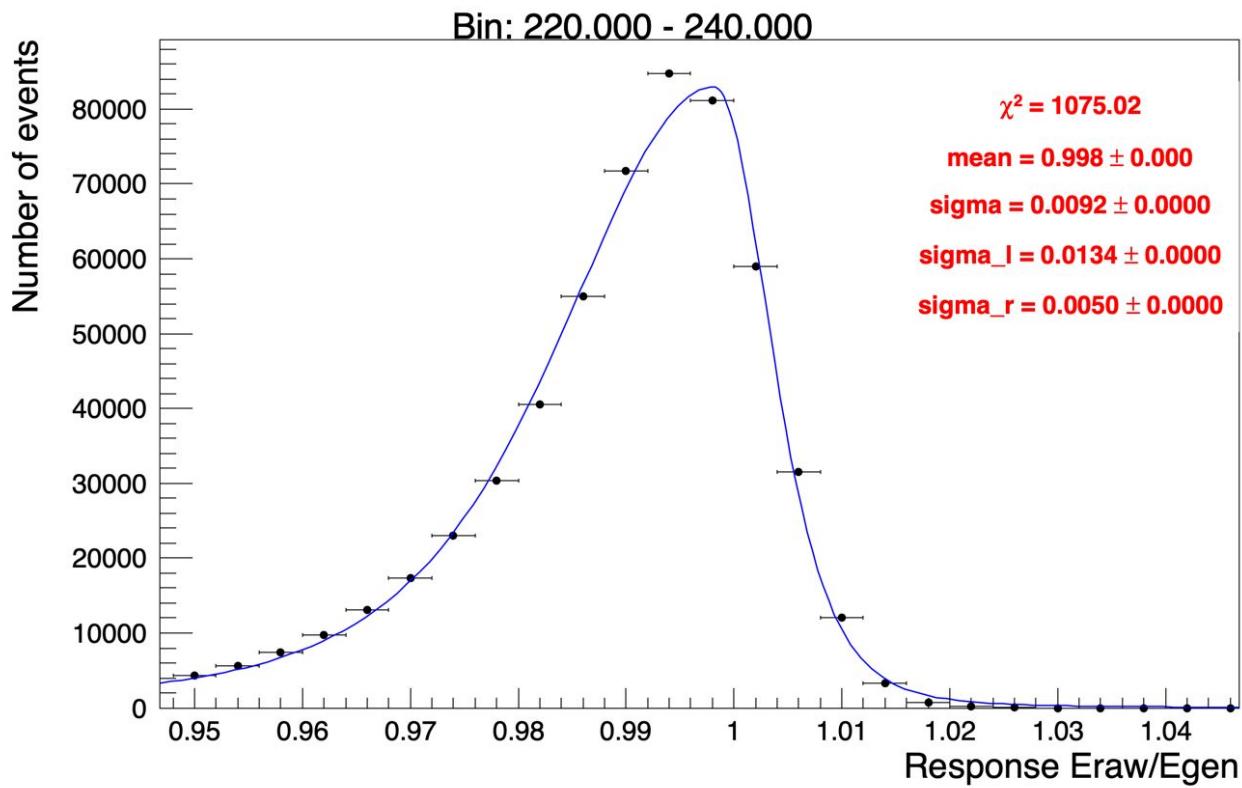




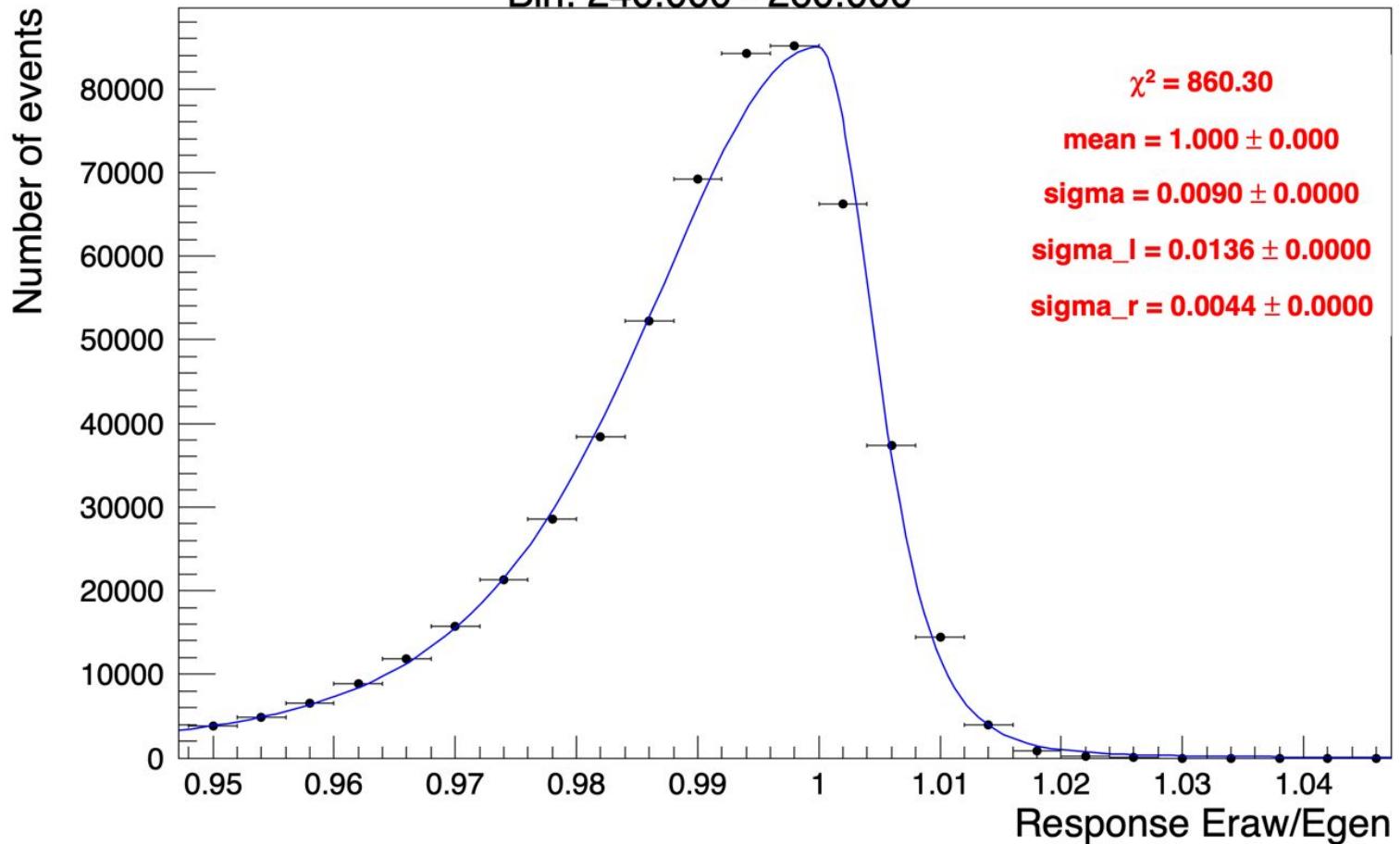




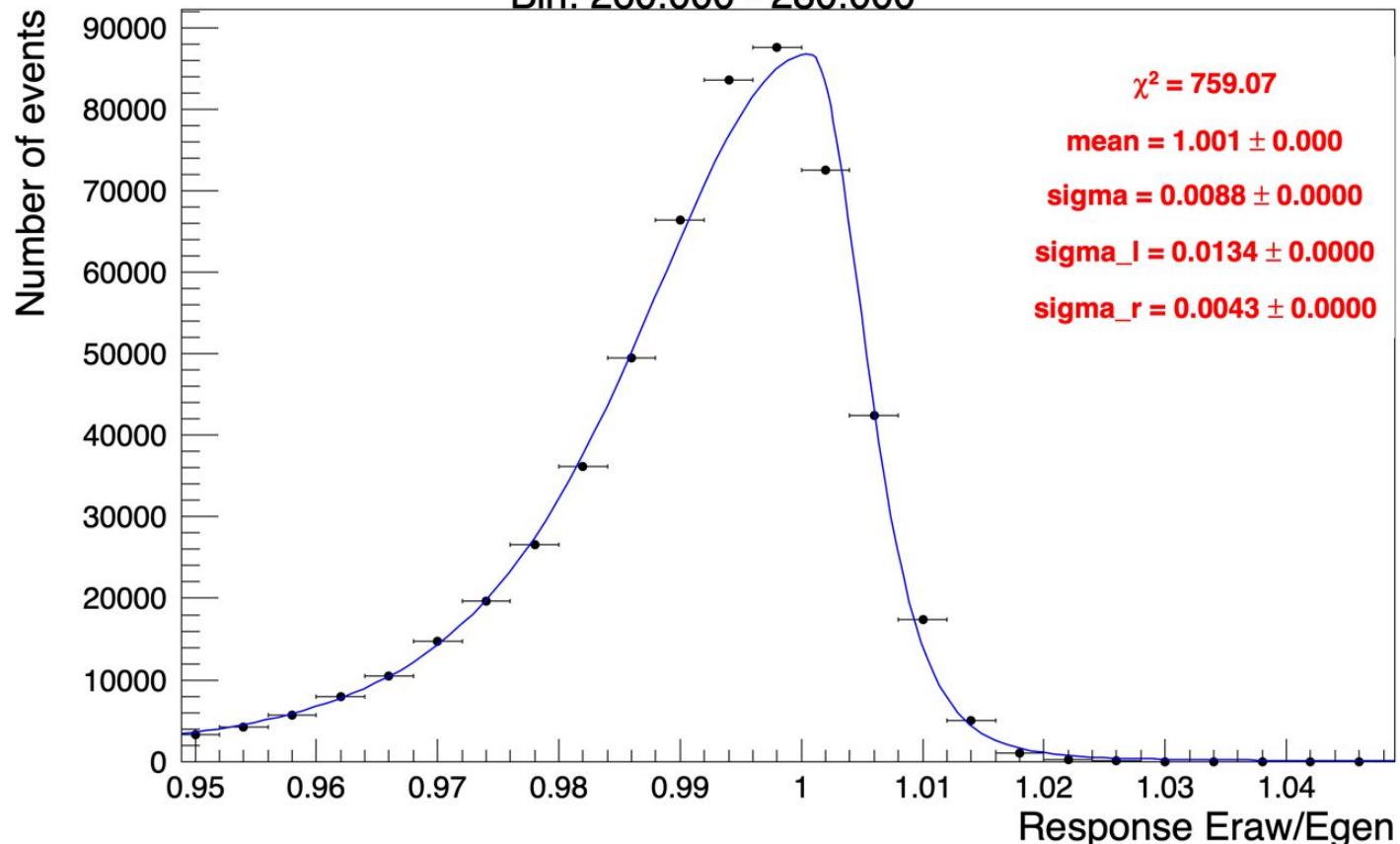




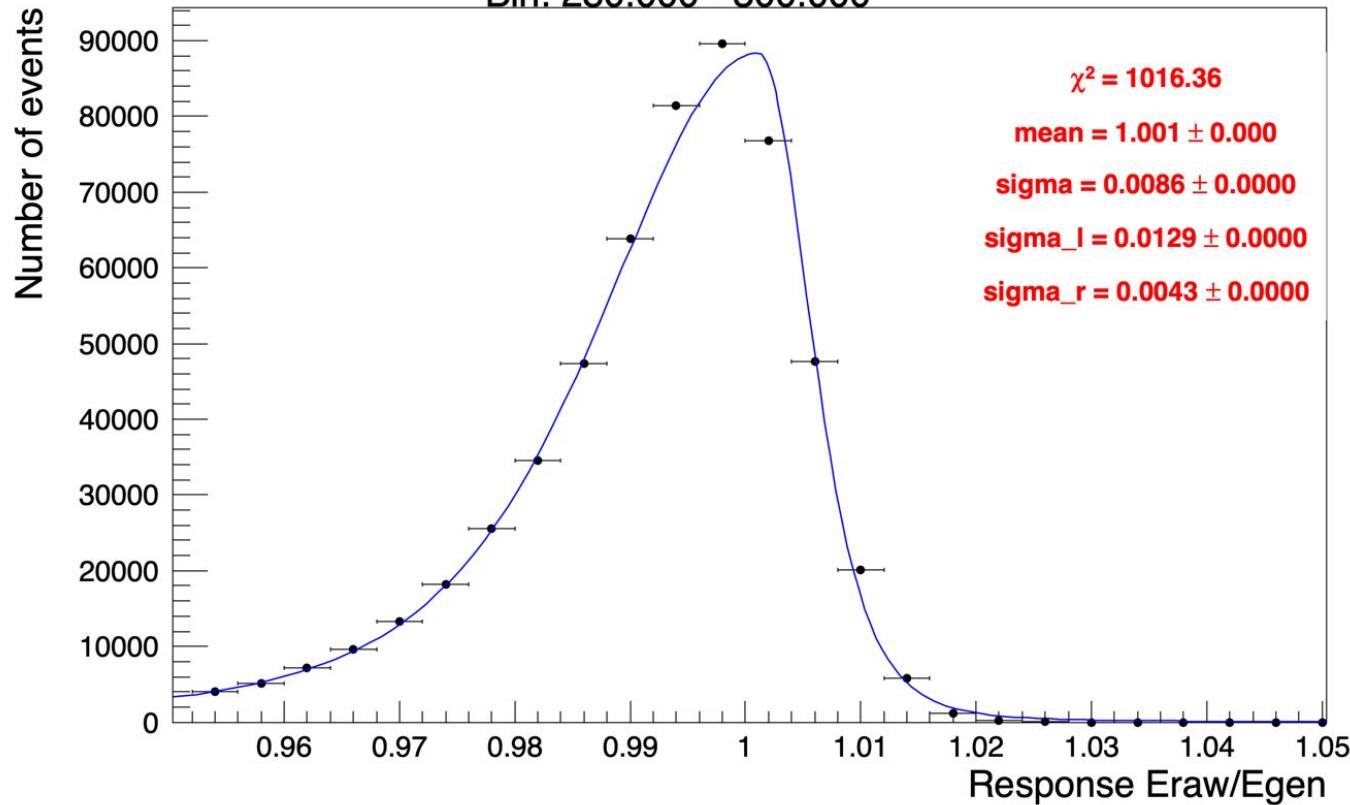
Bin: 240.000 - 260.000



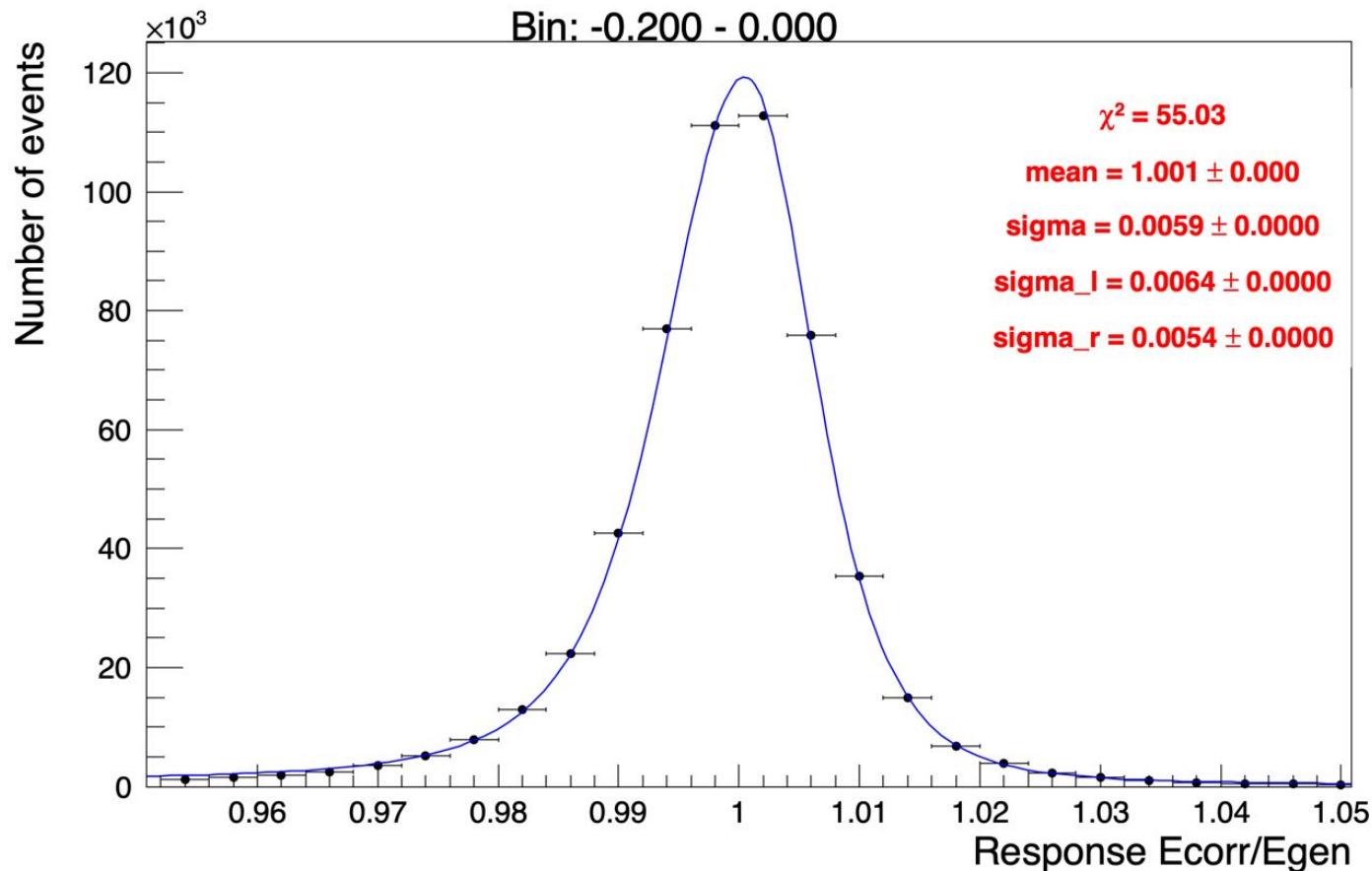
Bin: 260.000 - 280.000

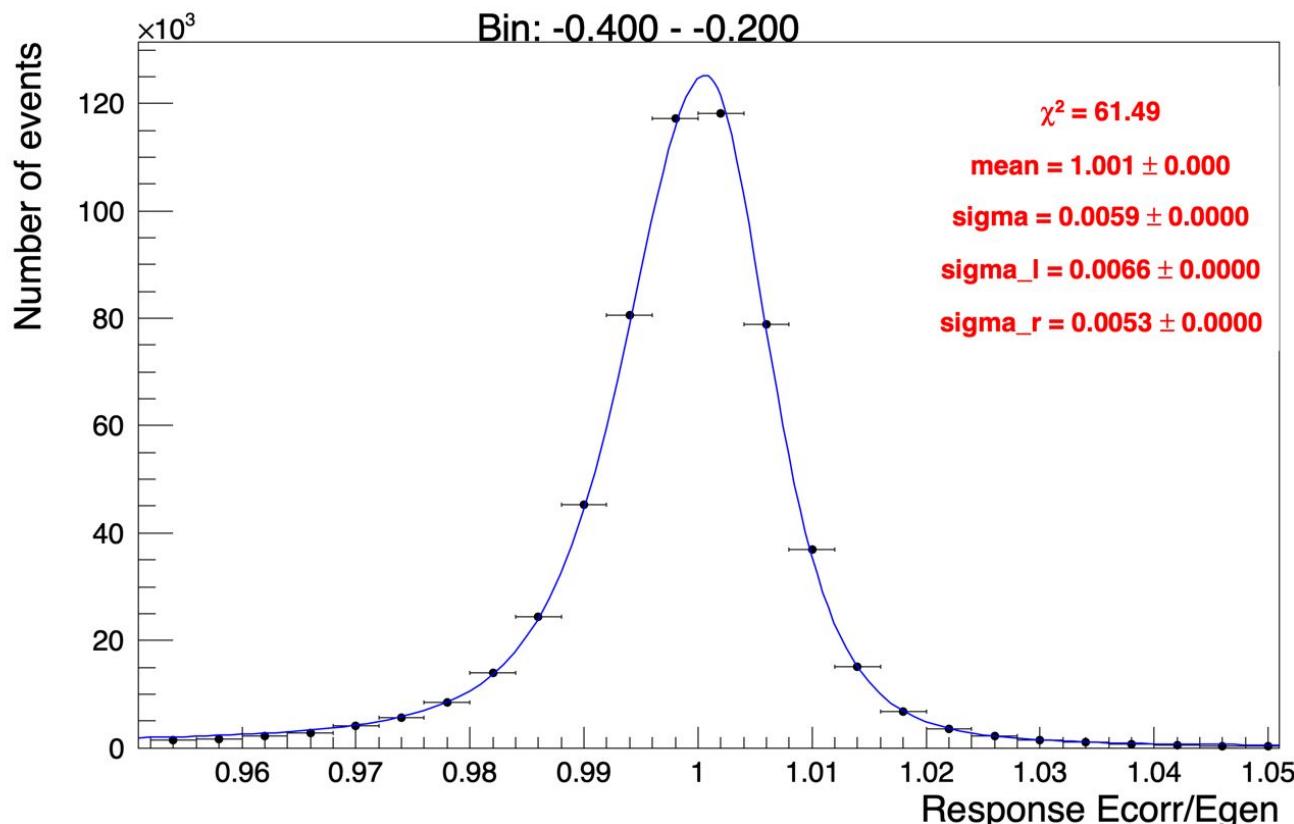


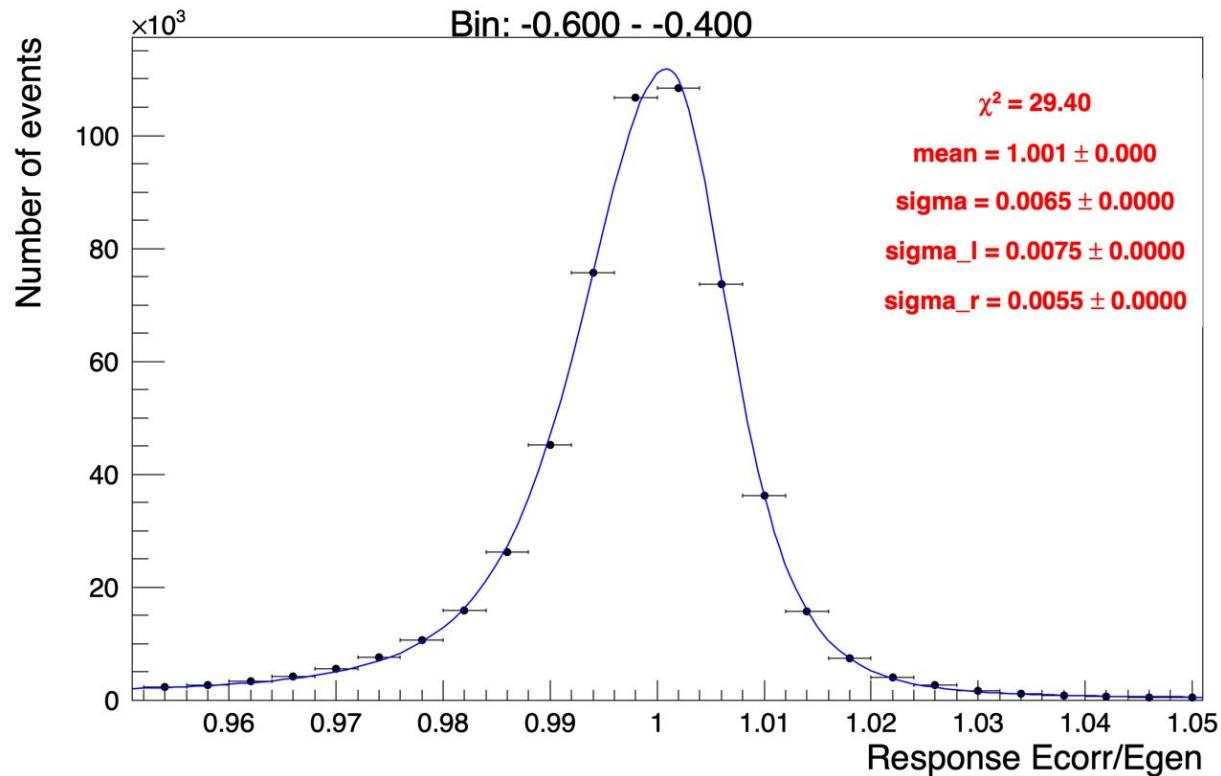
Bin: 280.000 - 300.000

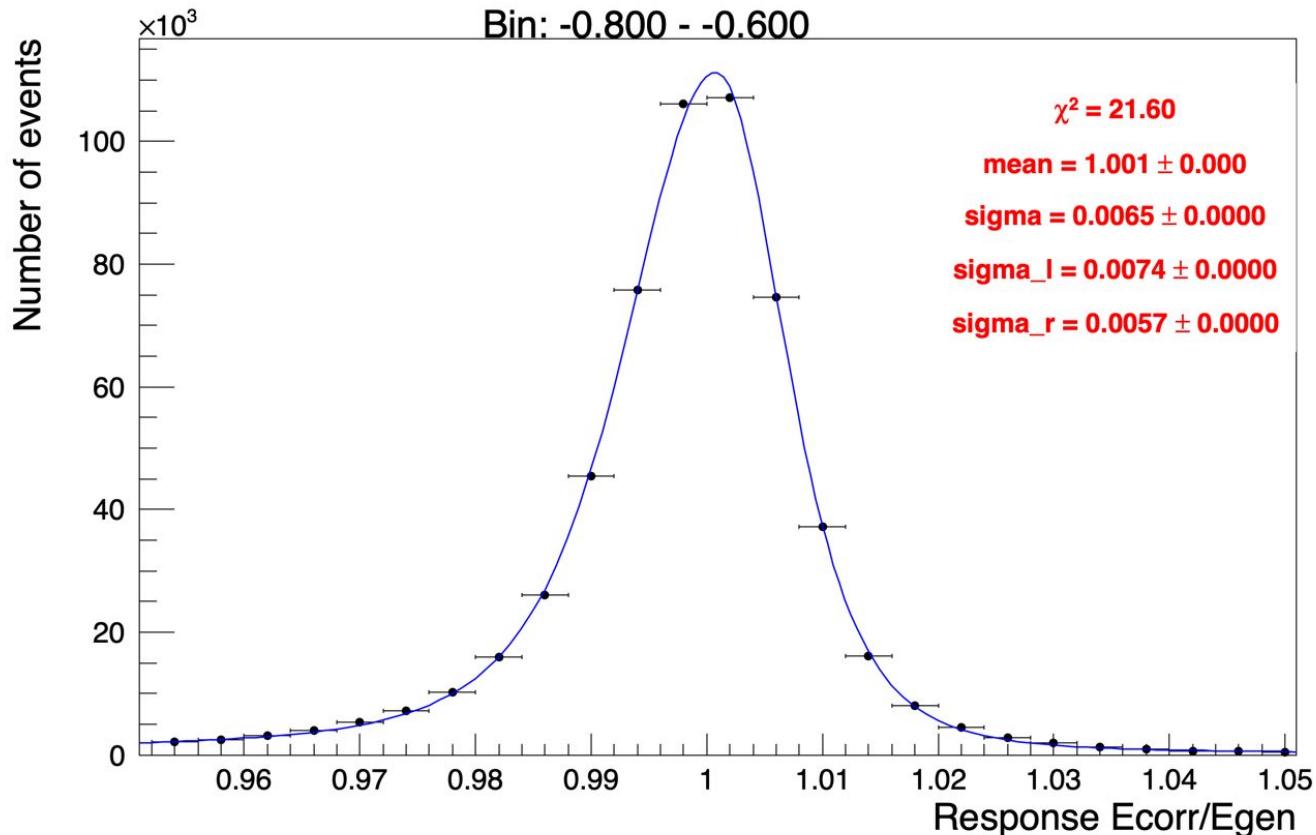


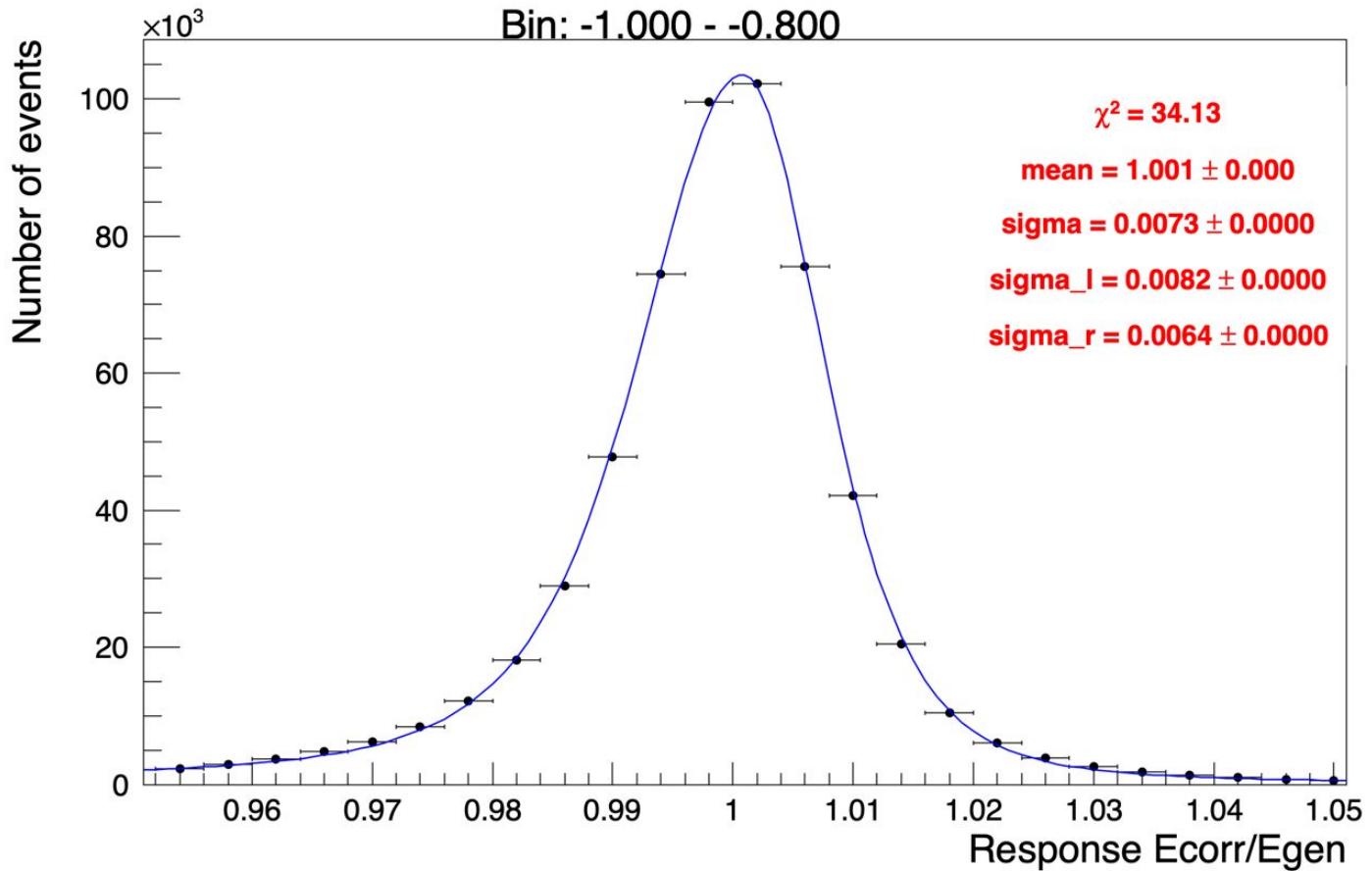
# Eta corrected

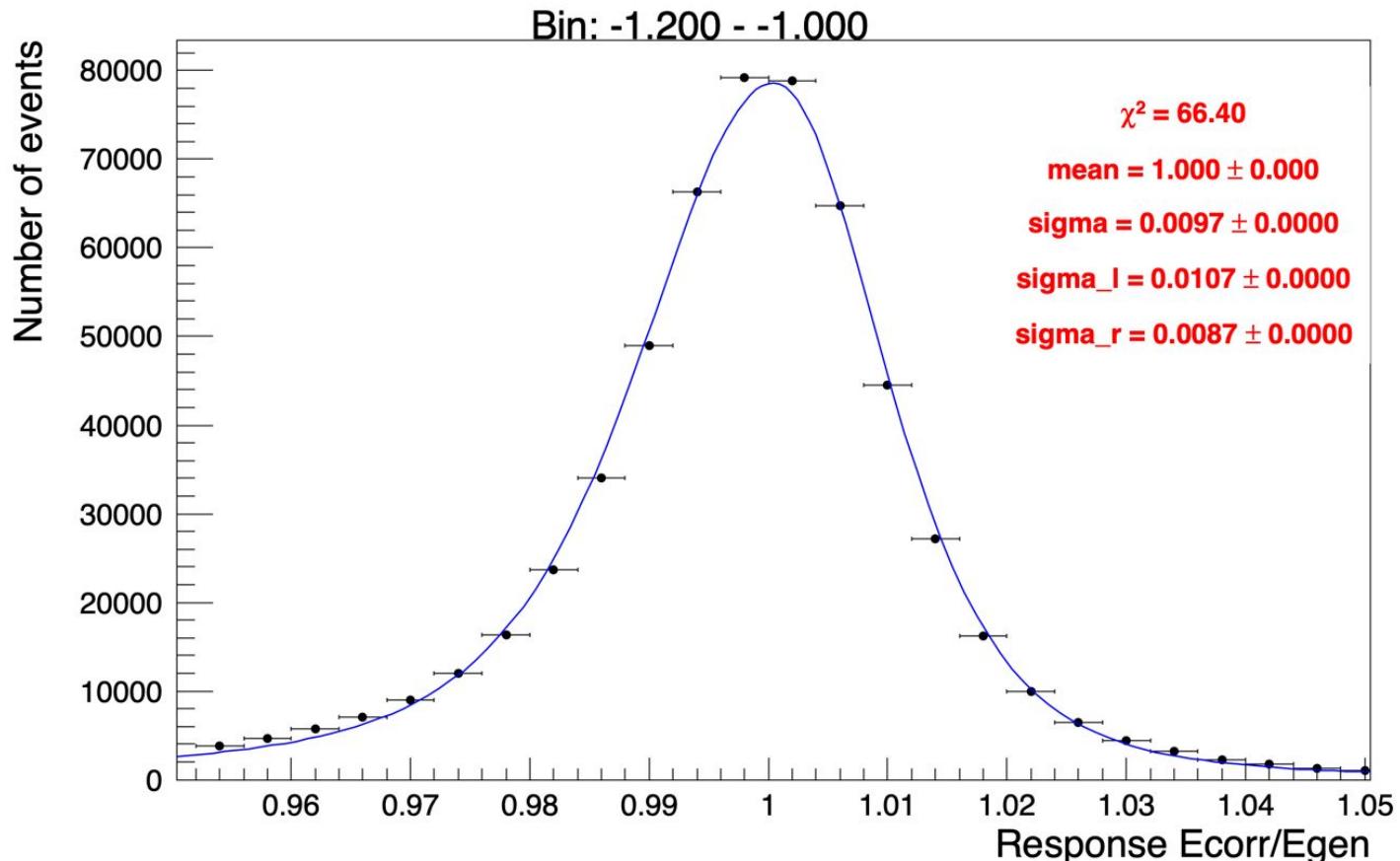


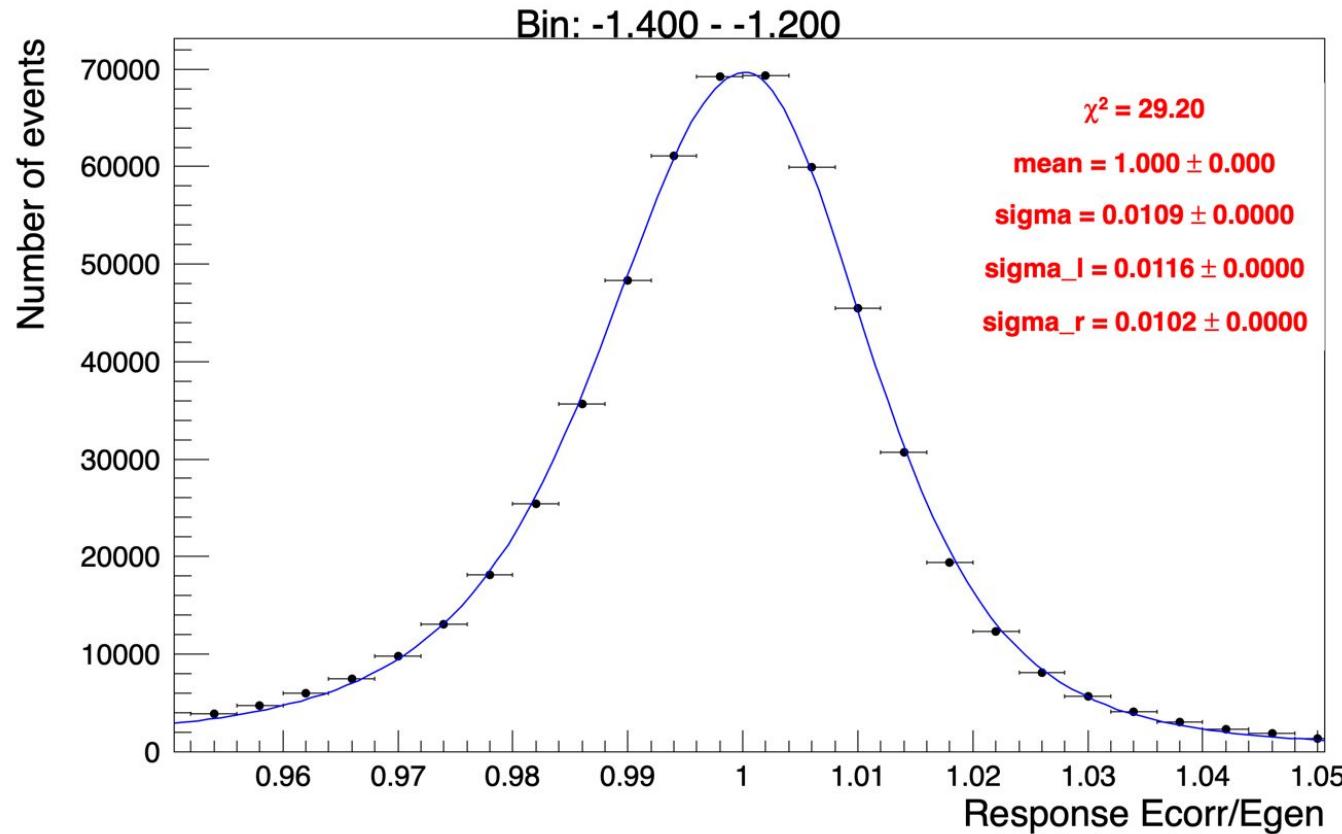


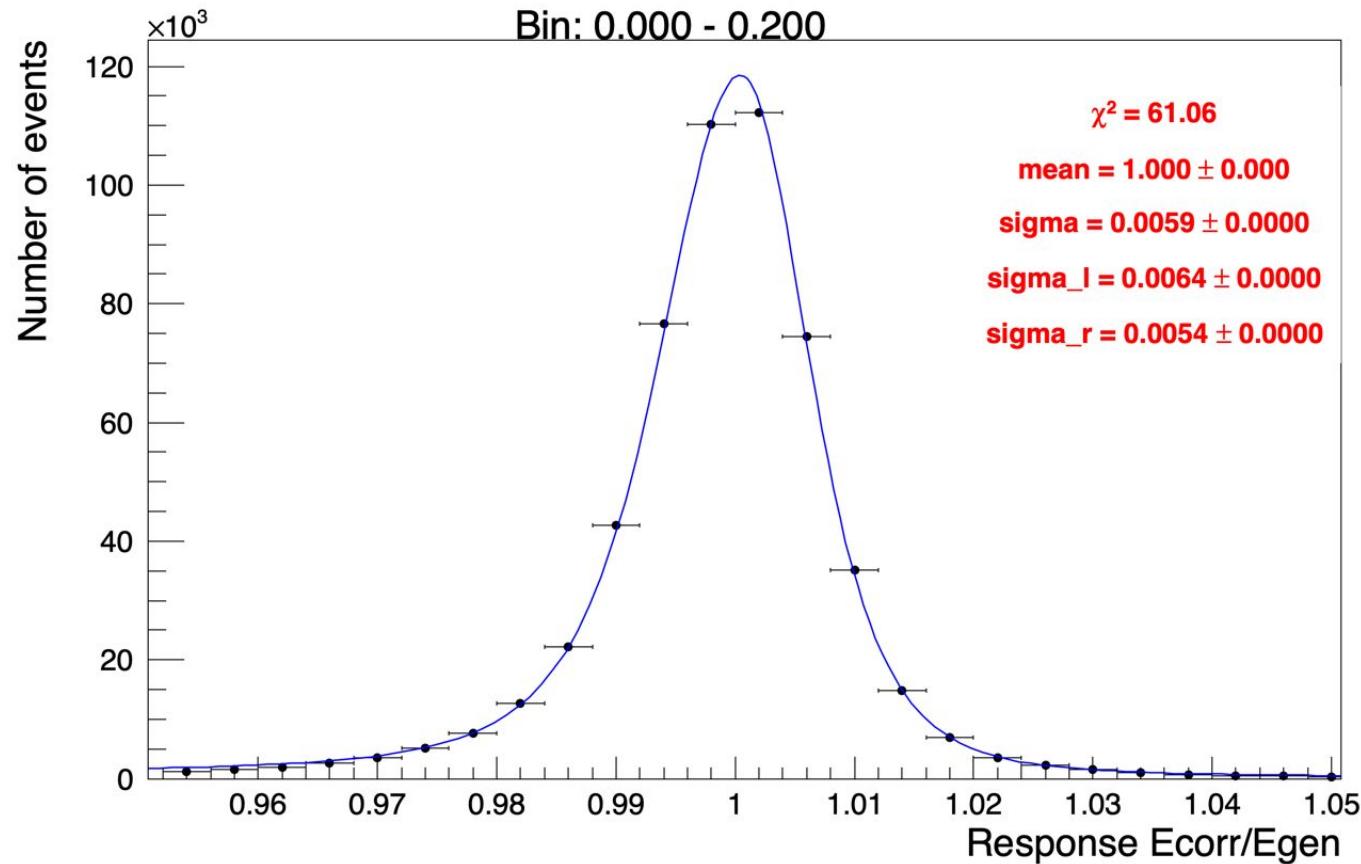


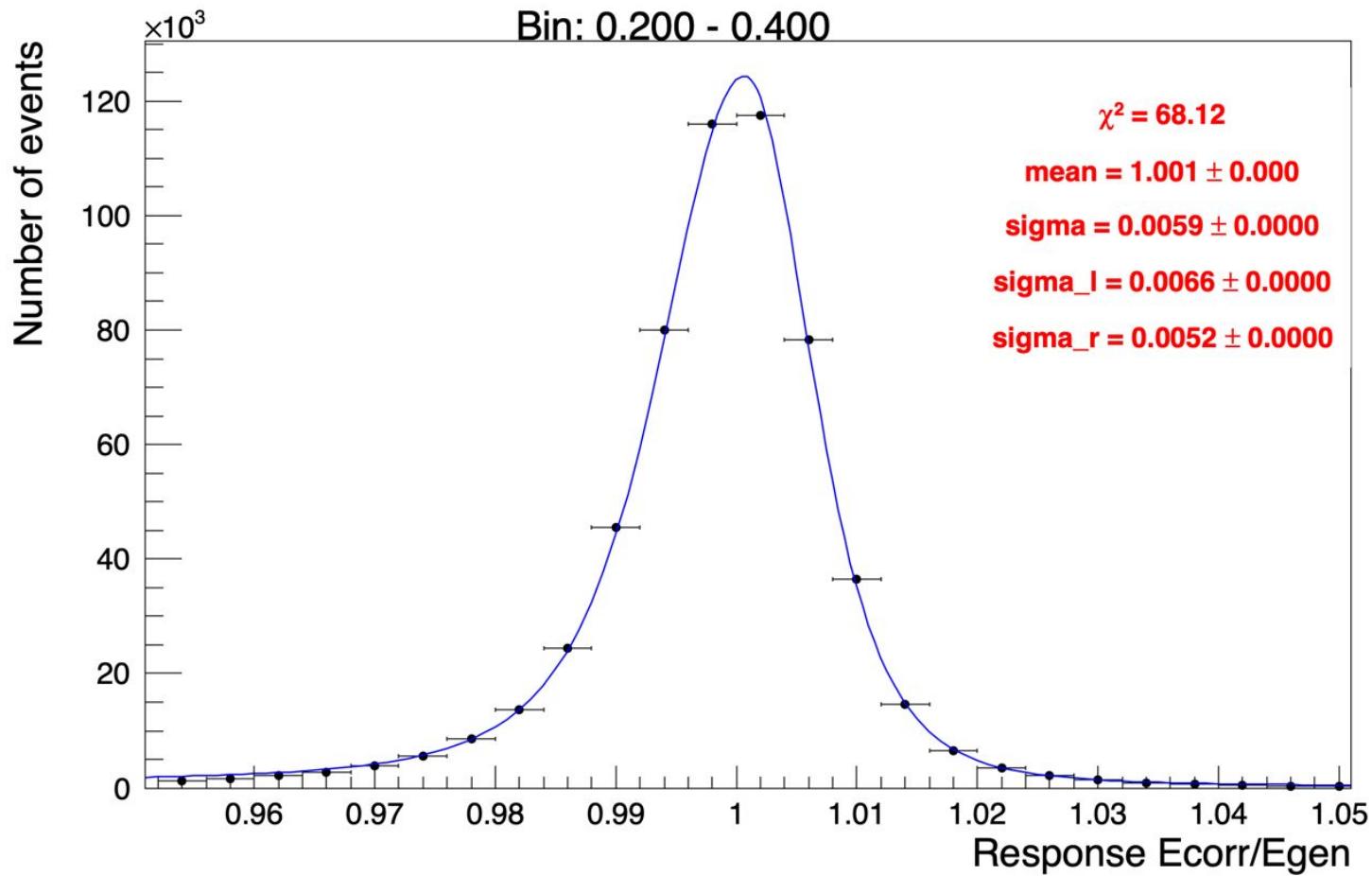


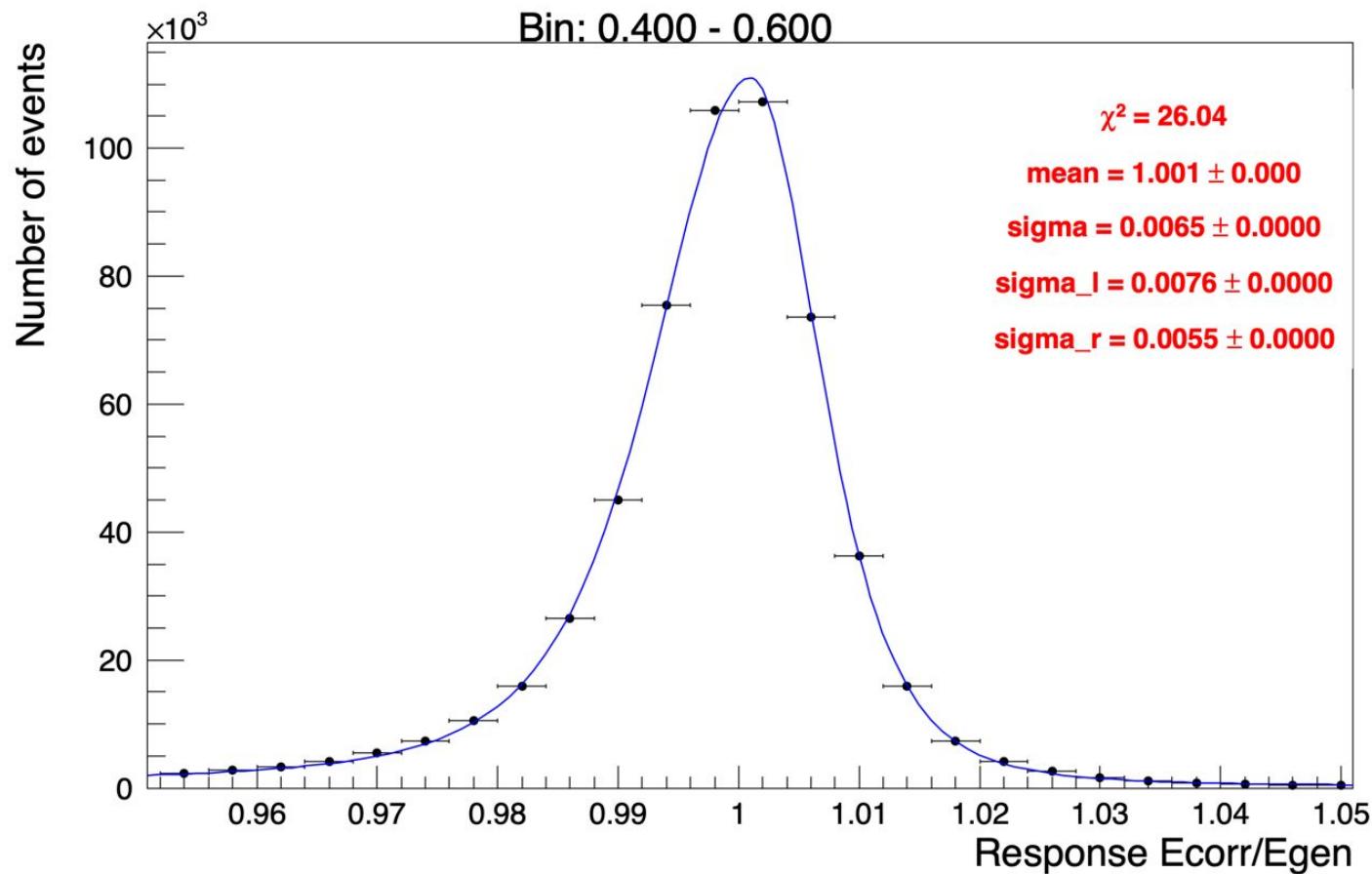


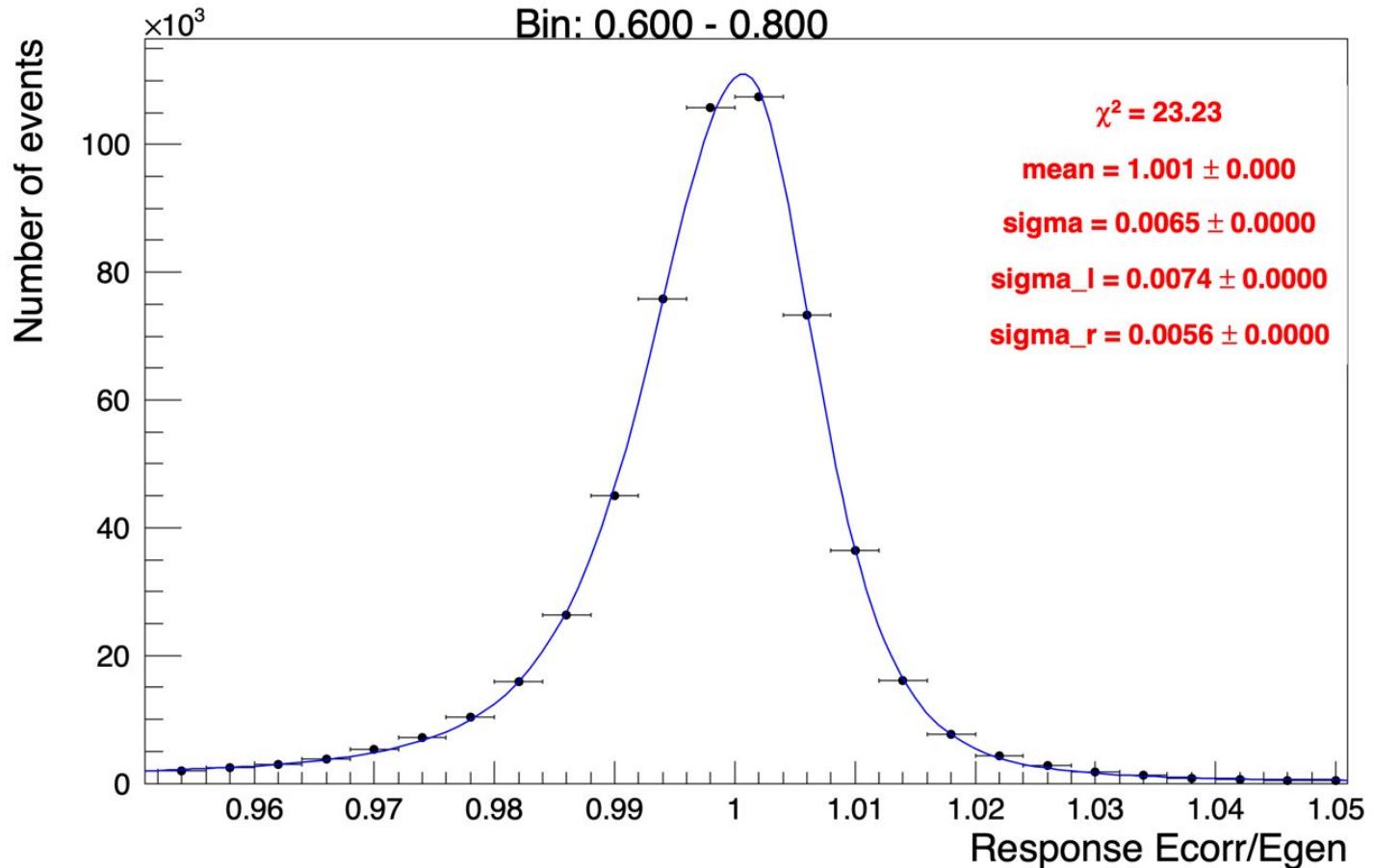


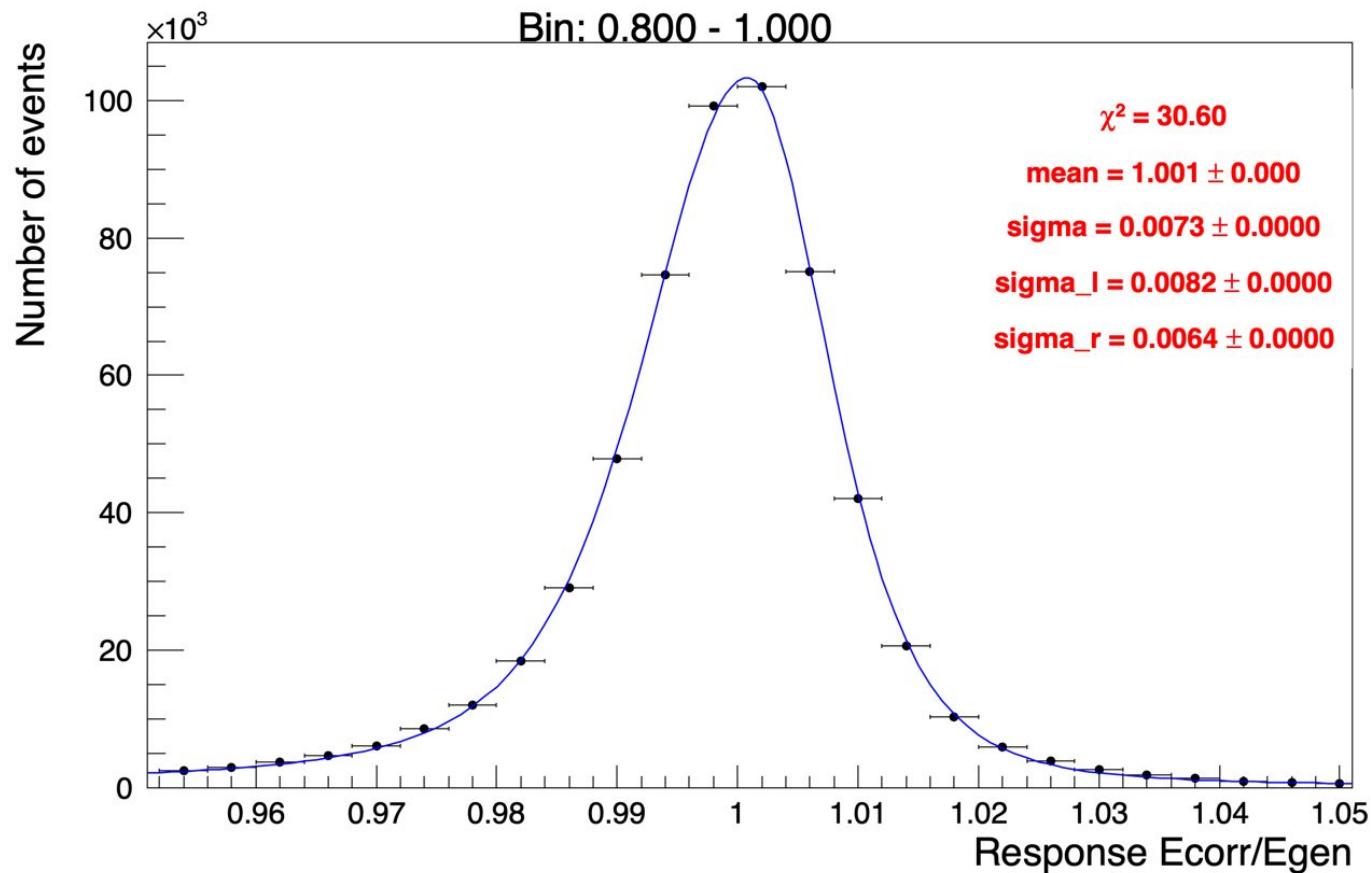


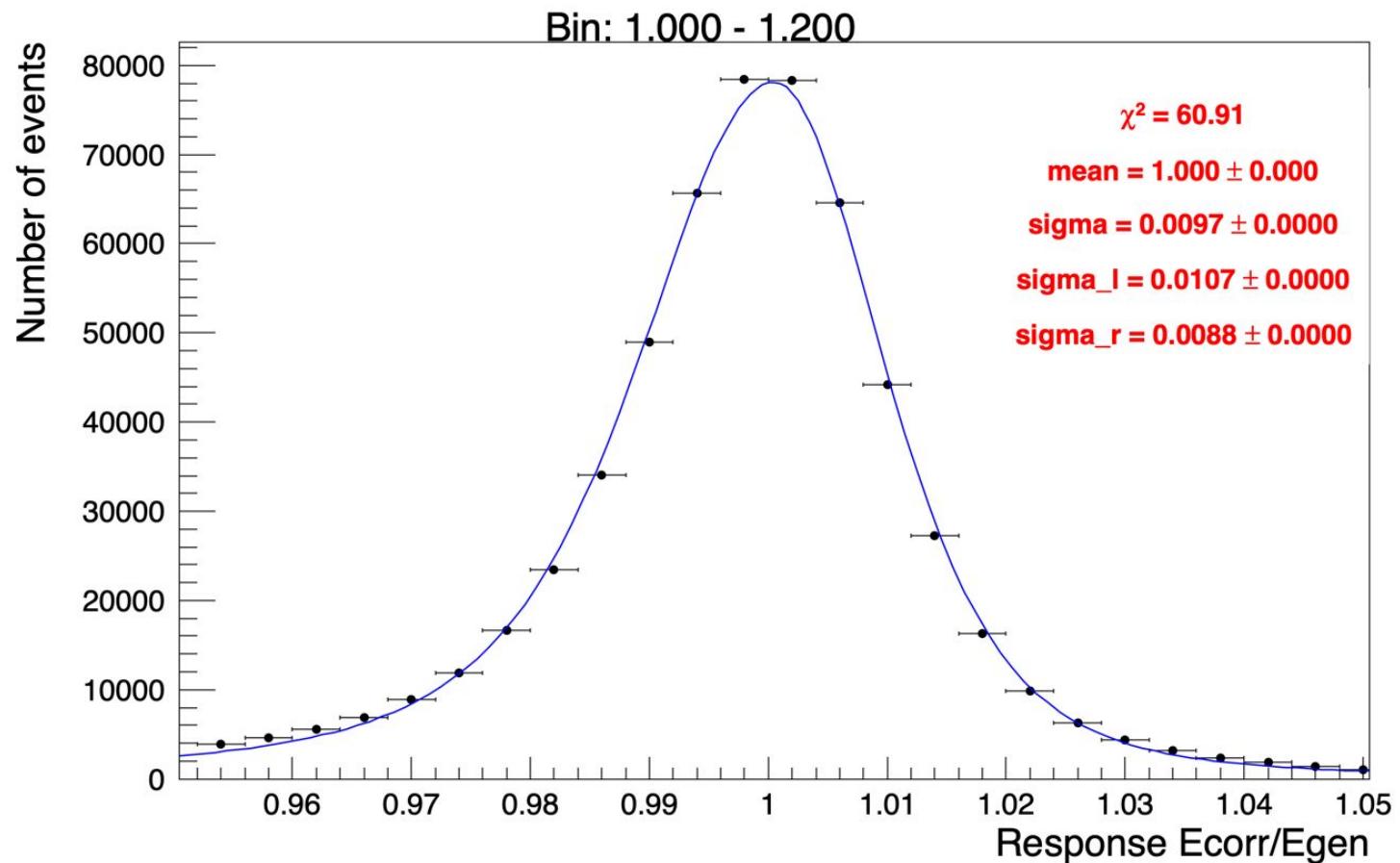


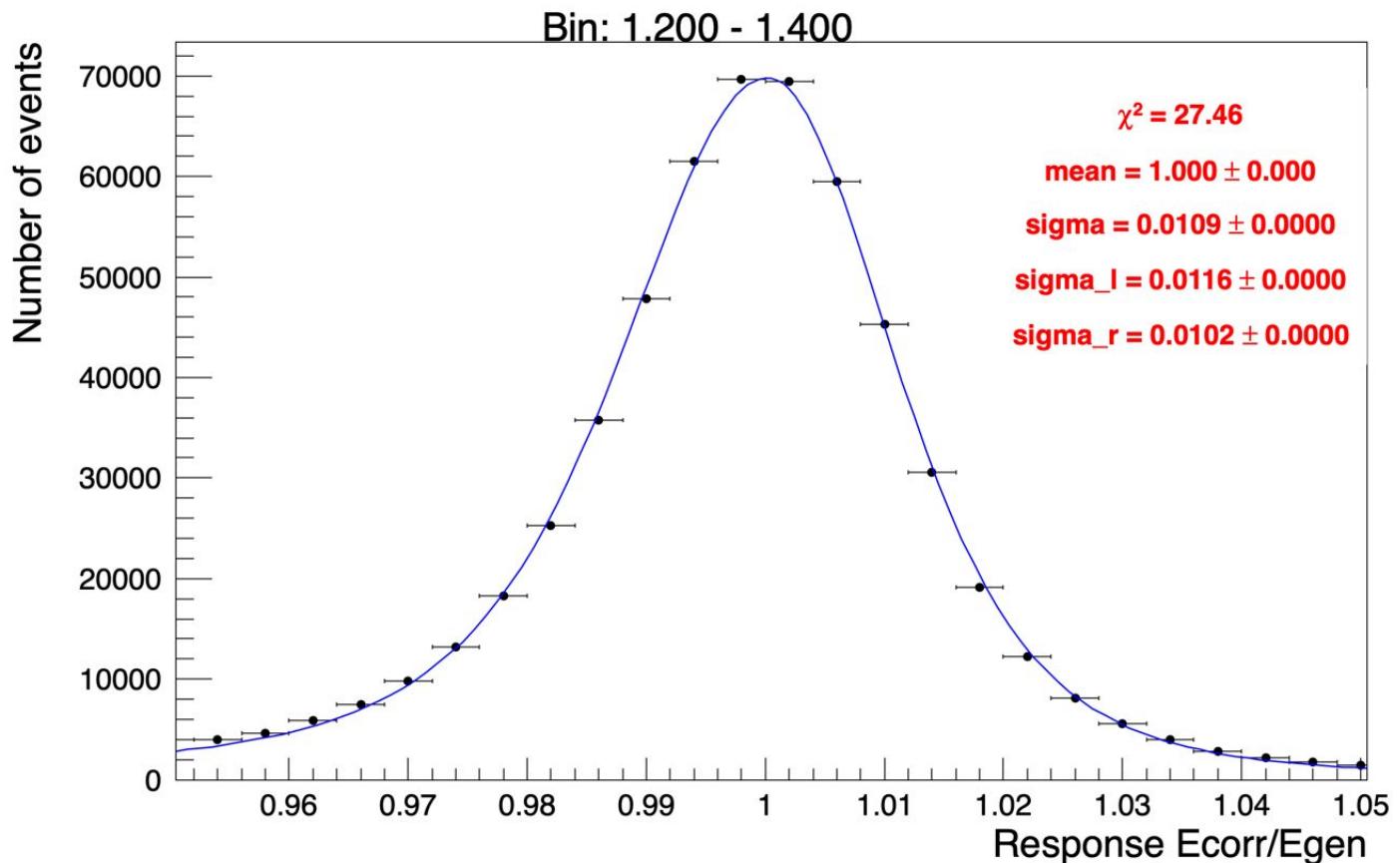




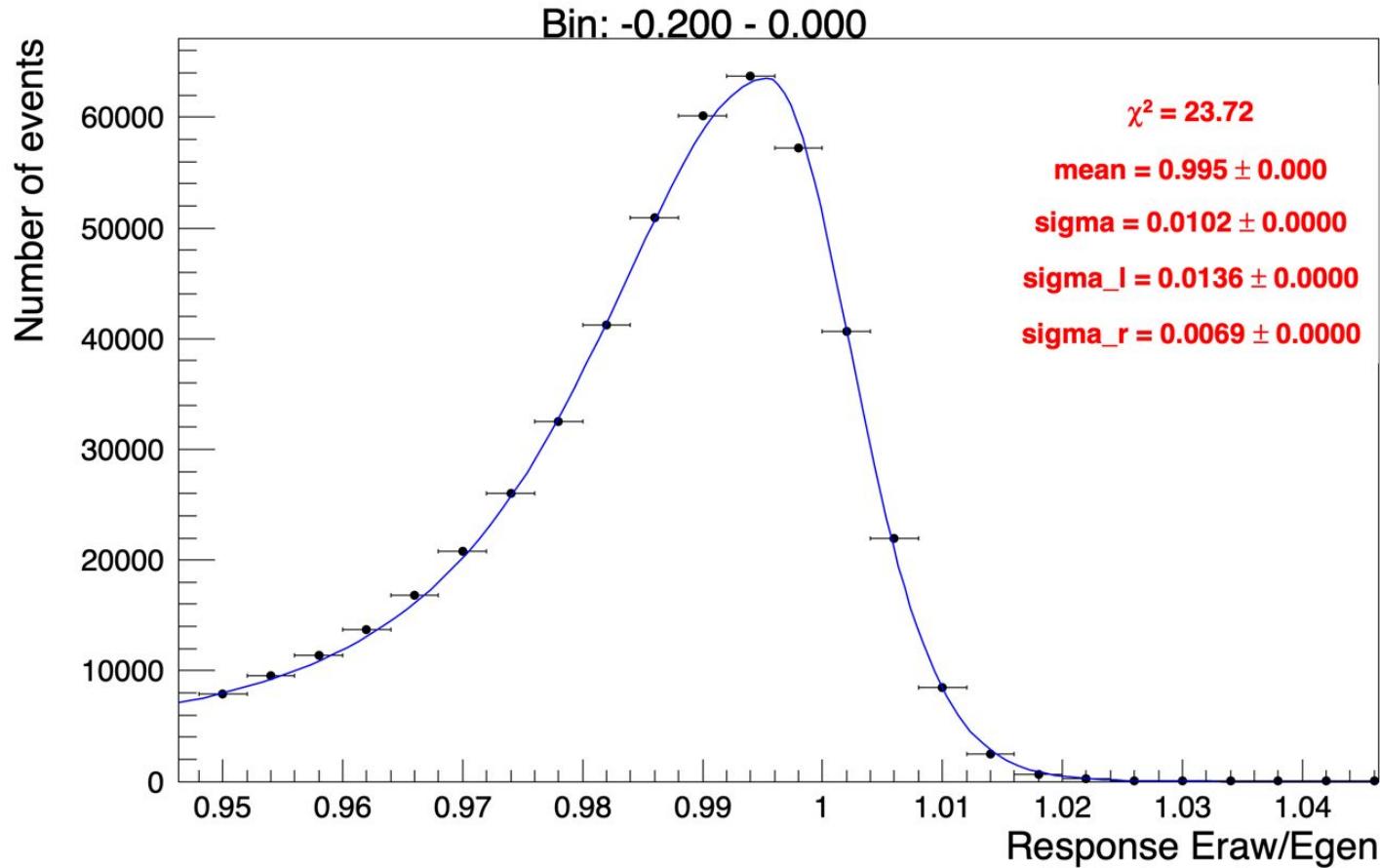




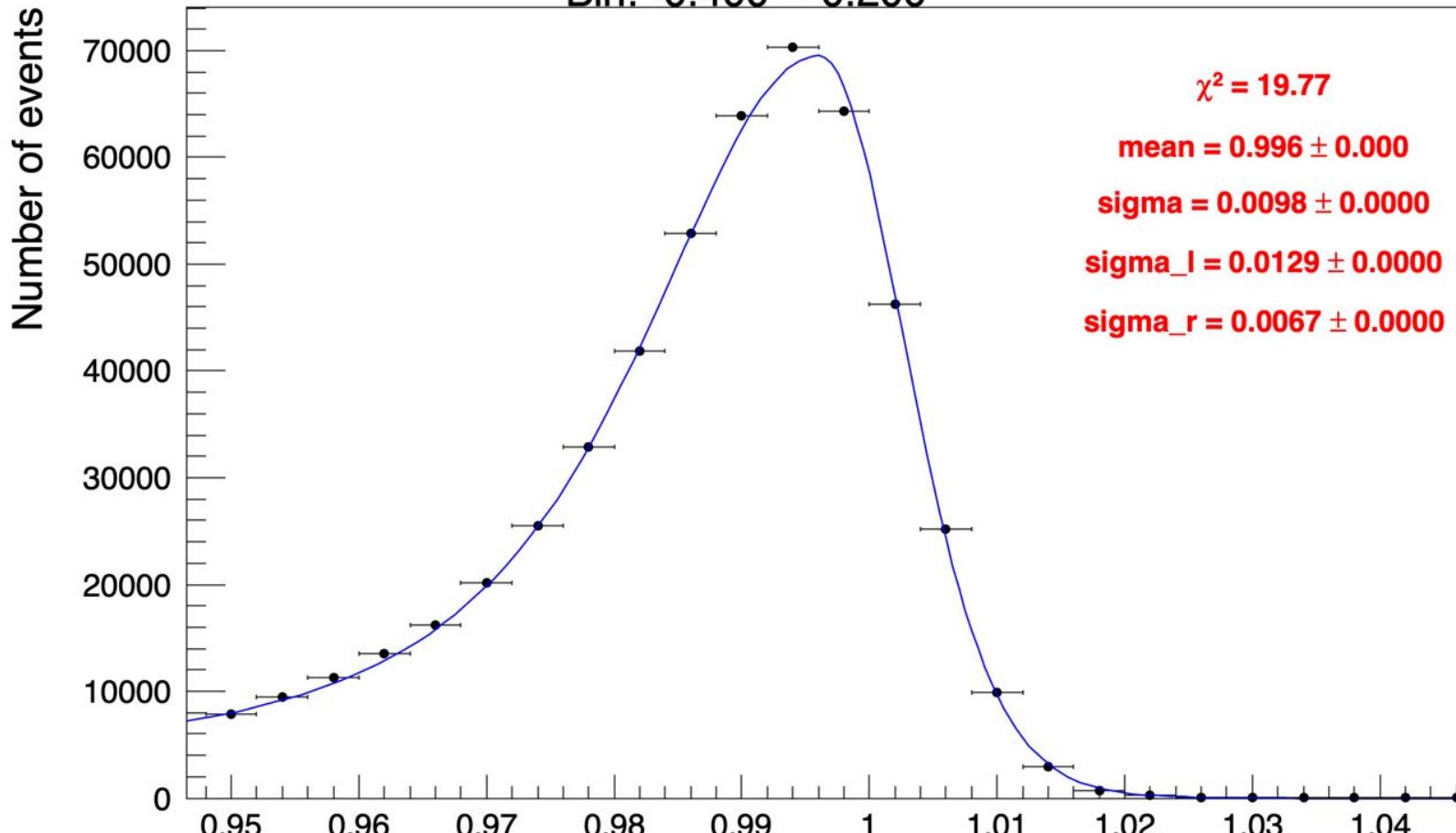




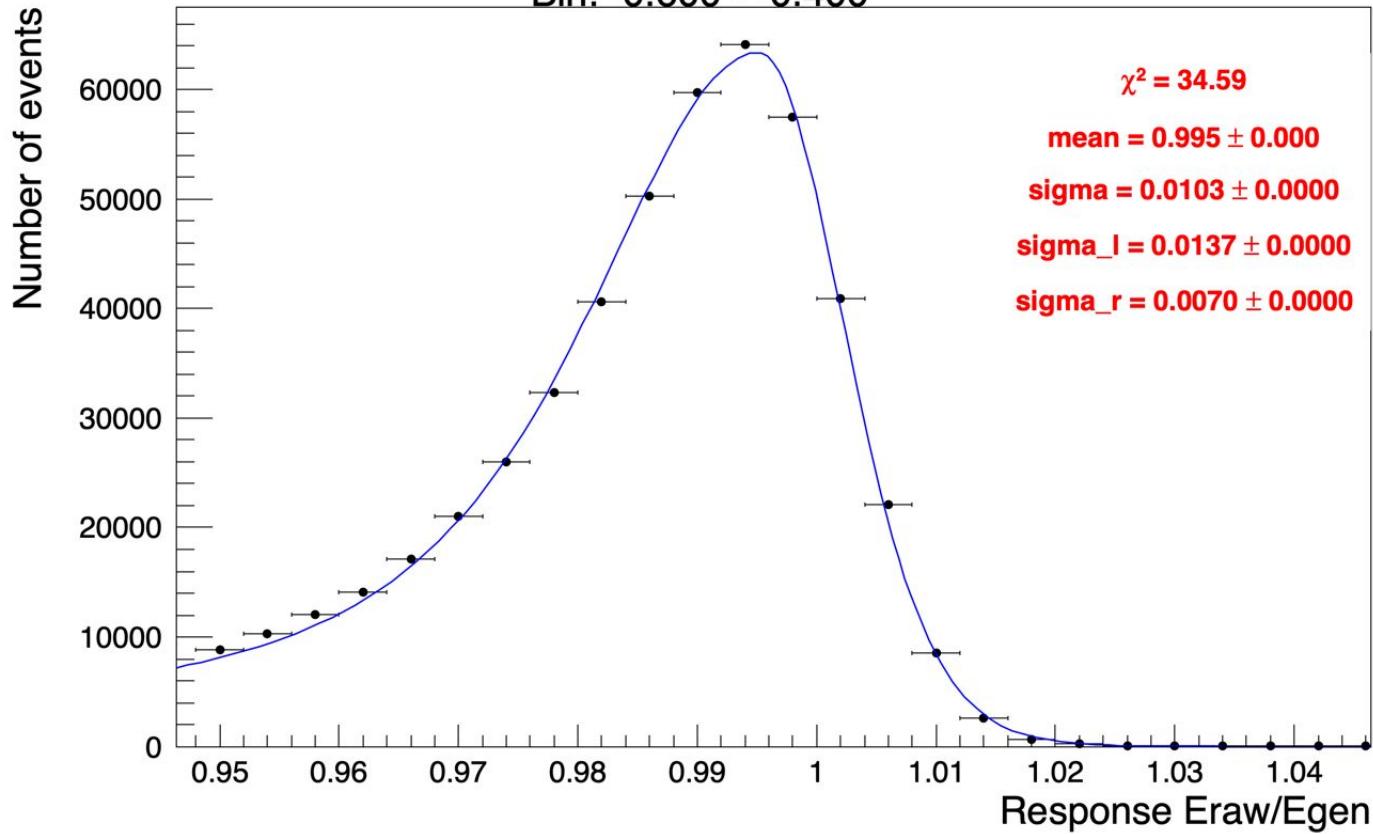
# Eta raw

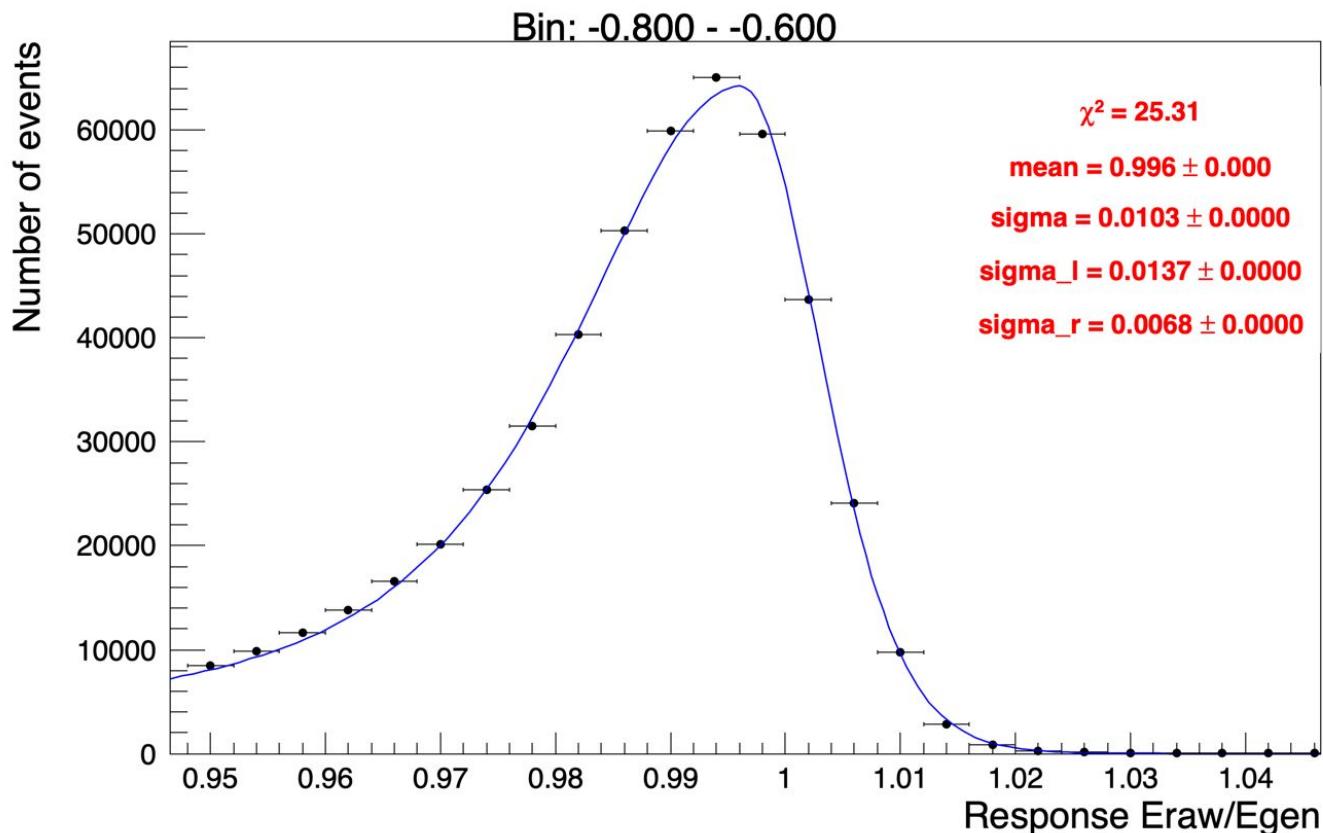


**Bin: -0.400 - -0.200**

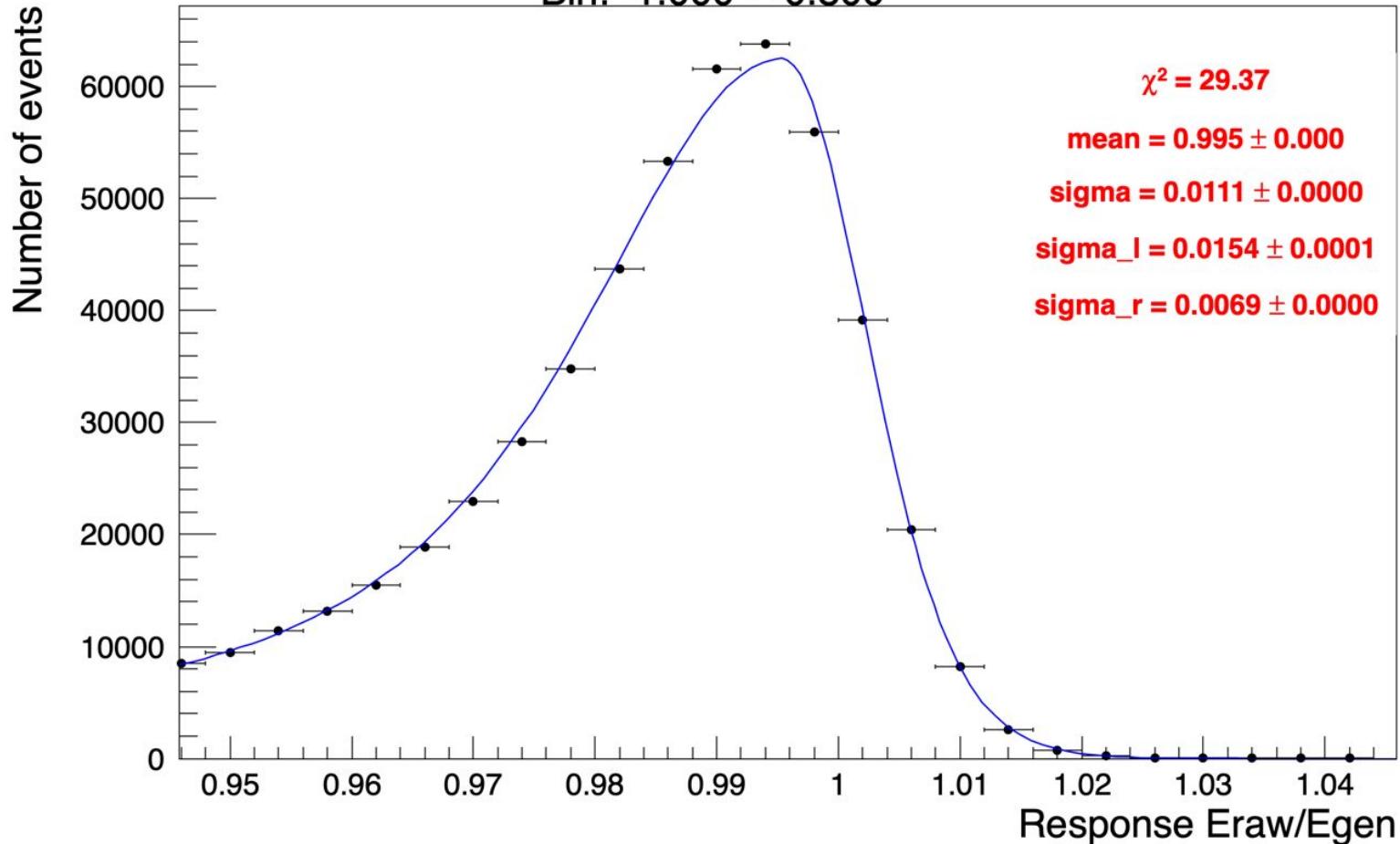


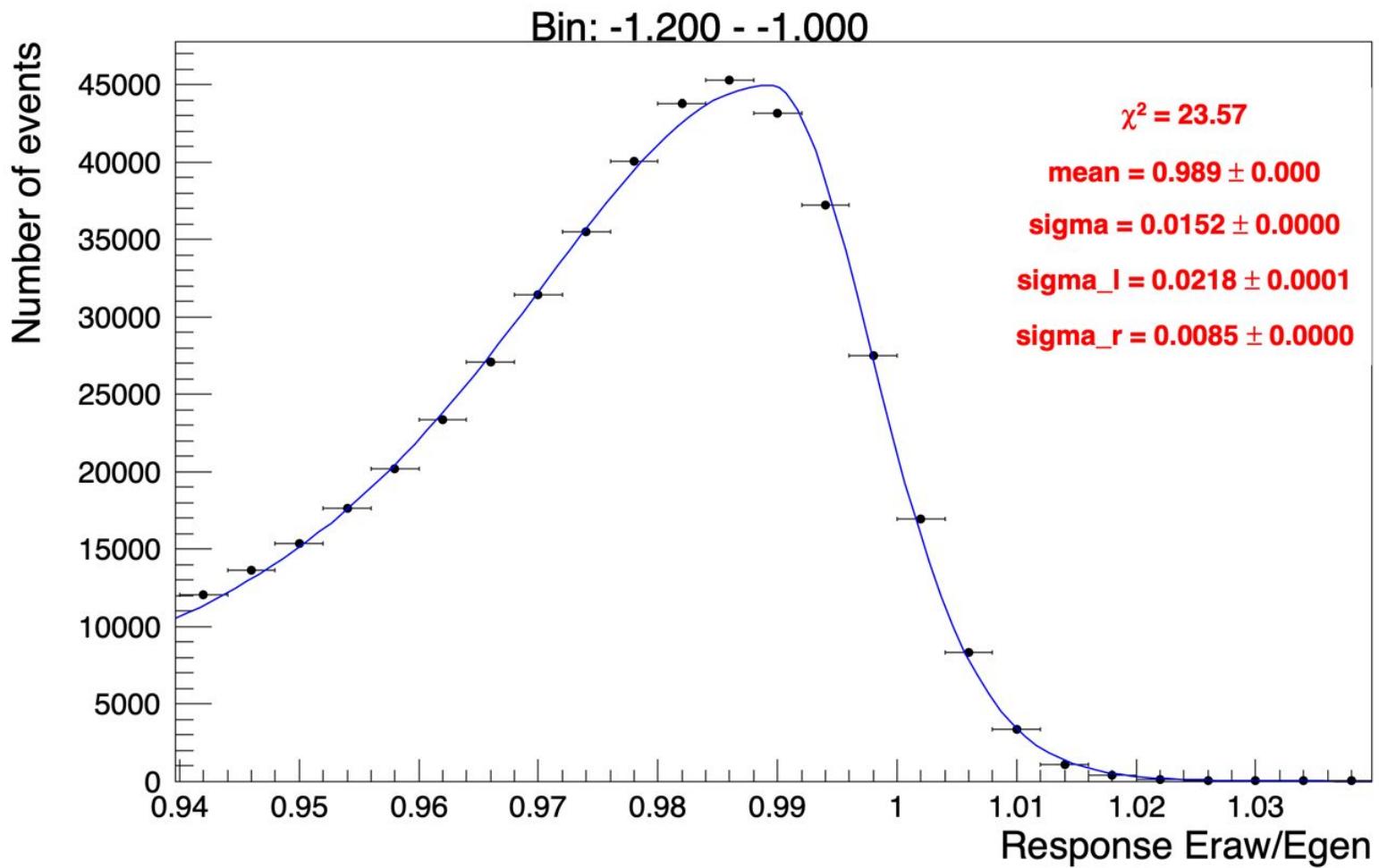
Bin: -0.600 - -0.400

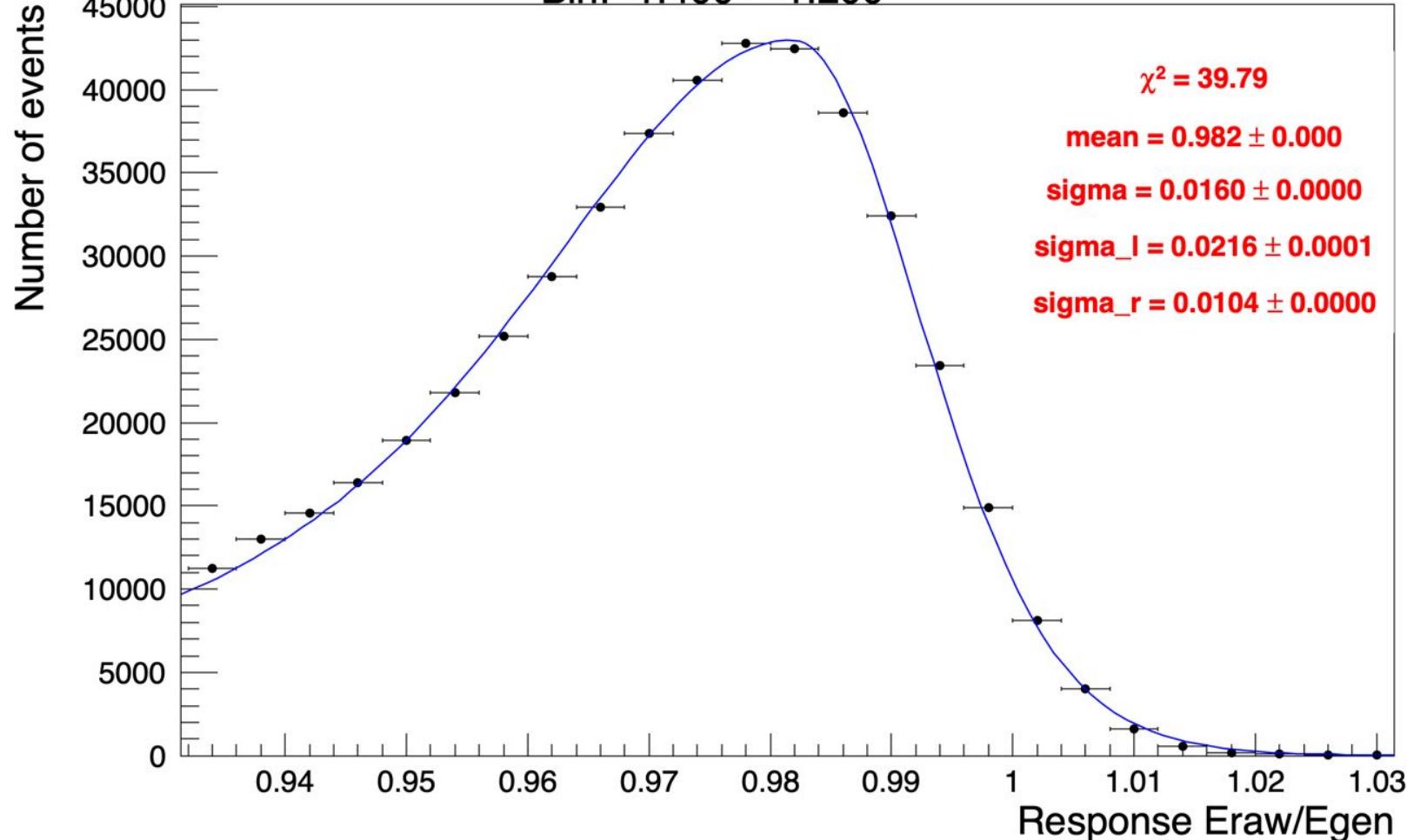


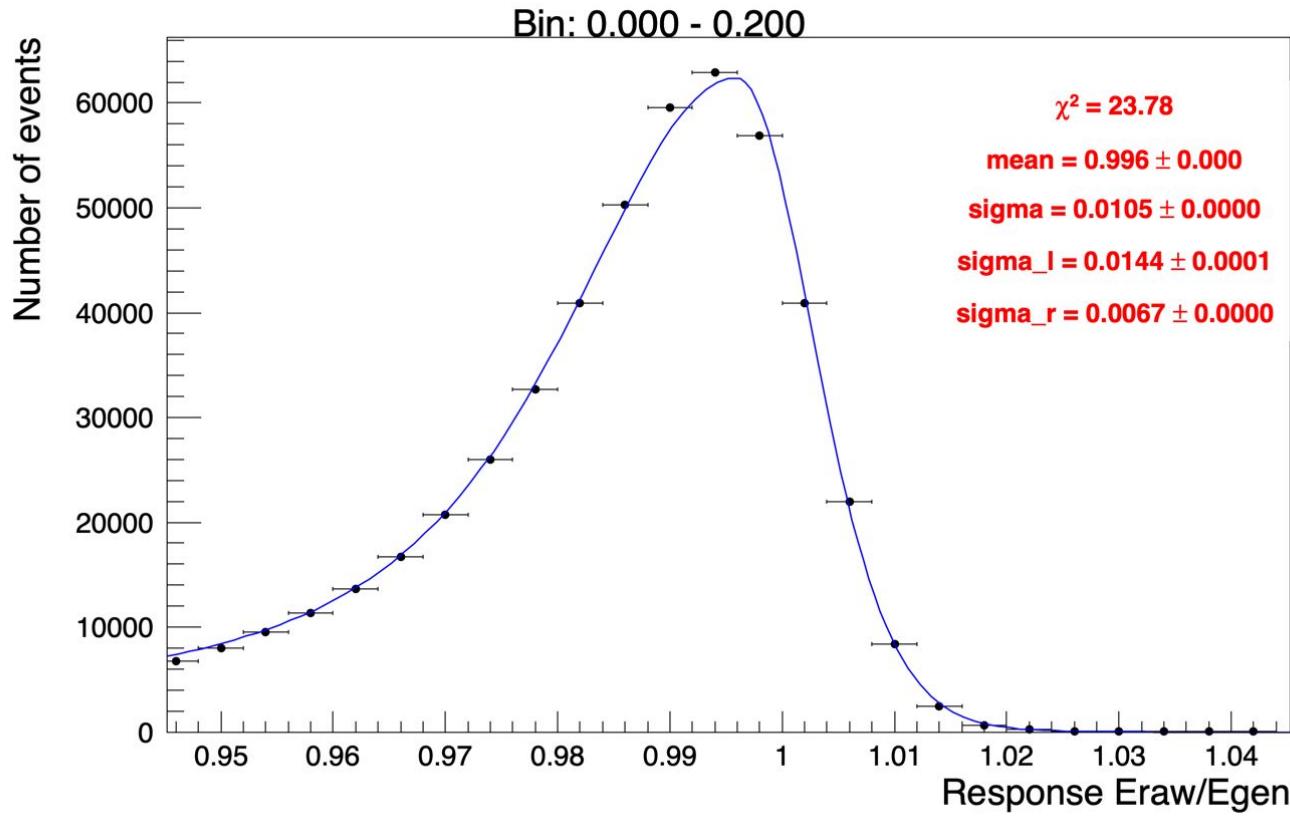


Bin: -1.000 - -0.800

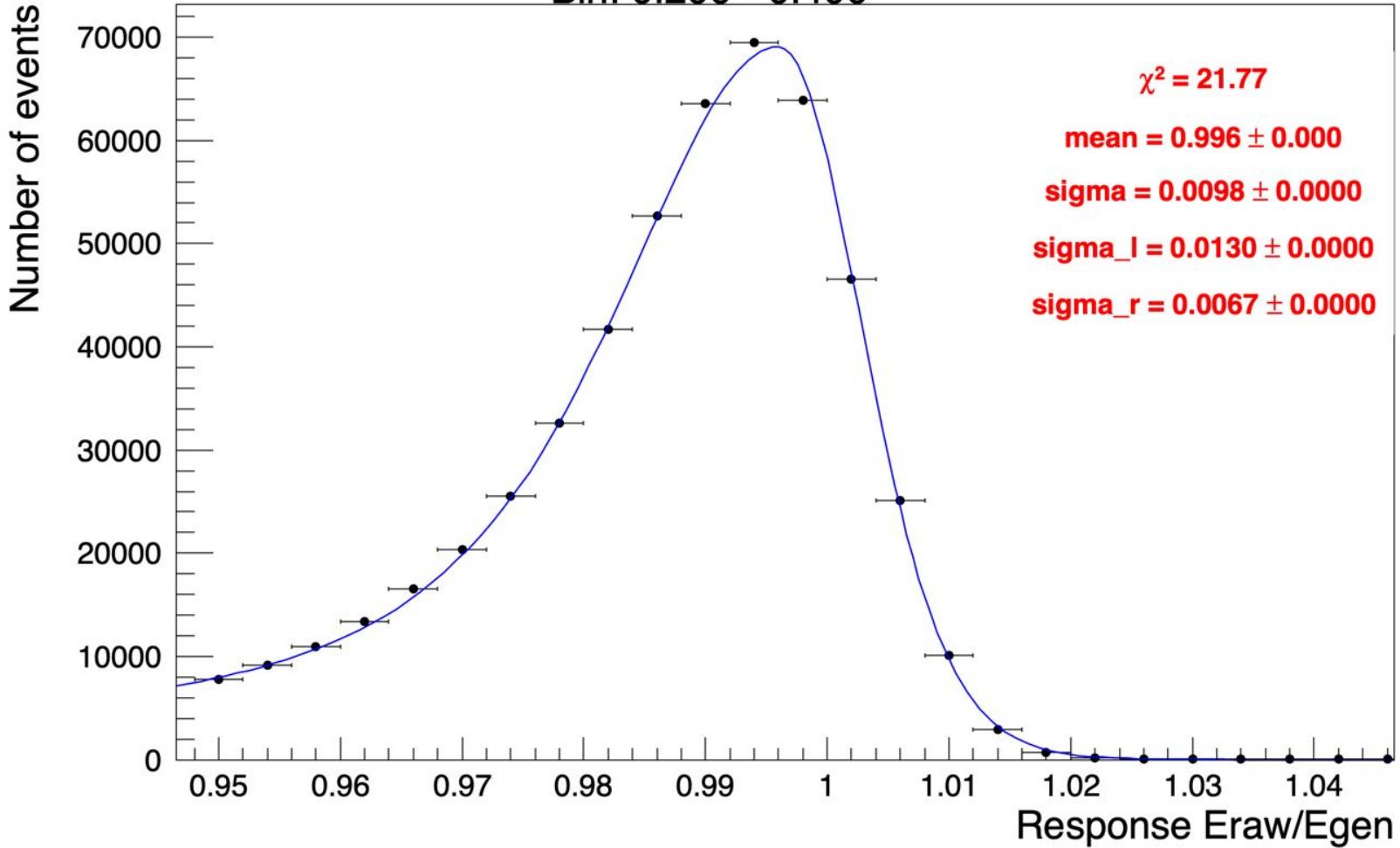




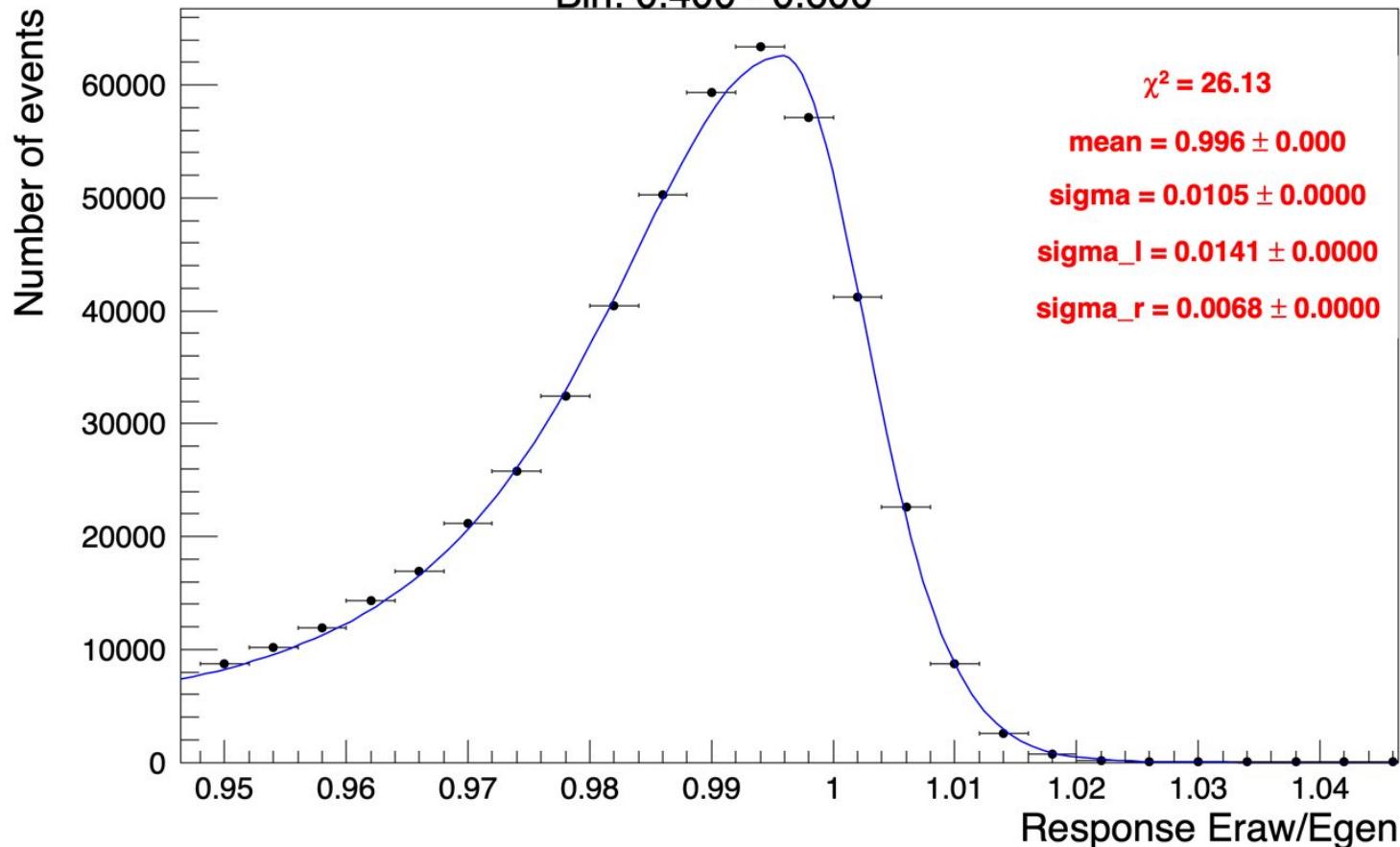




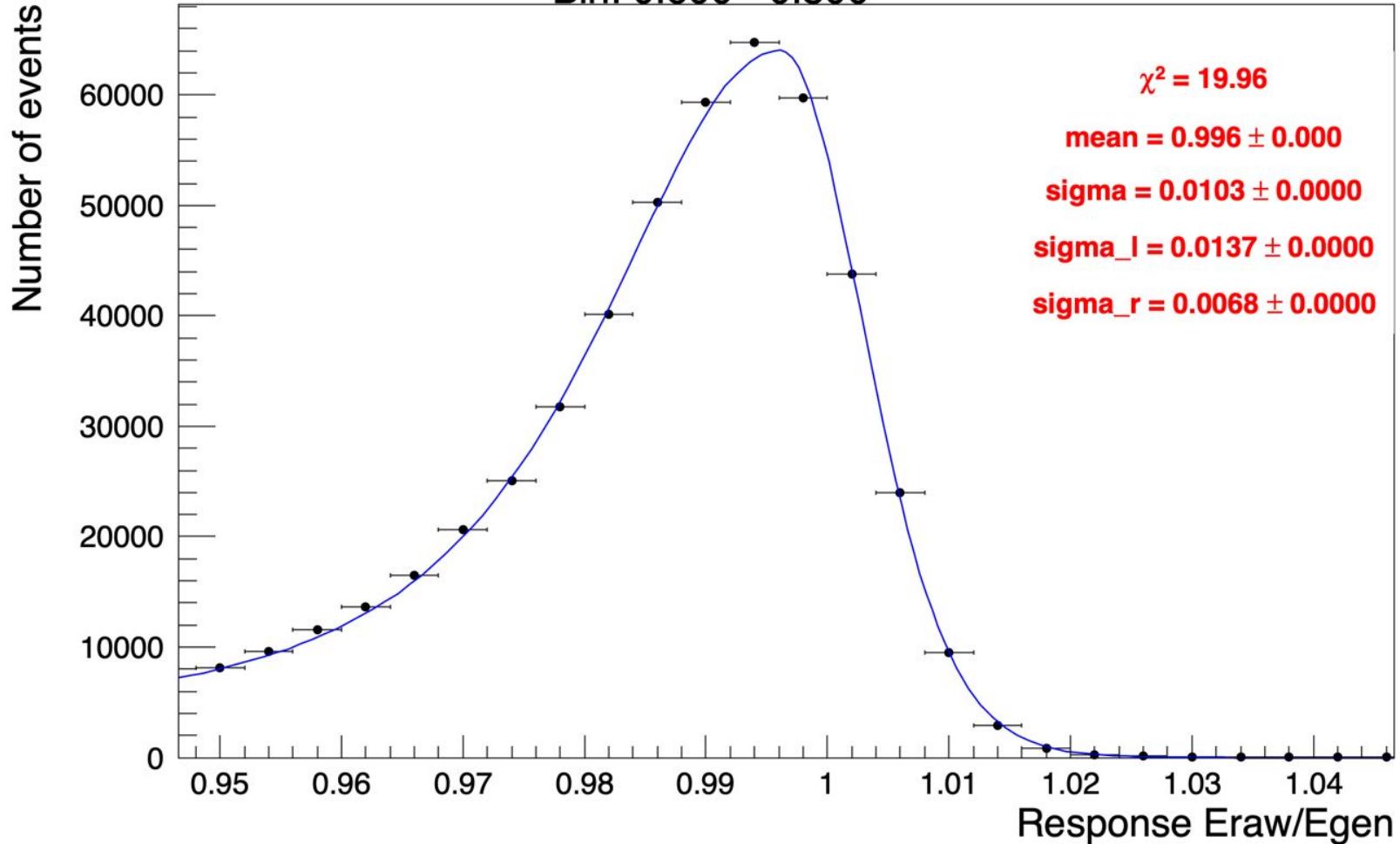
Bin: 0.200 - 0.400

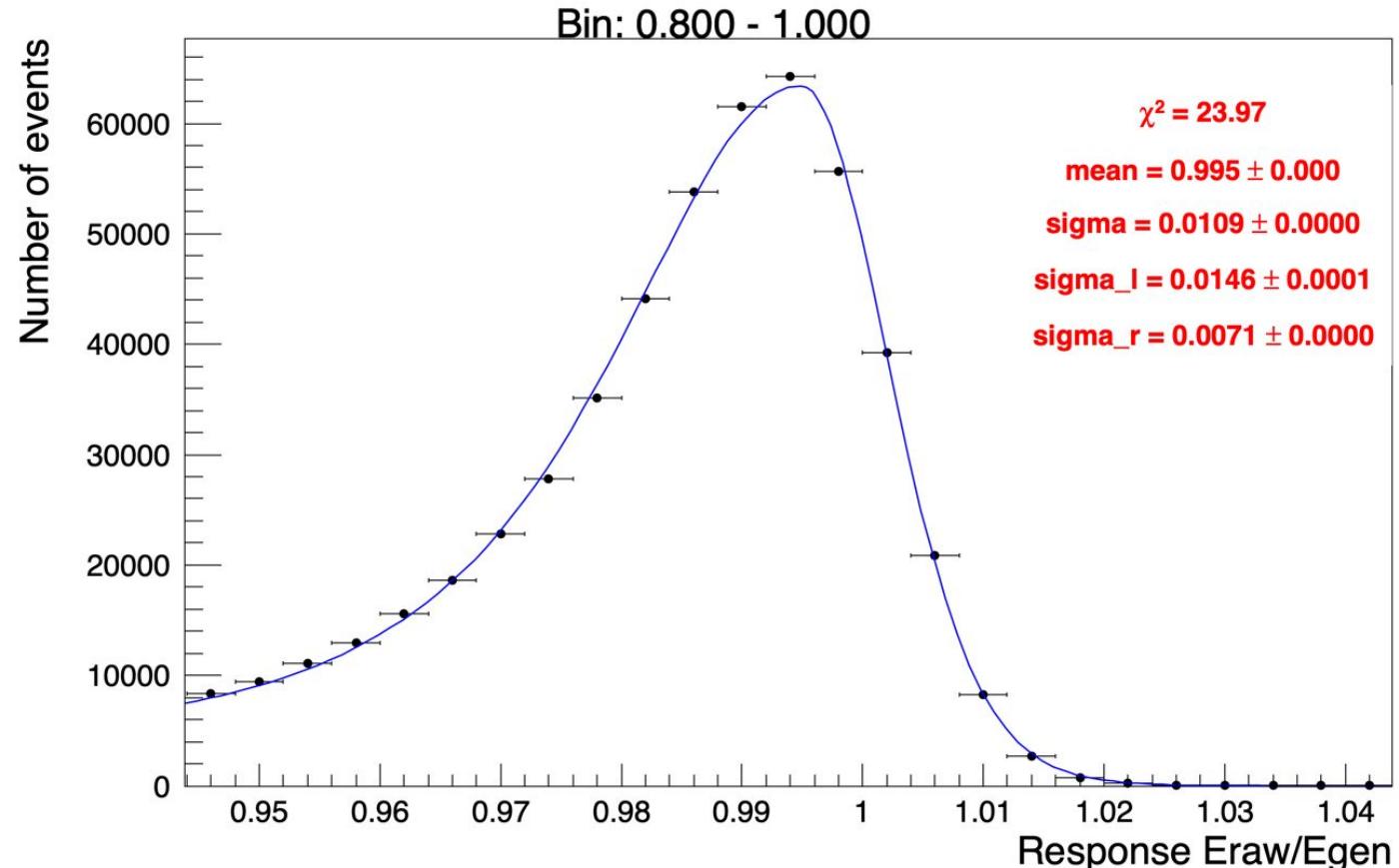


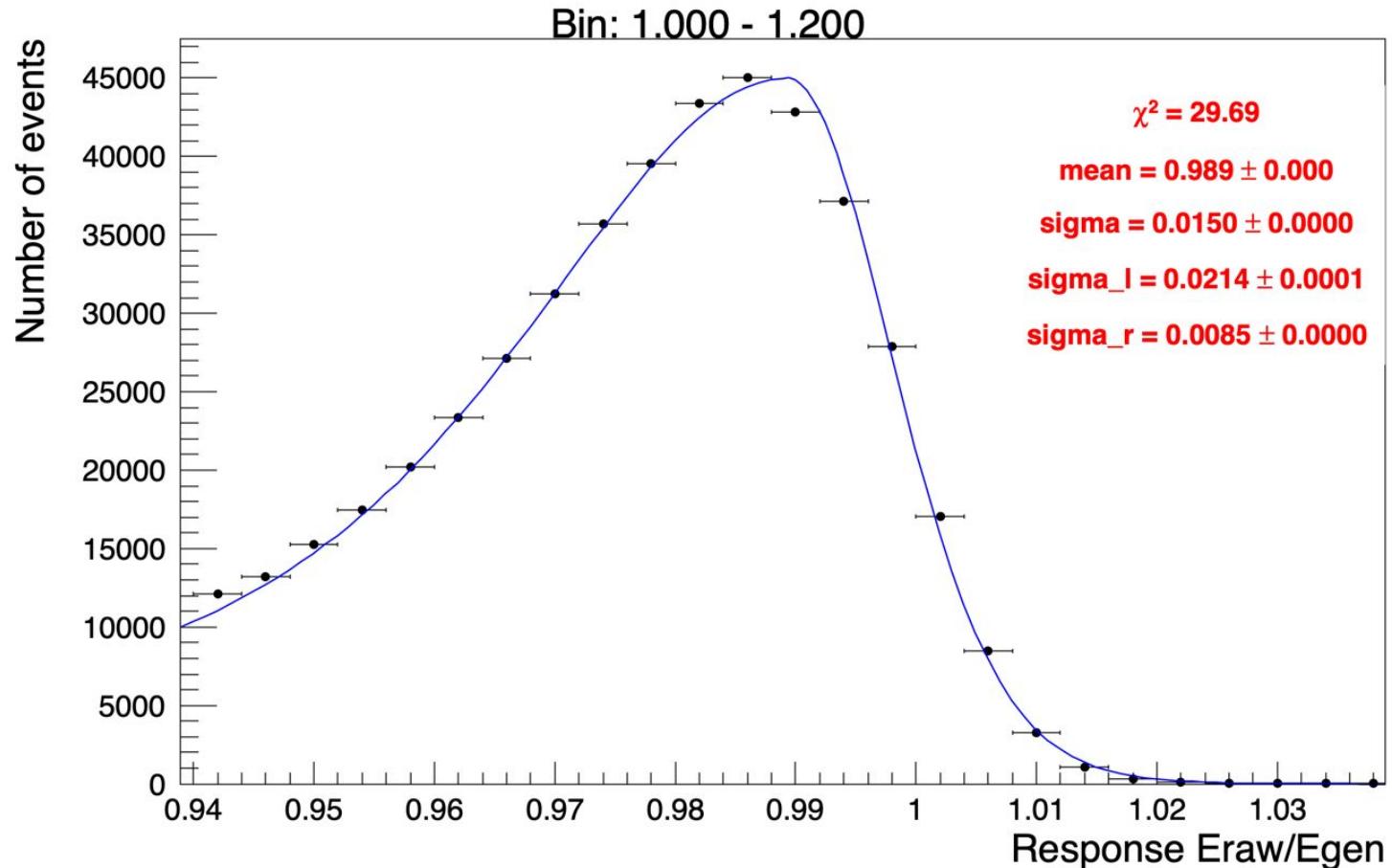
Bin: 0.400 - 0.600

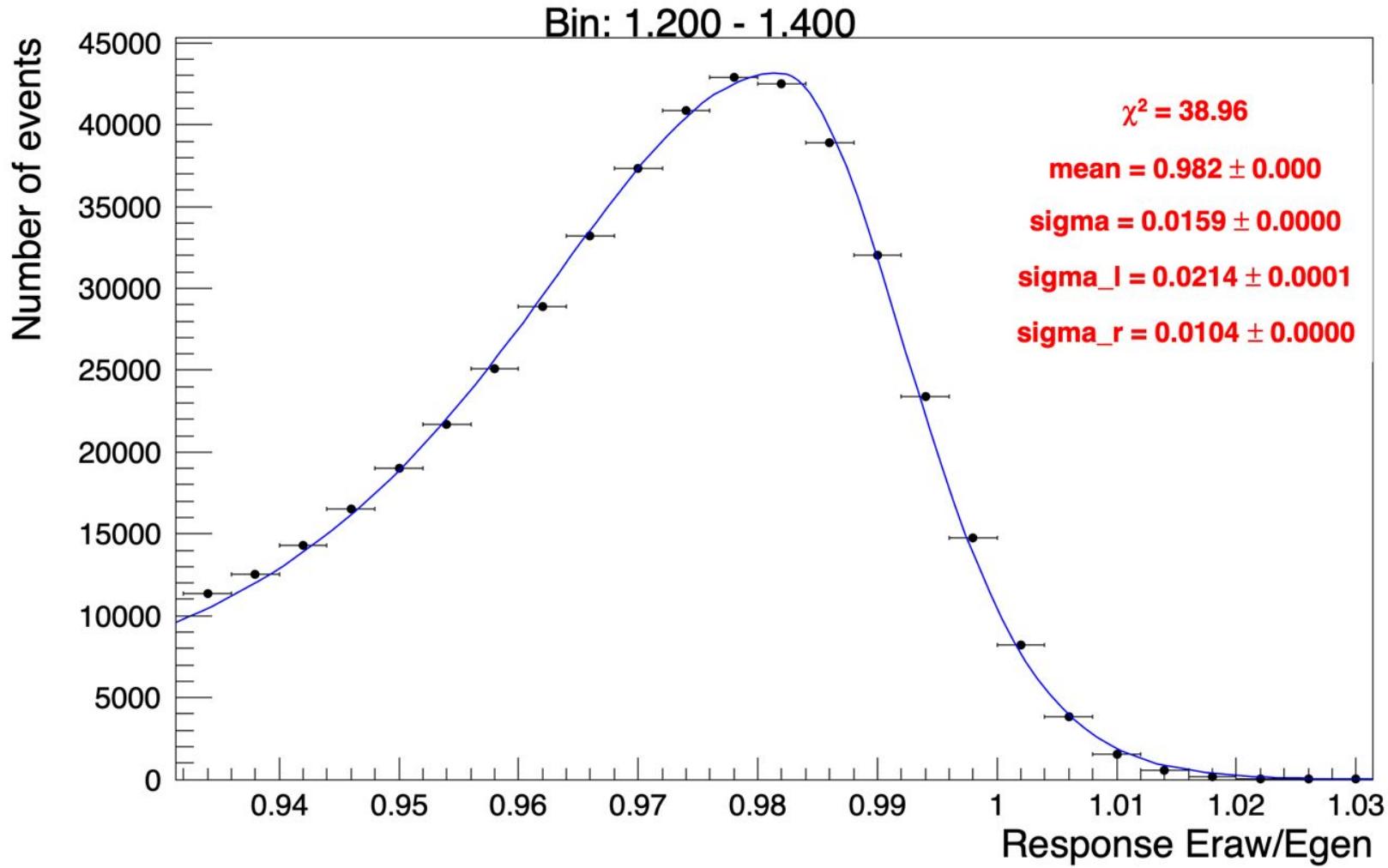


Bin: 0.600 - 0.800

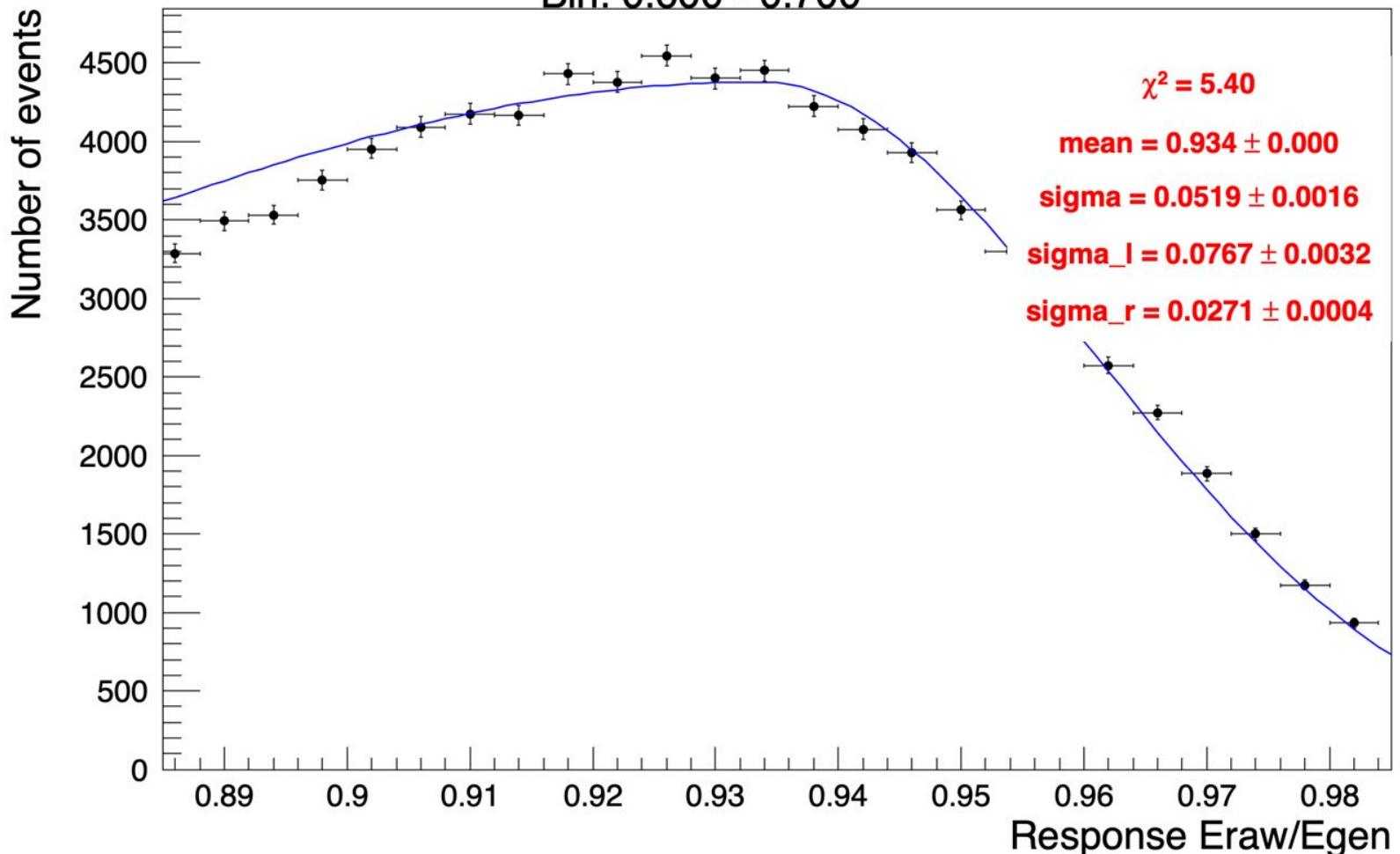




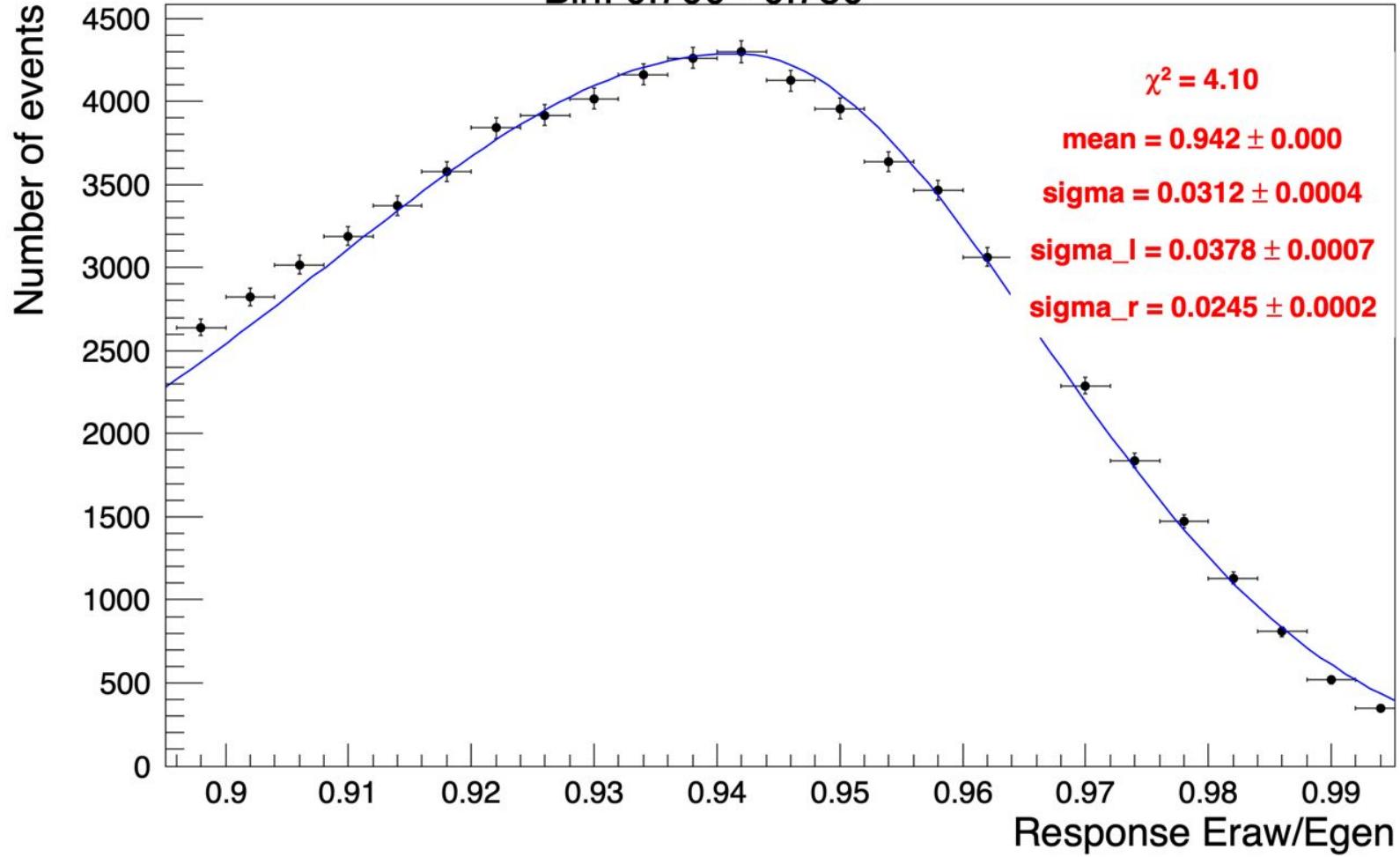


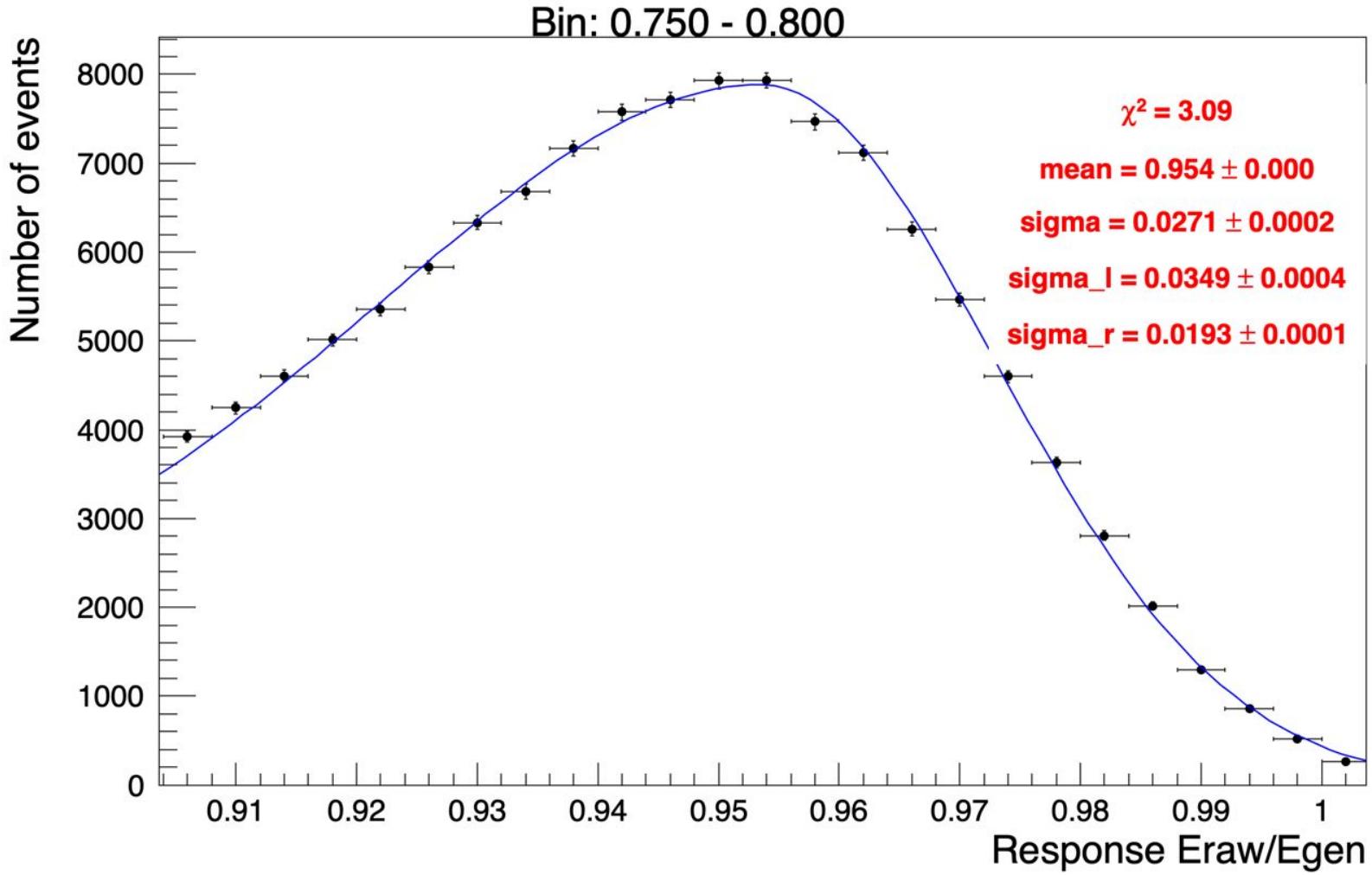


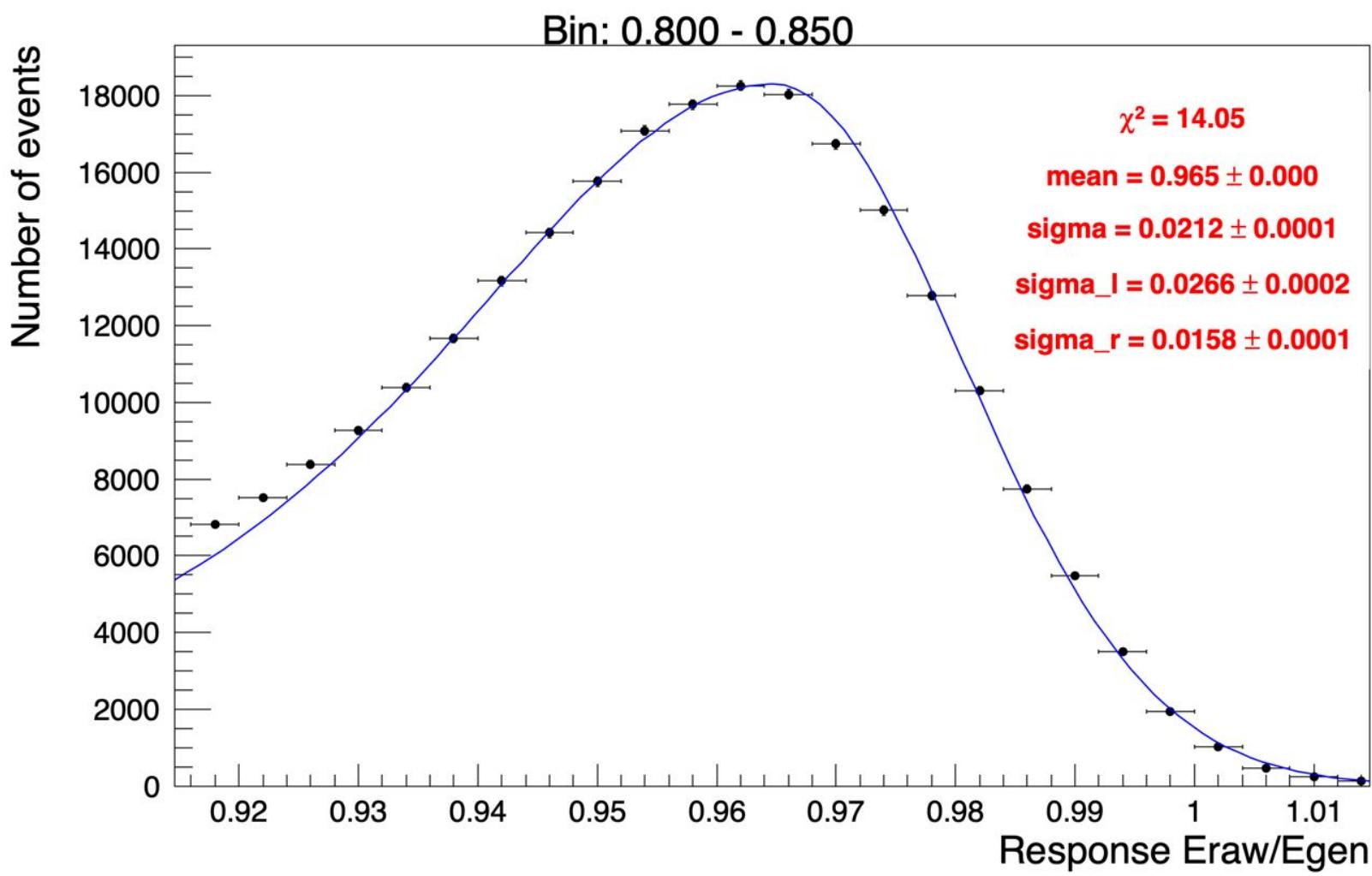
# R9 raw



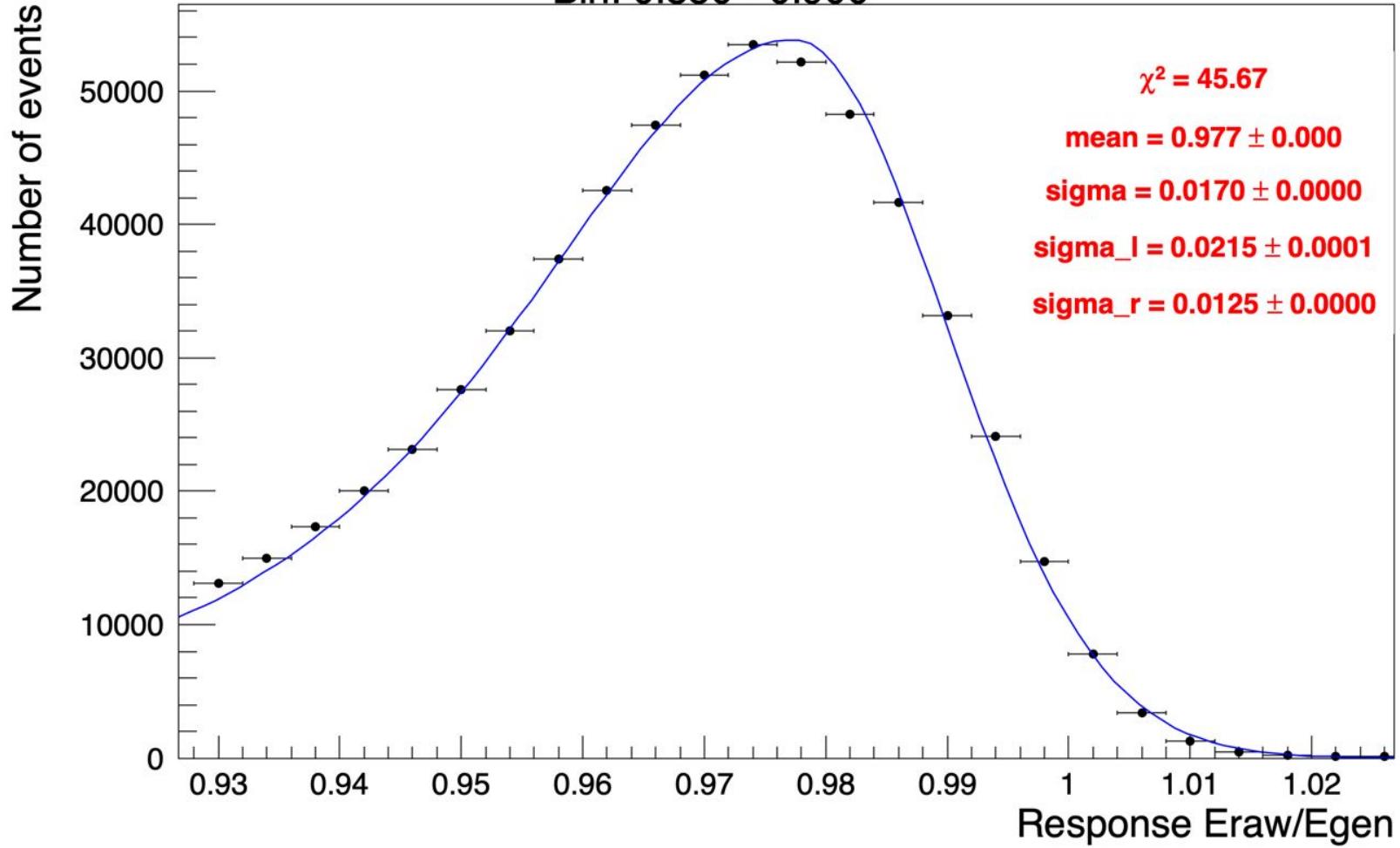
Bin: 0.700 - 0.750

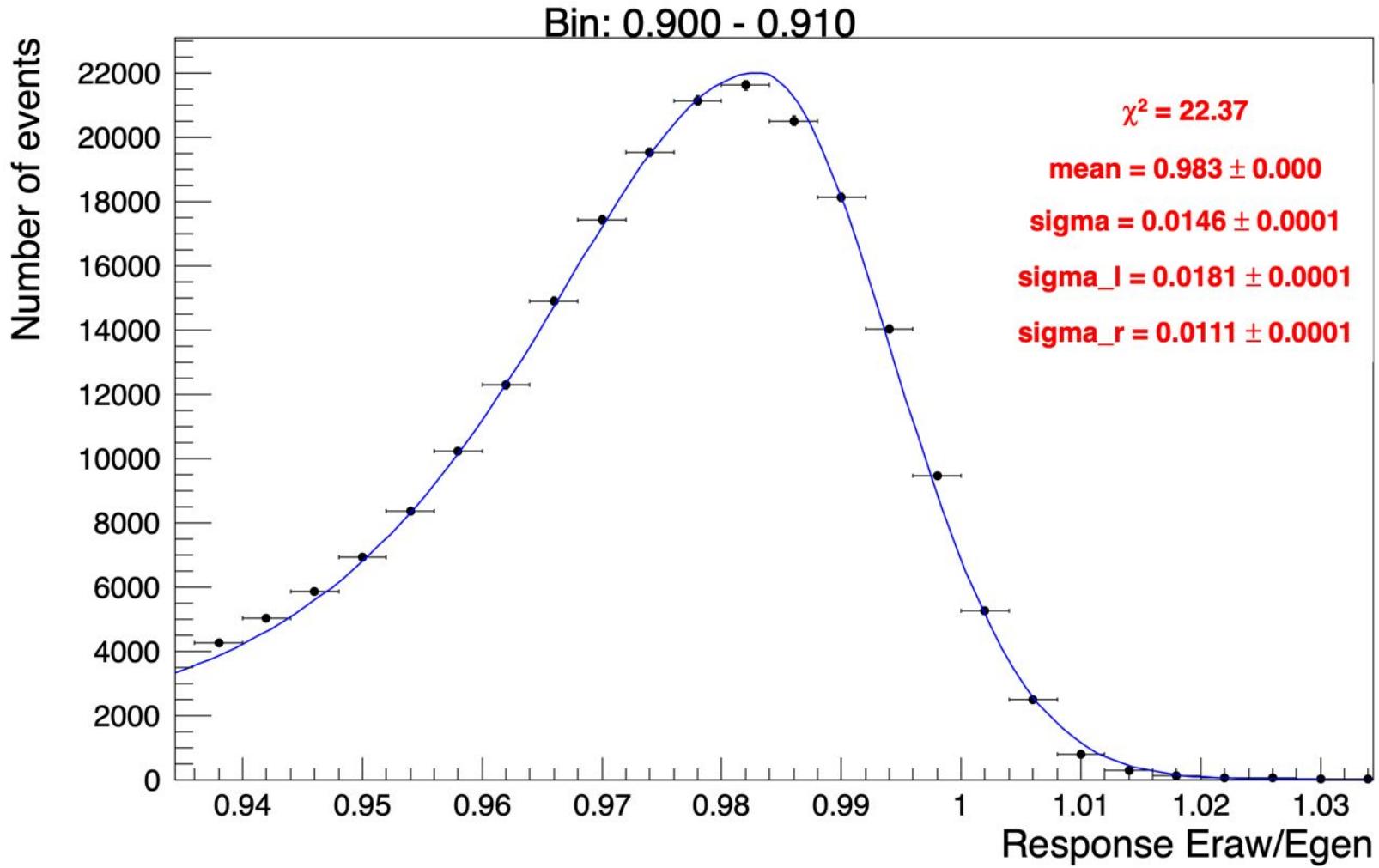


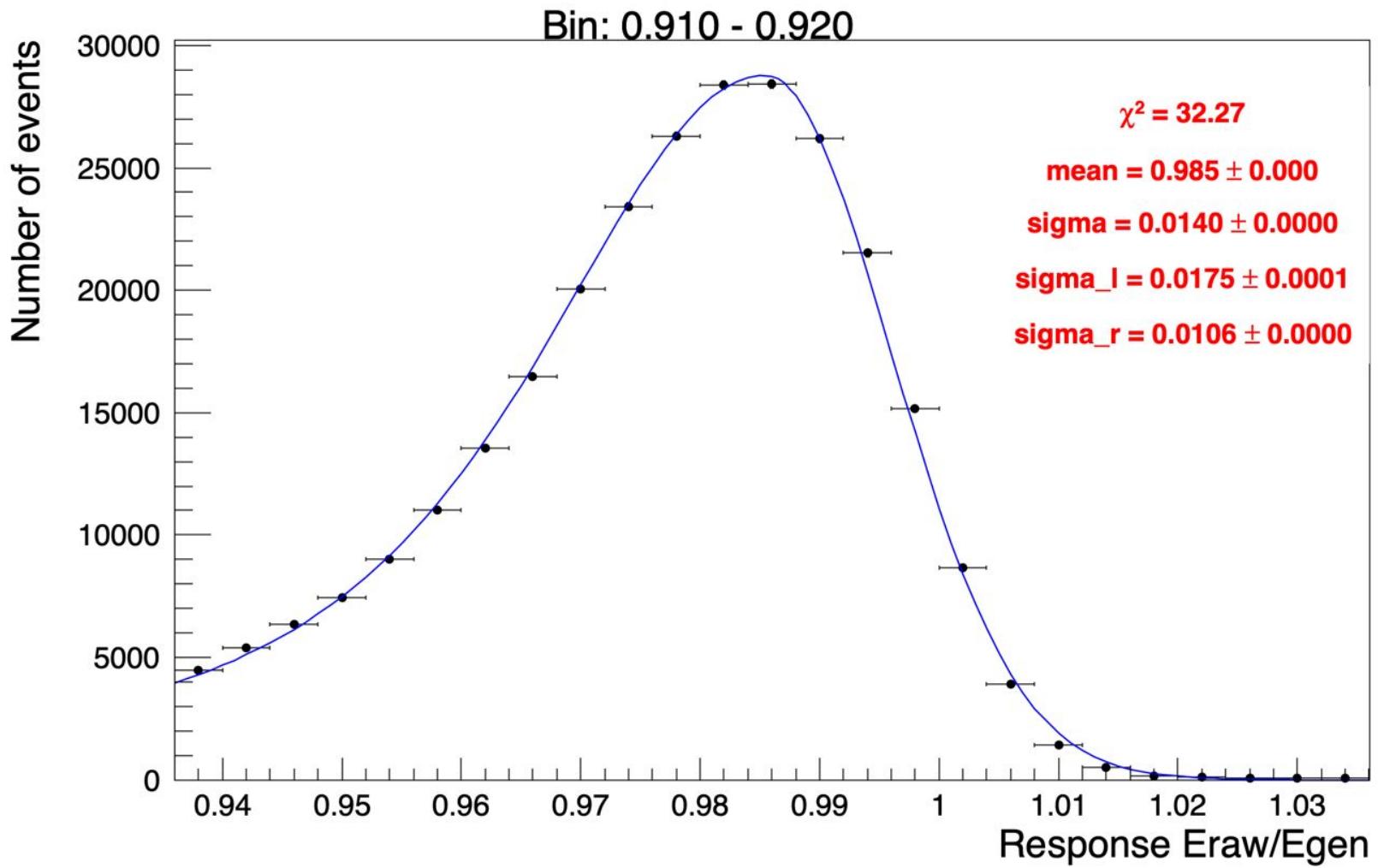




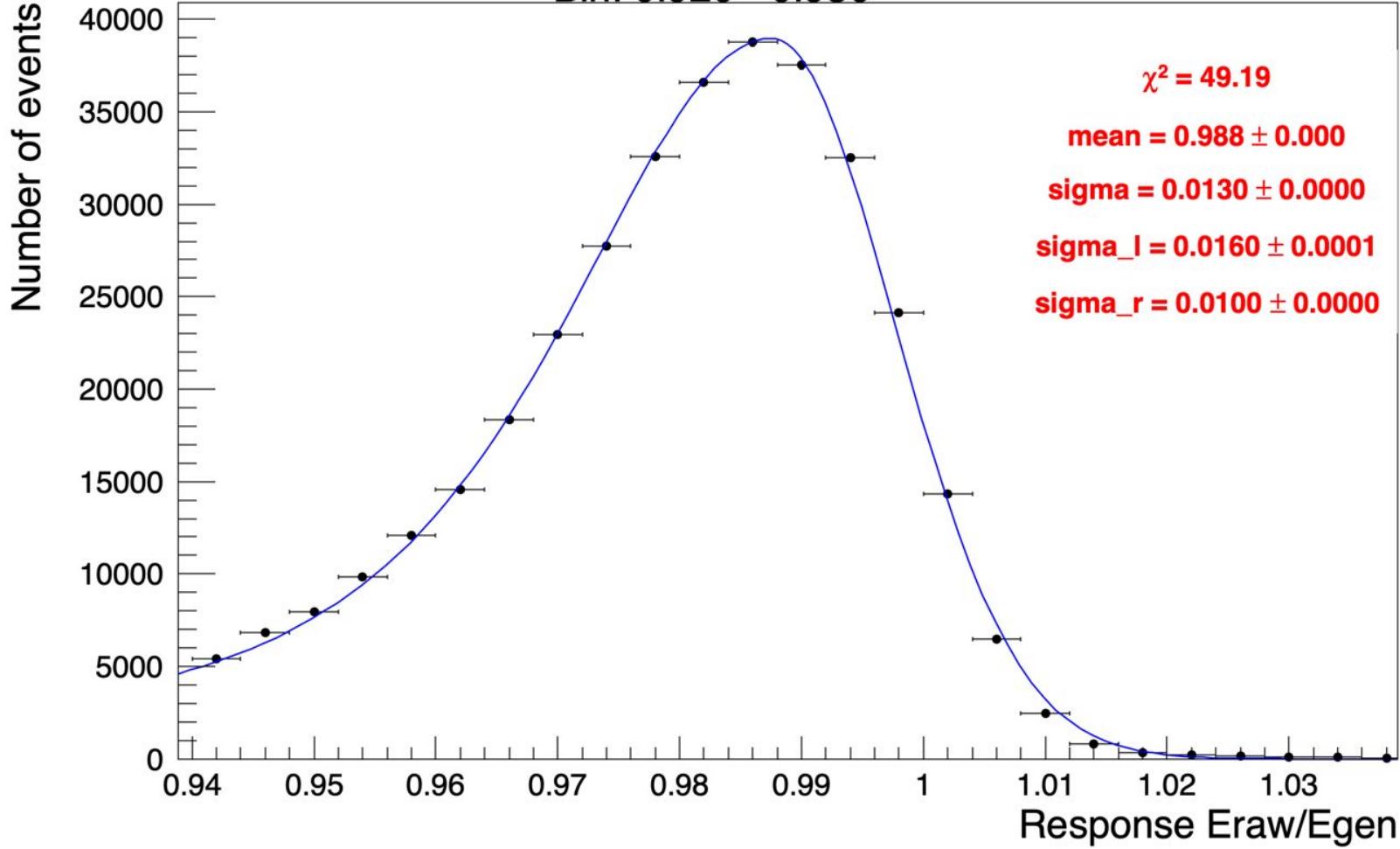
Bin: 0.850 - 0.900



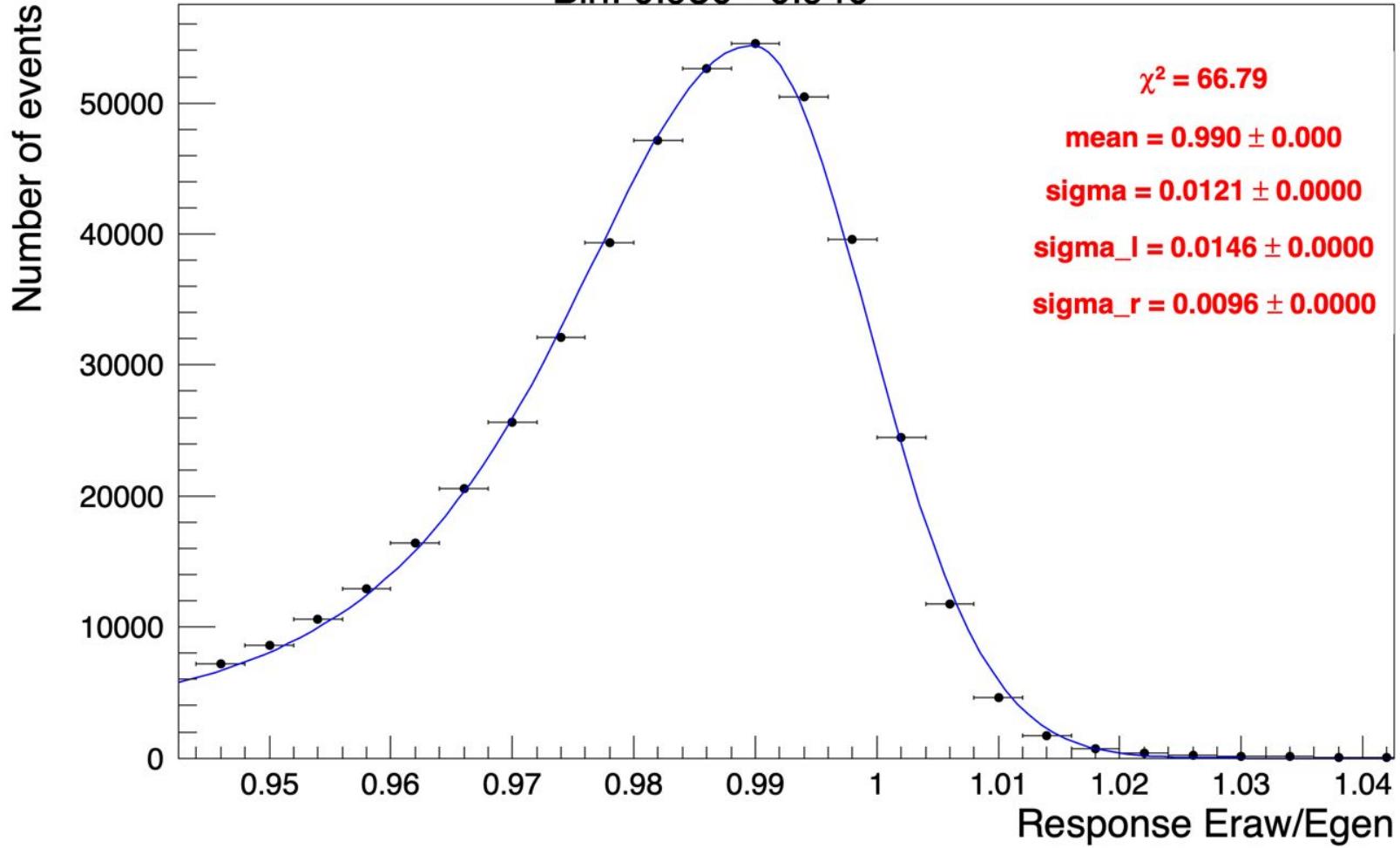


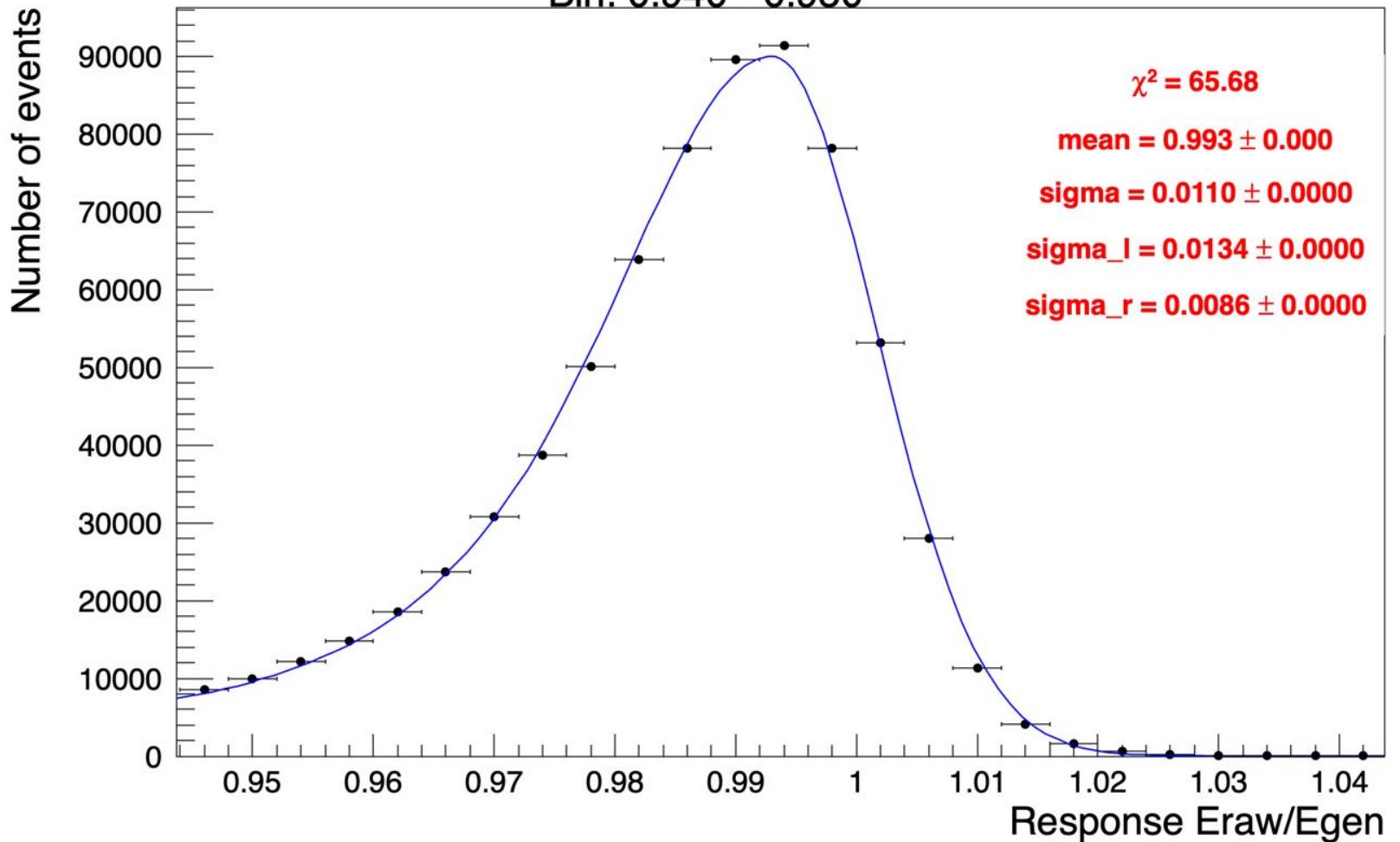


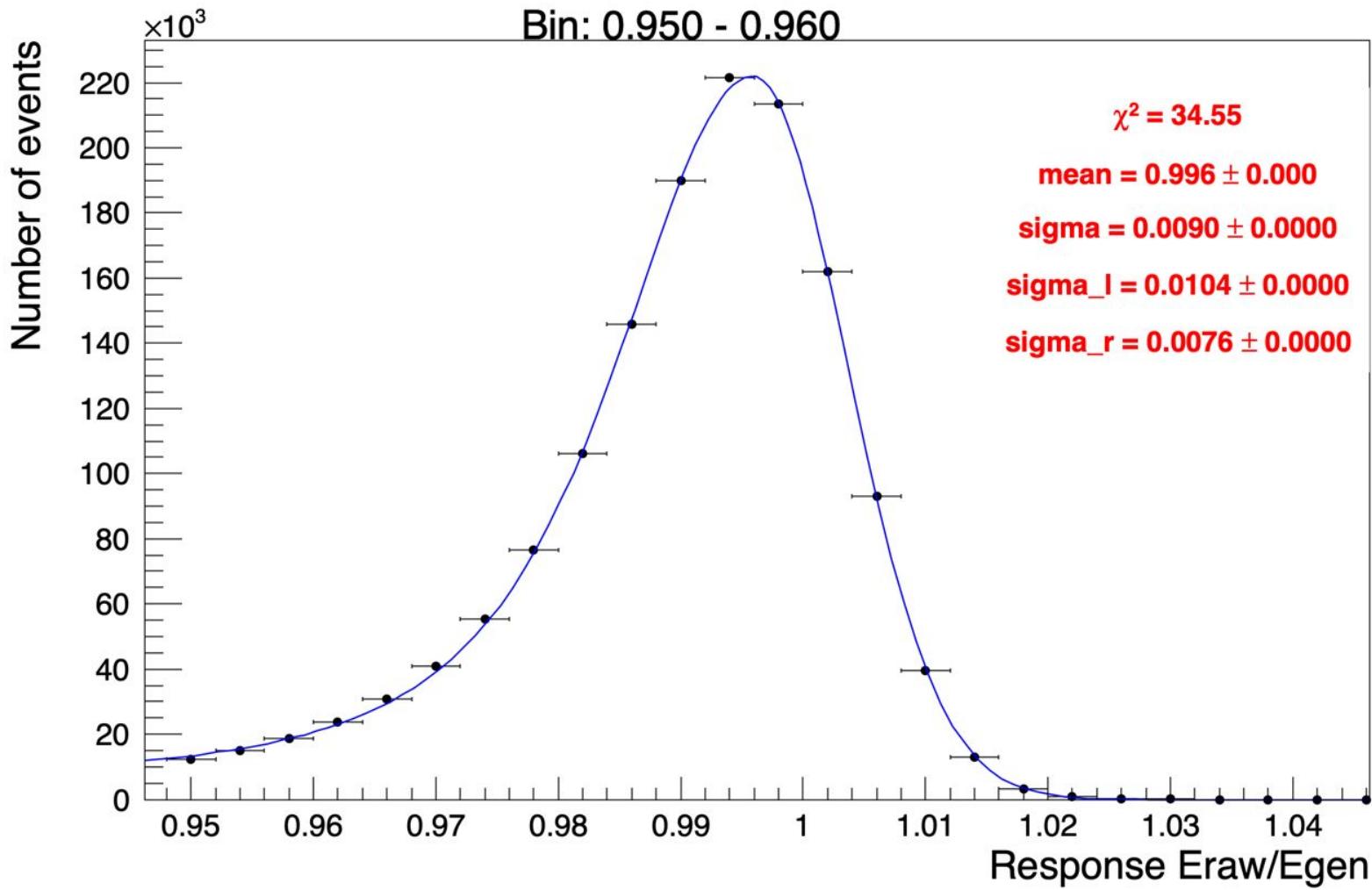
Bin: 0.920 - 0.930

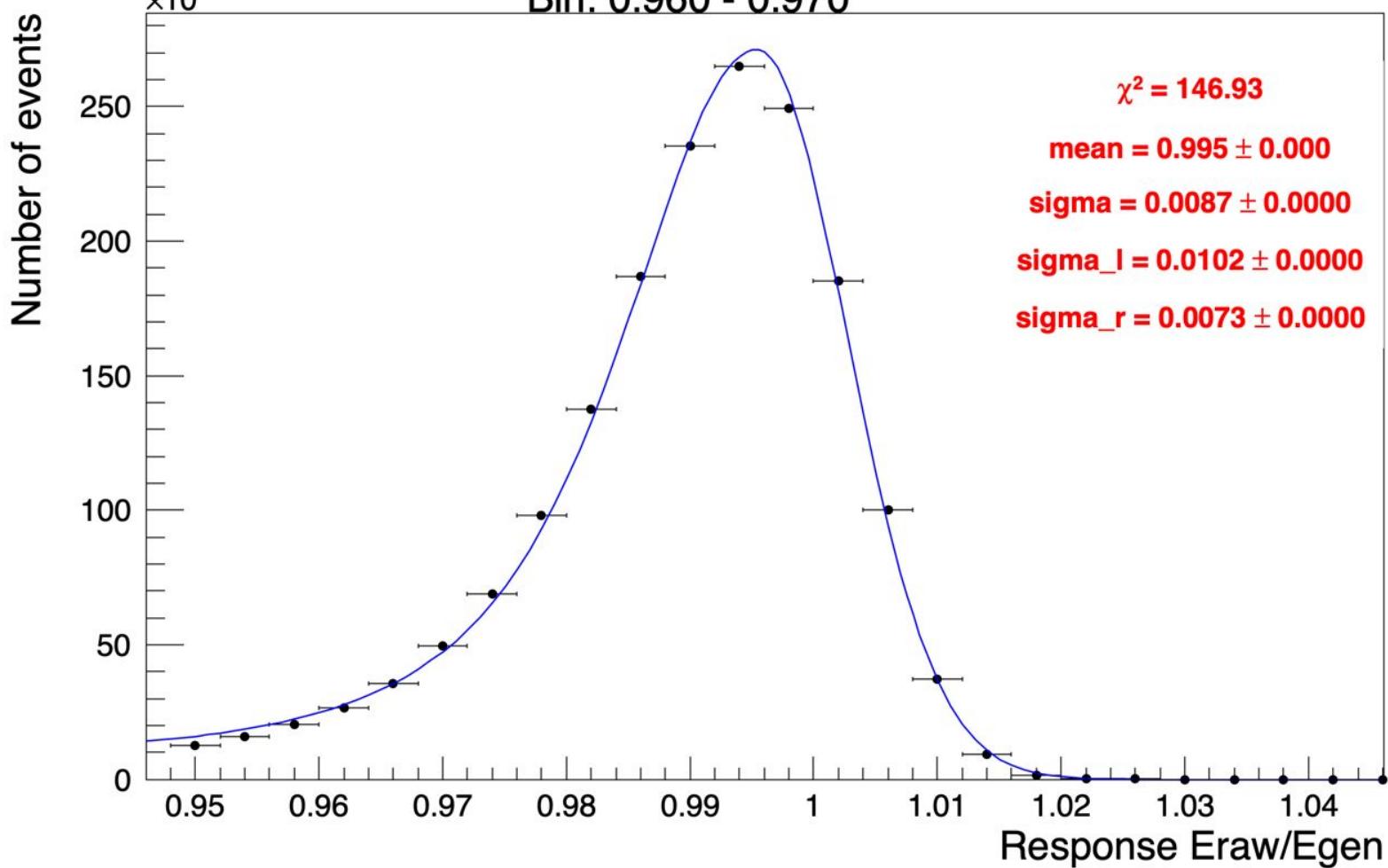


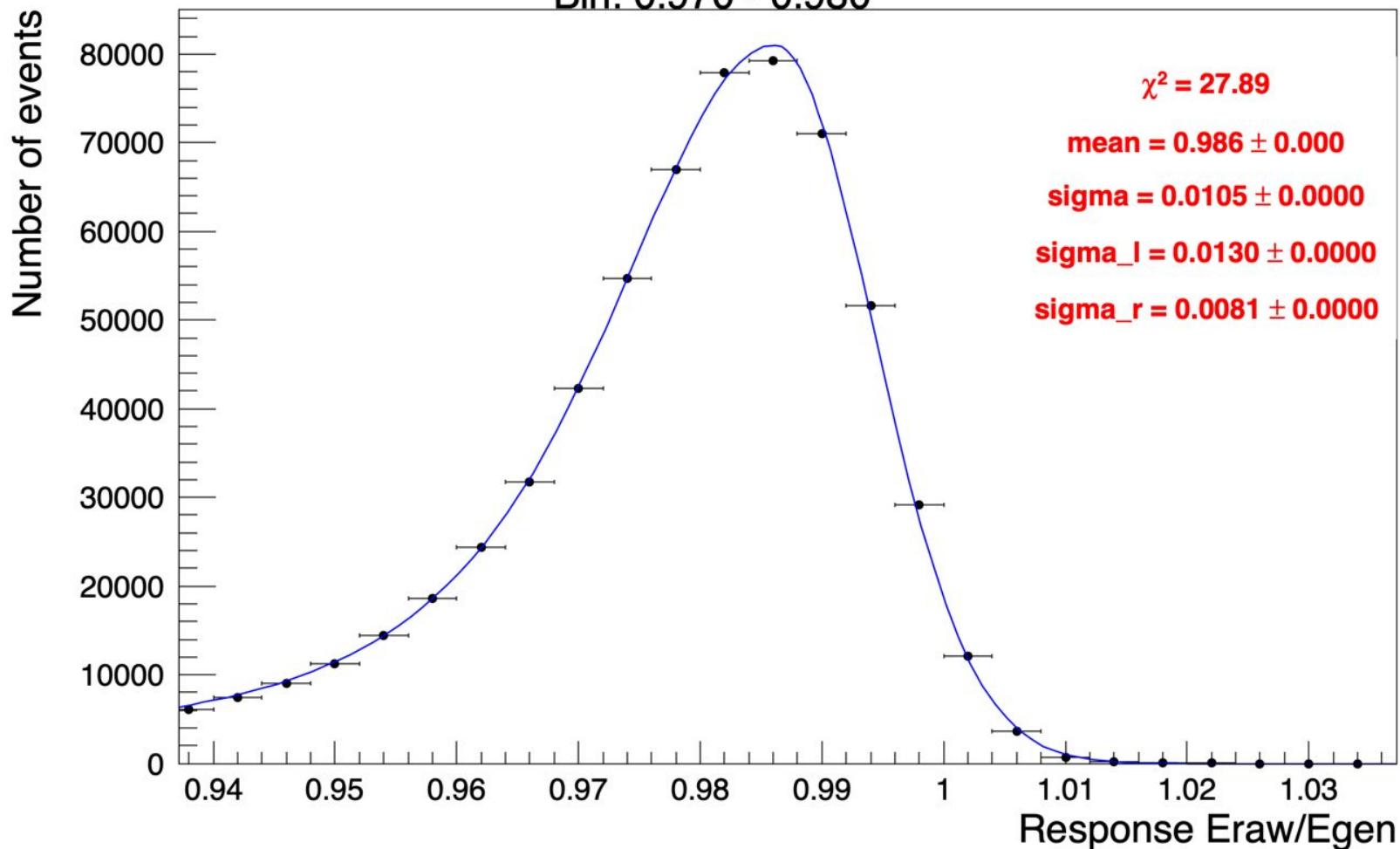
Bin: 0.930 - 0.940





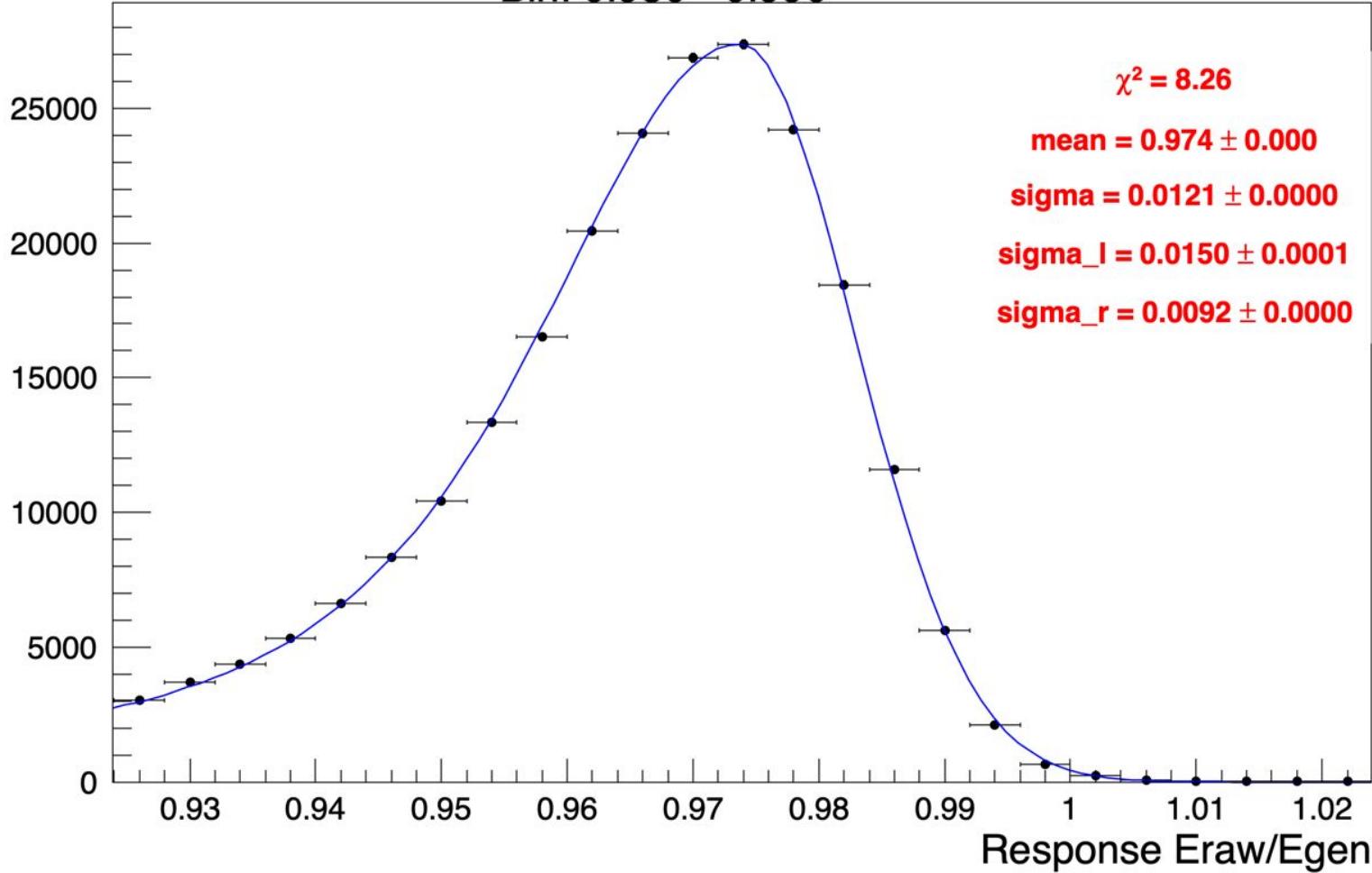


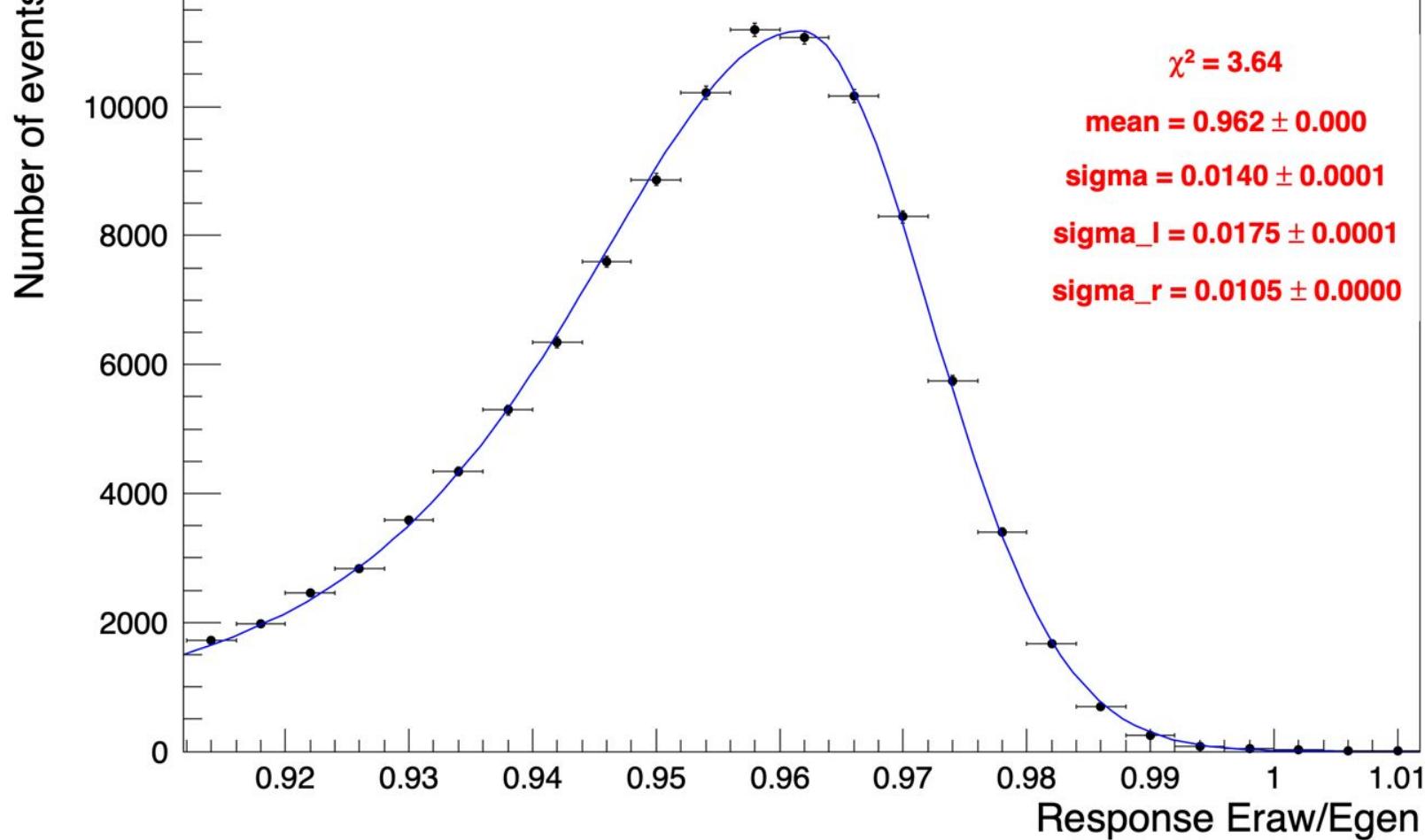




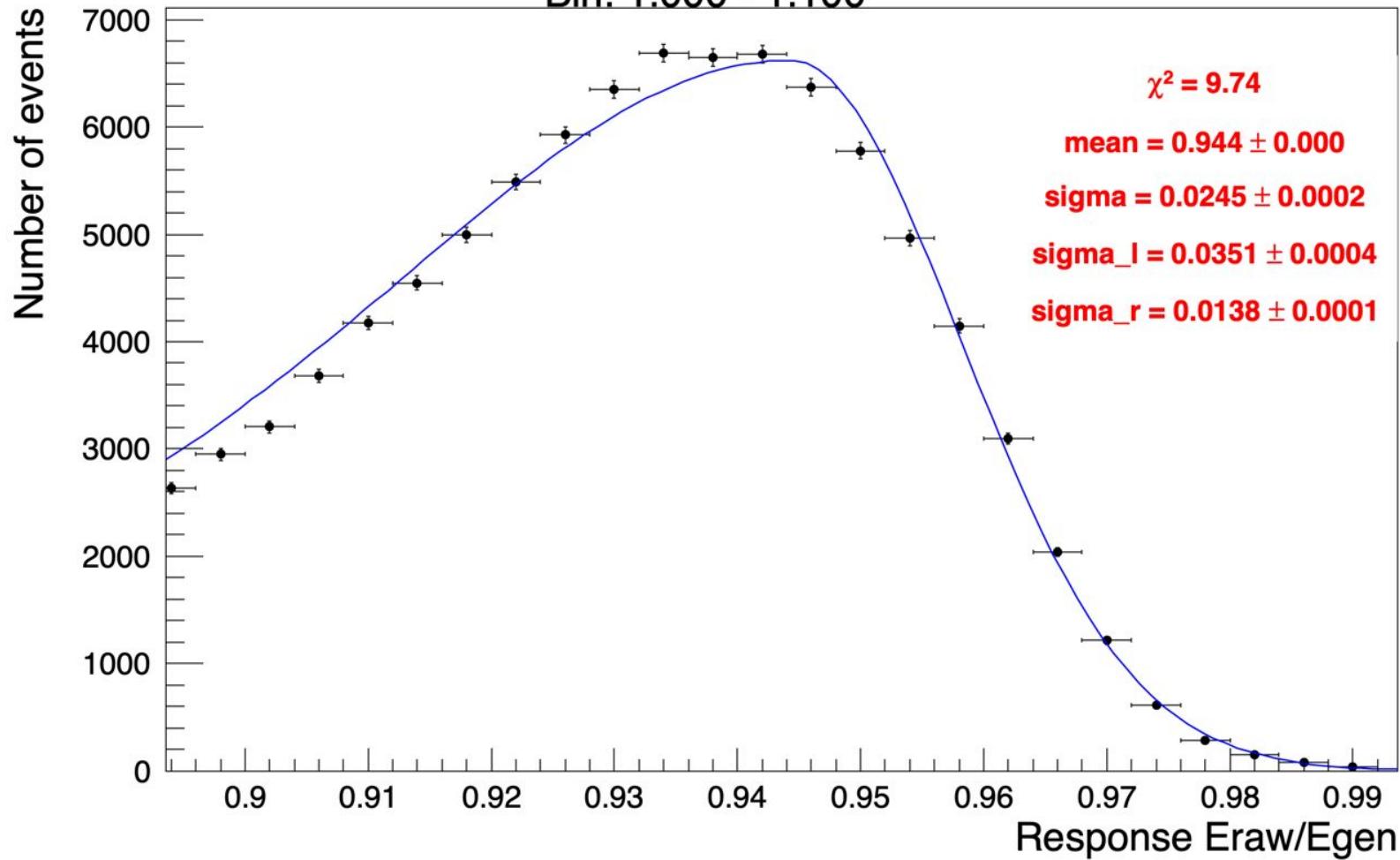
Bin: 0.980 - 0.990

Number of events



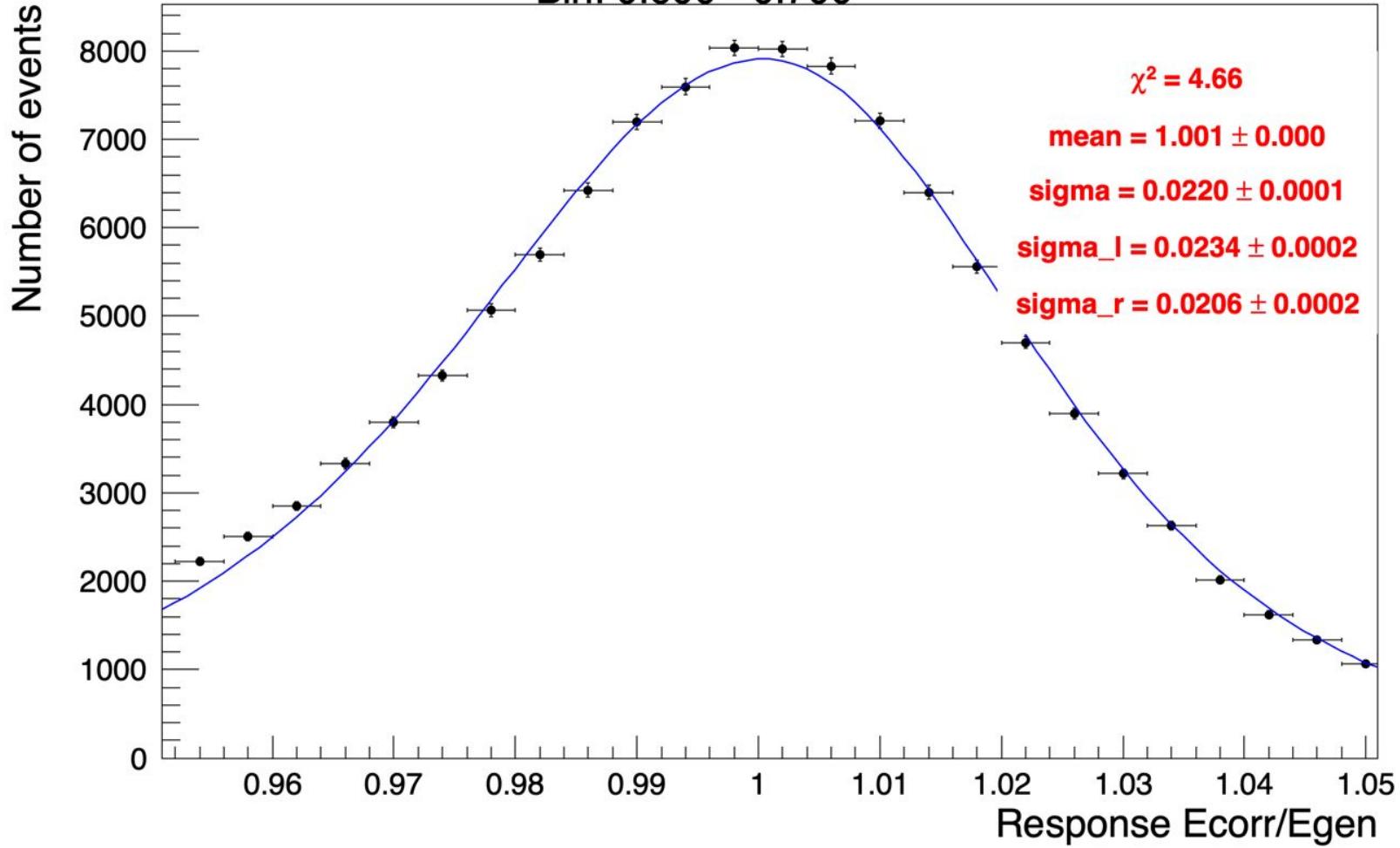


Bin: 1.000 - 1.100

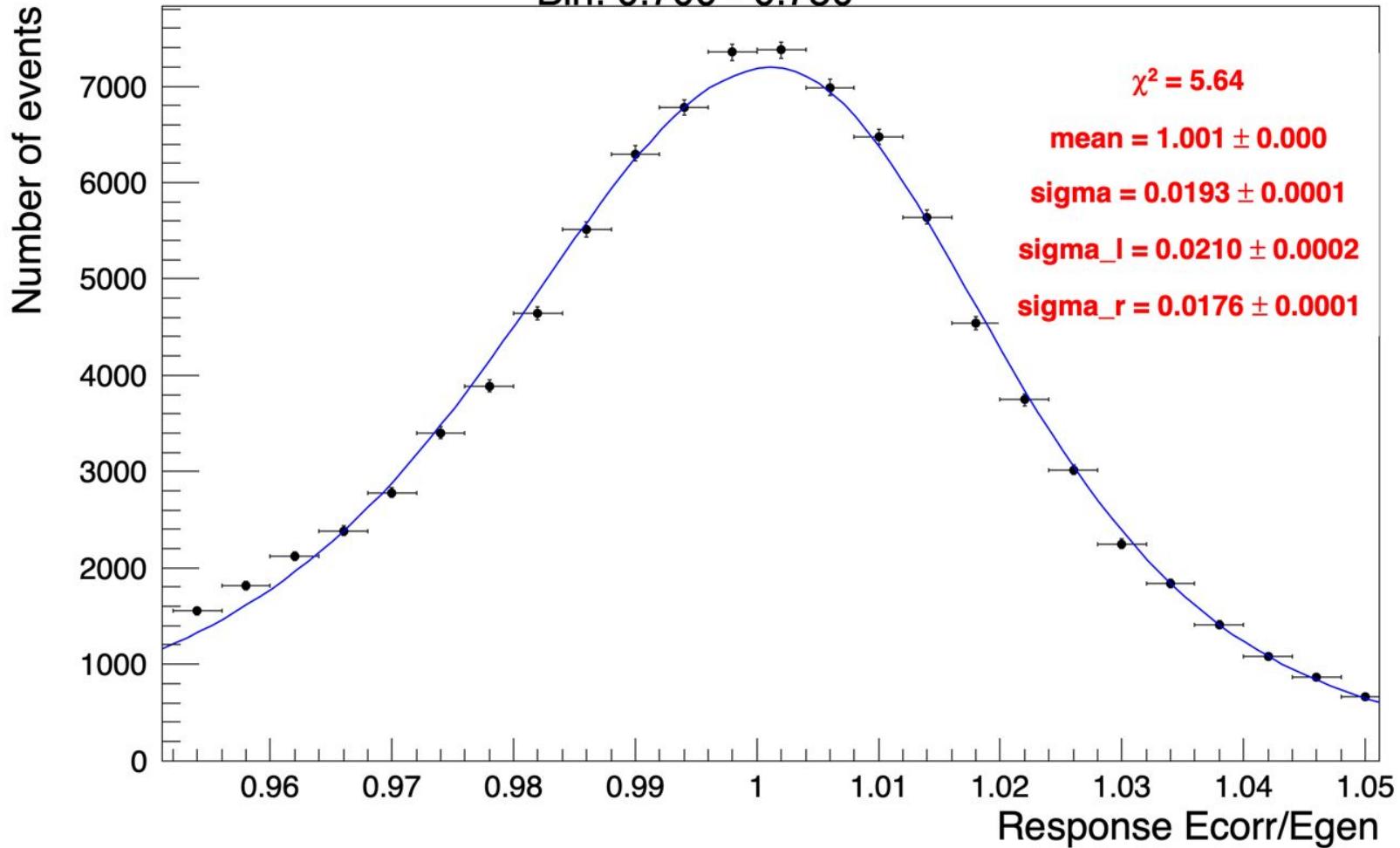


# R9 corrected

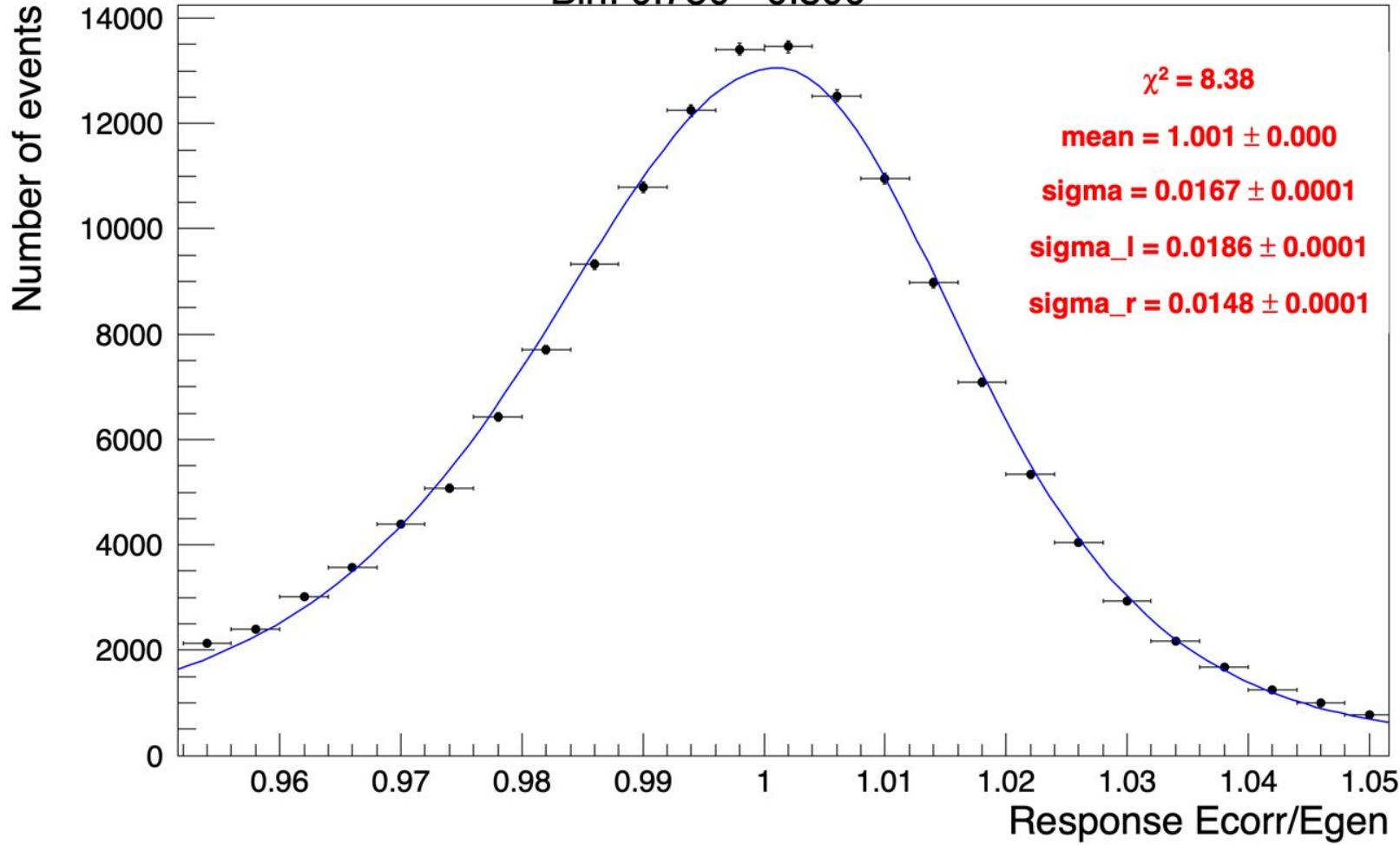
Bin: 0.600 - 0.700



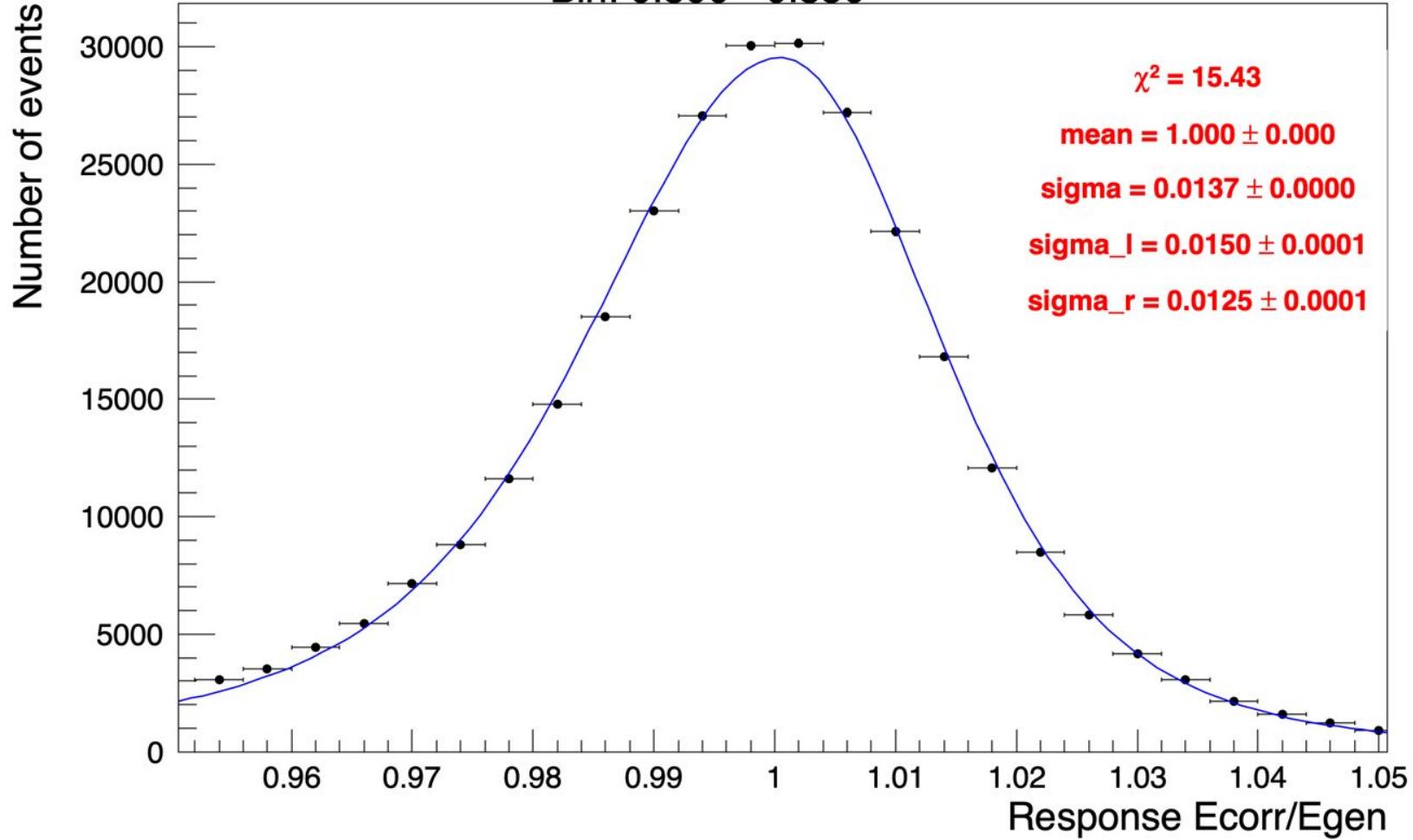
Bin: 0.700 - 0.750

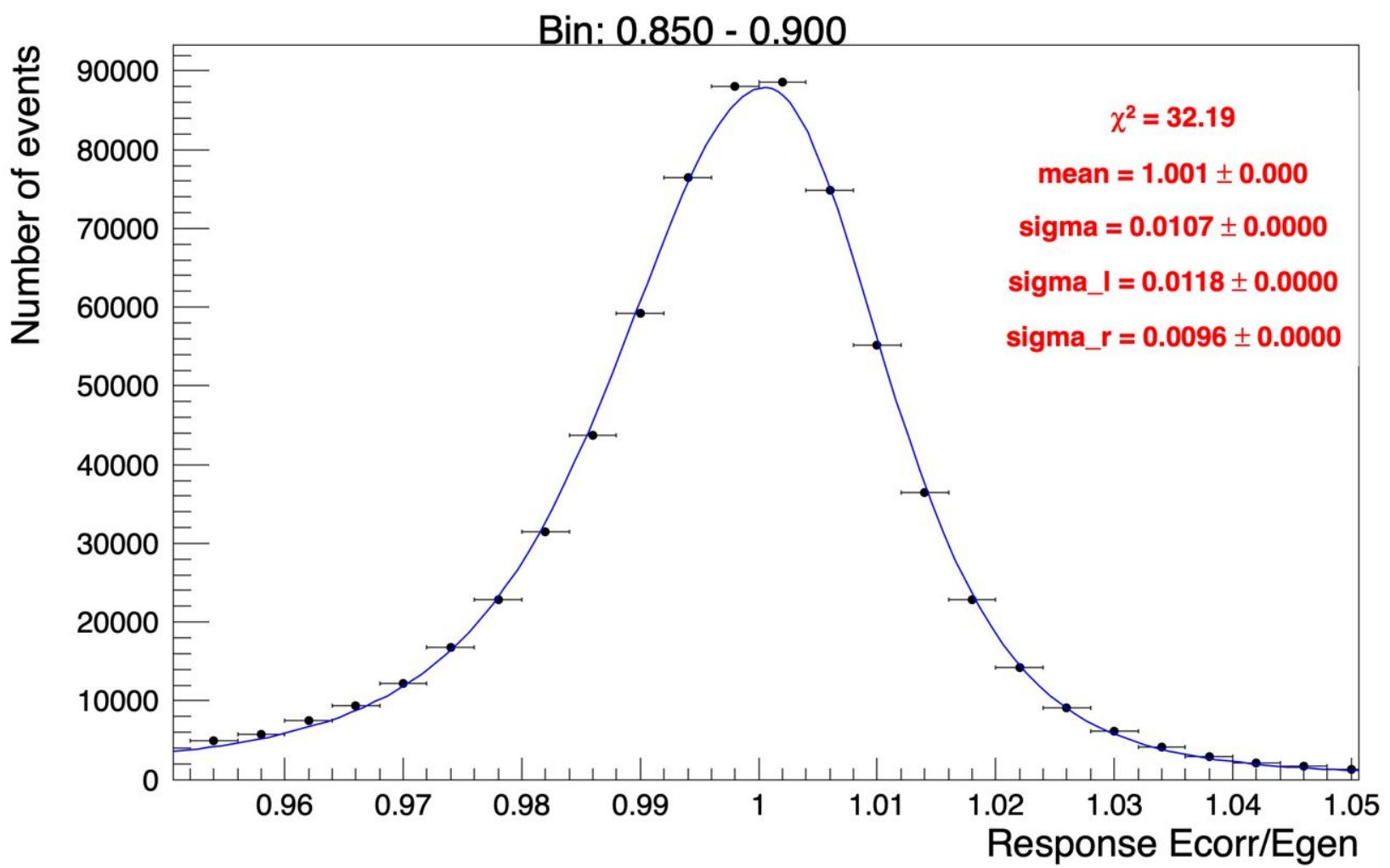


Bin: 0.750 - 0.800



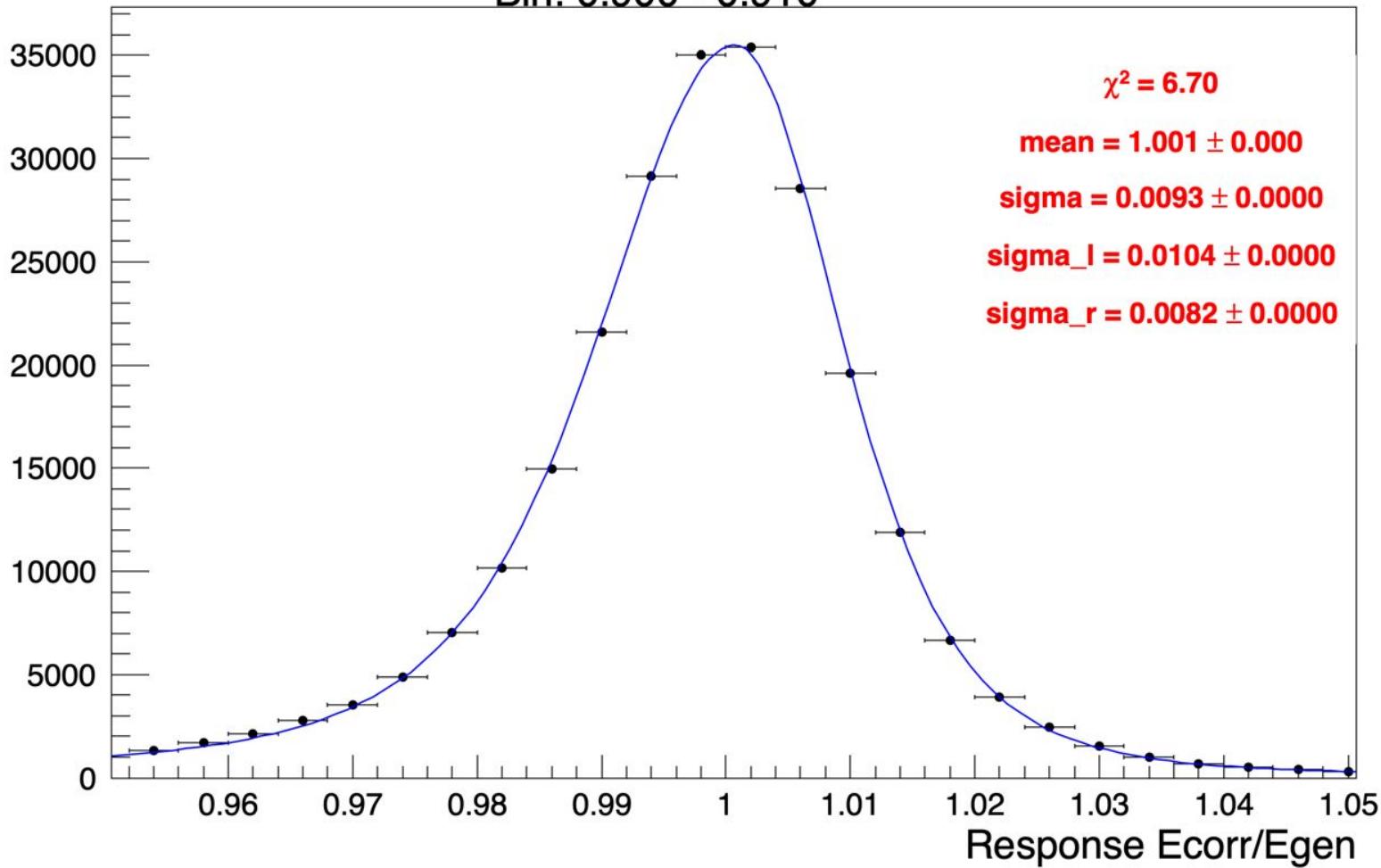
Bin: 0.800 - 0.850



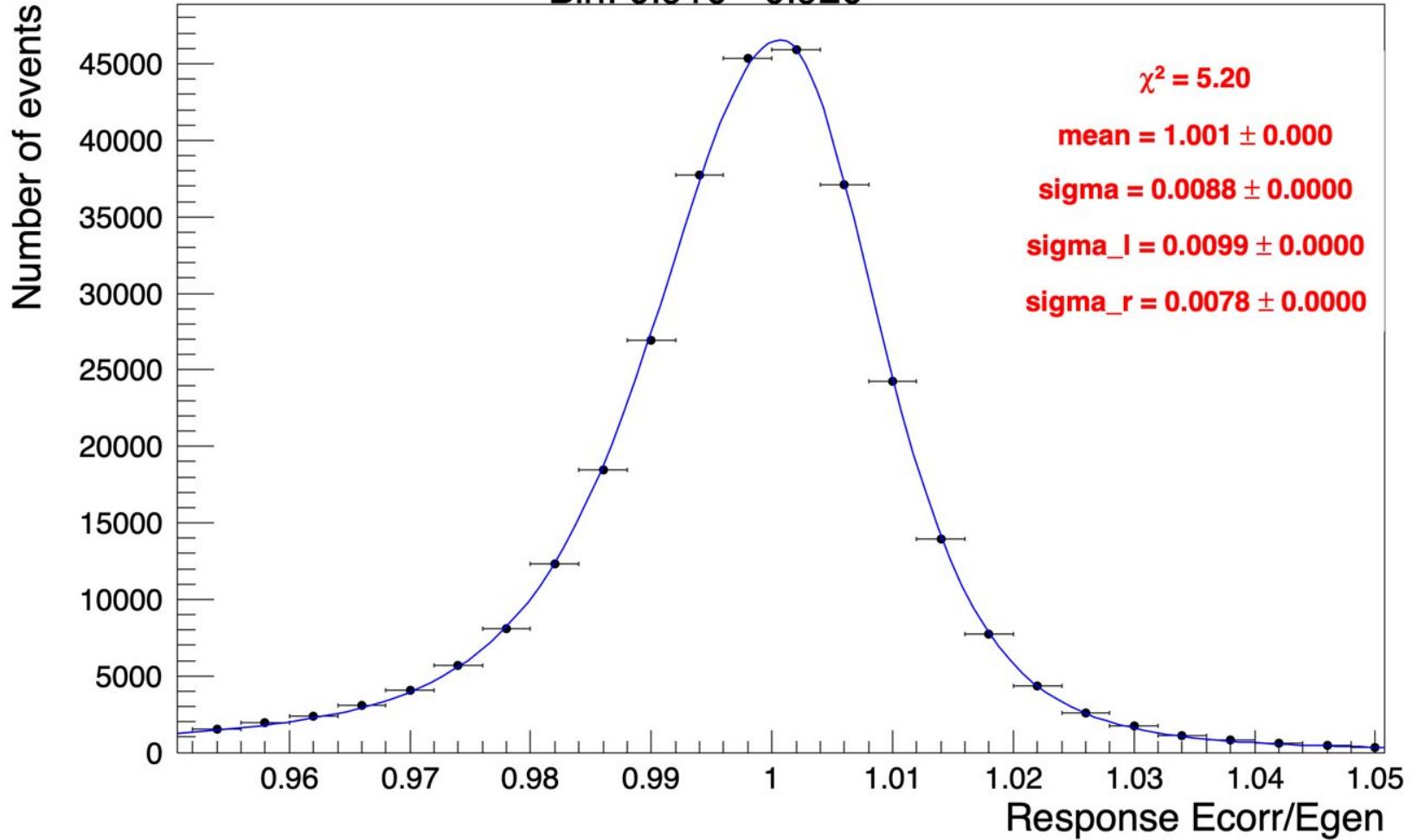


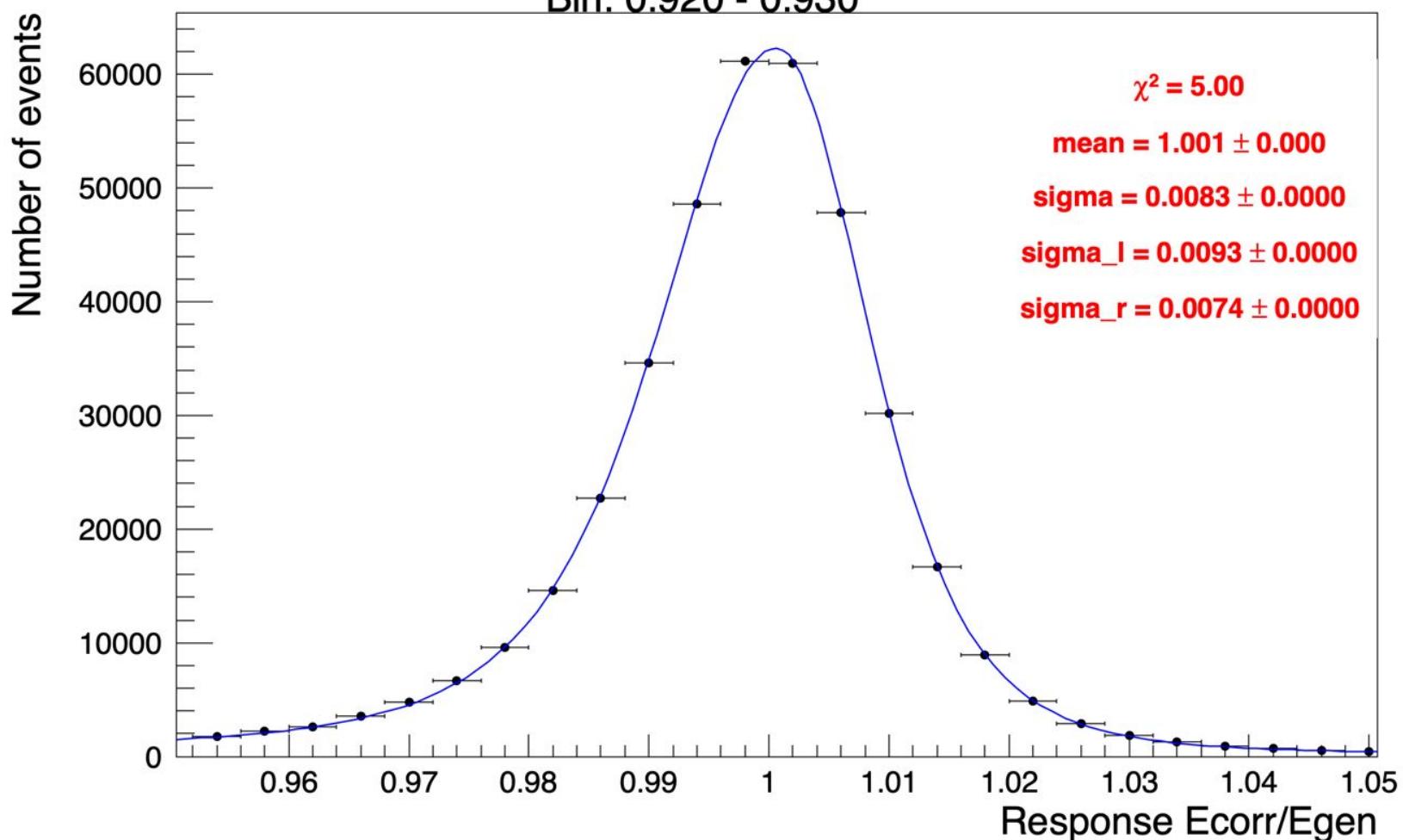
Number of events

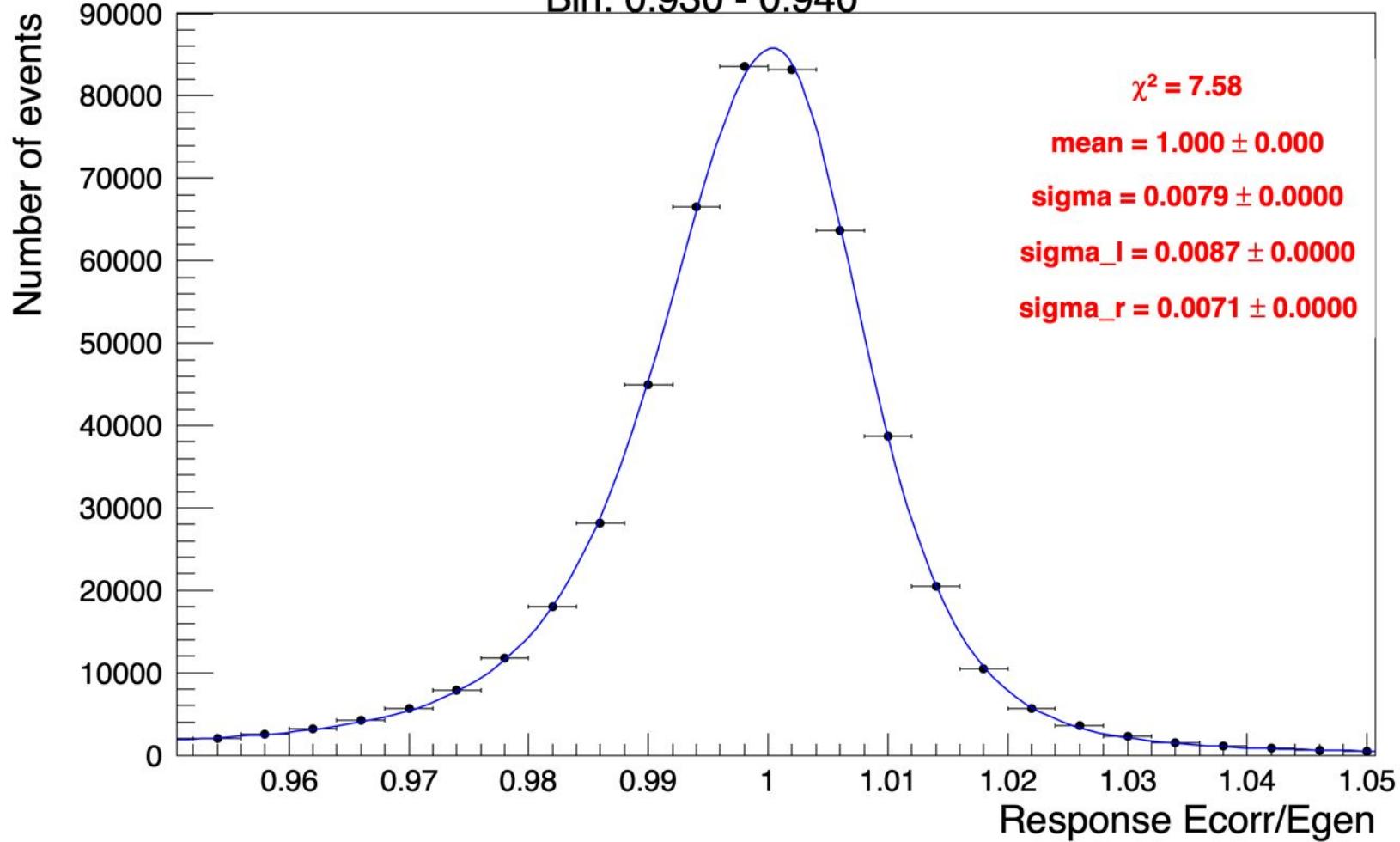
Bin: 0.900 - 0.910

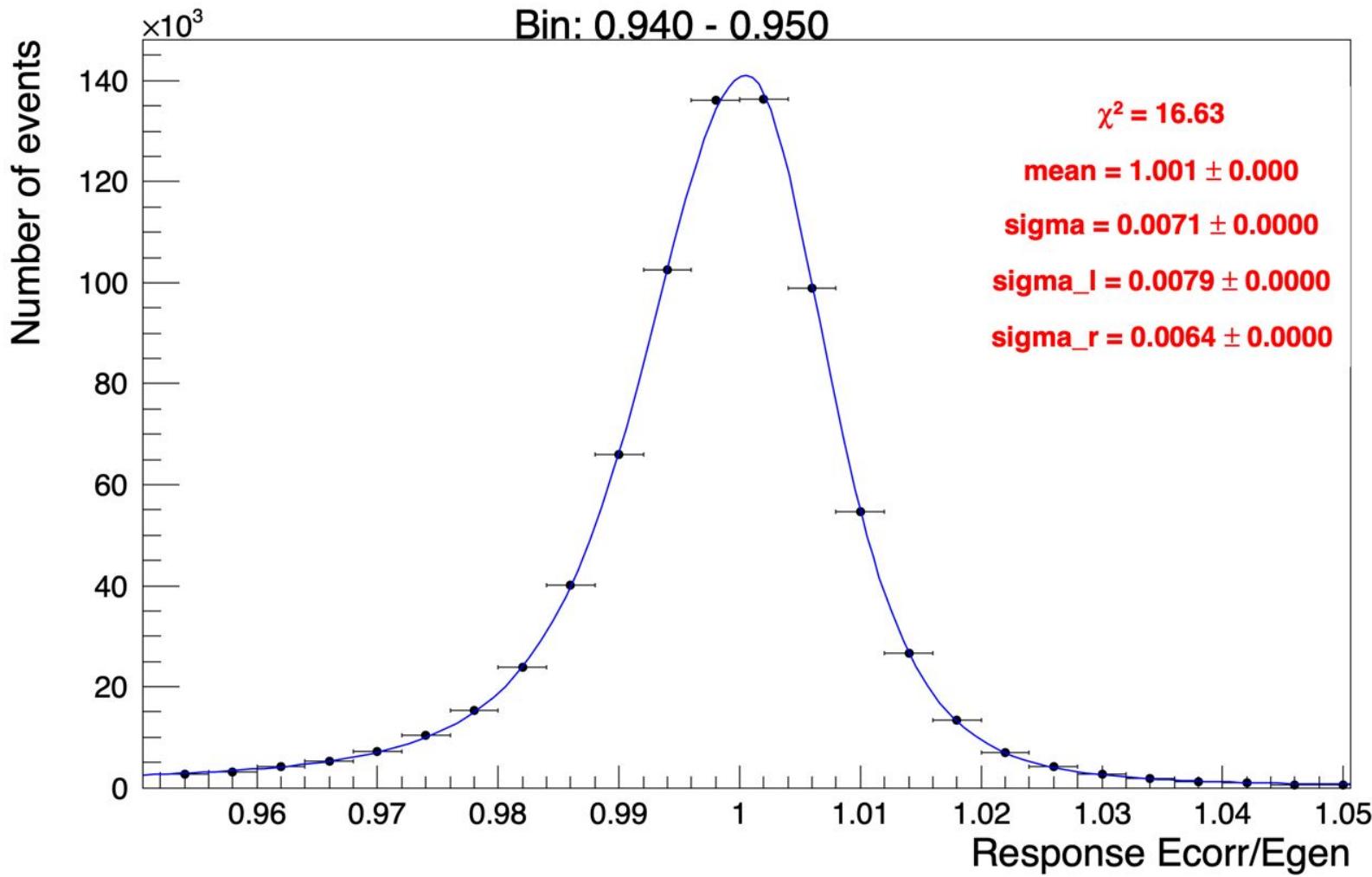


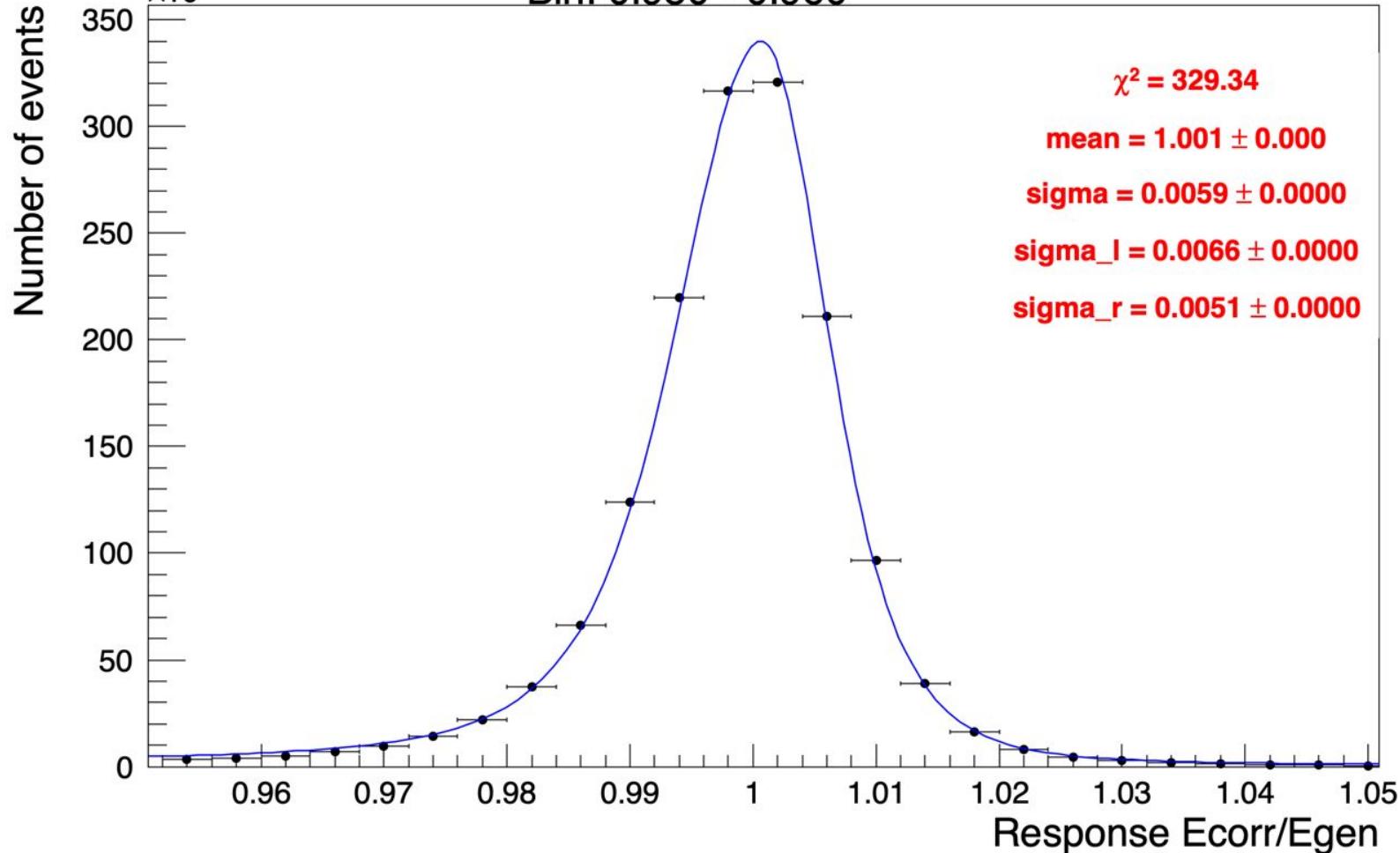
Bin: 0.910 - 0.920

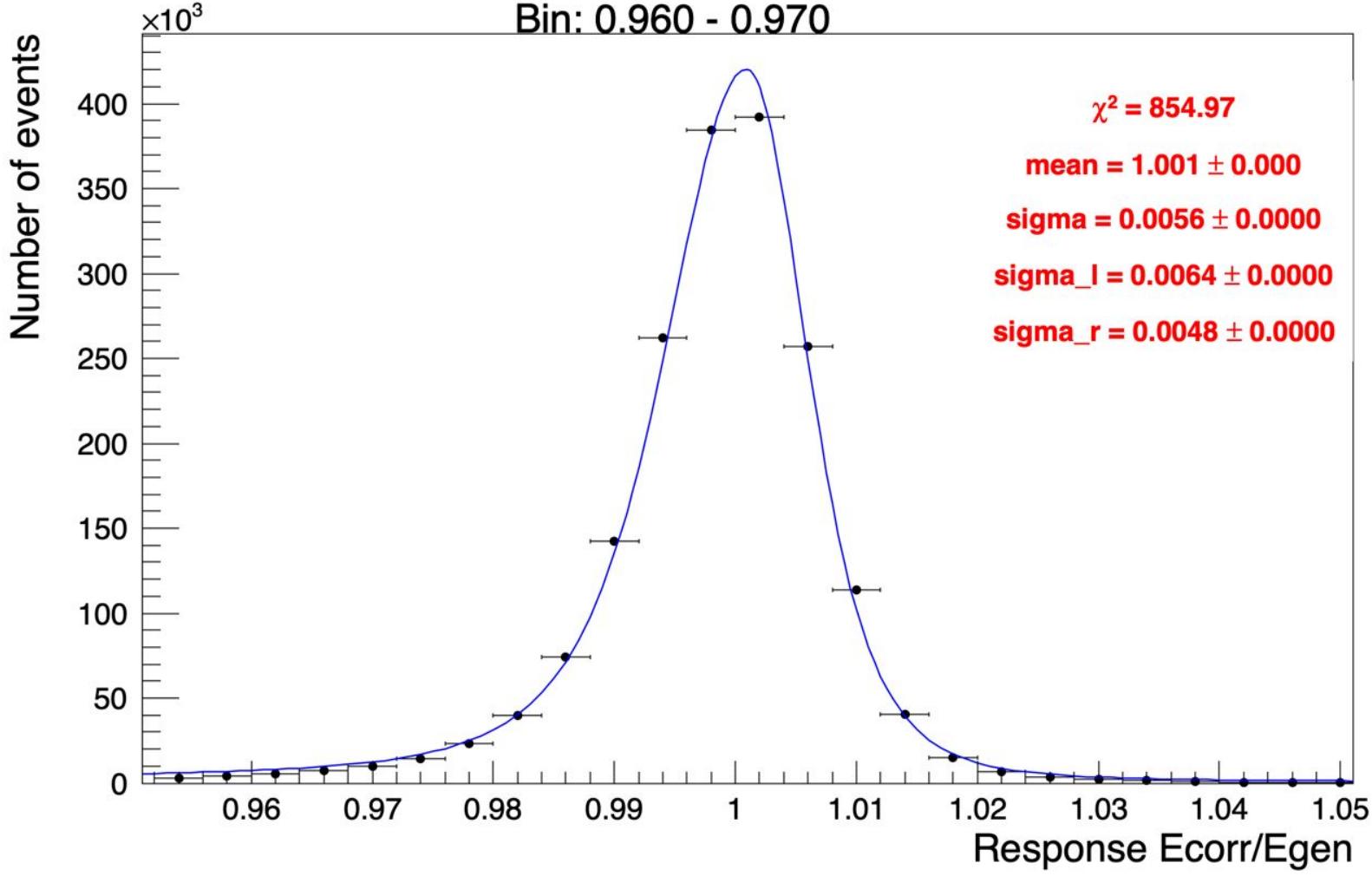


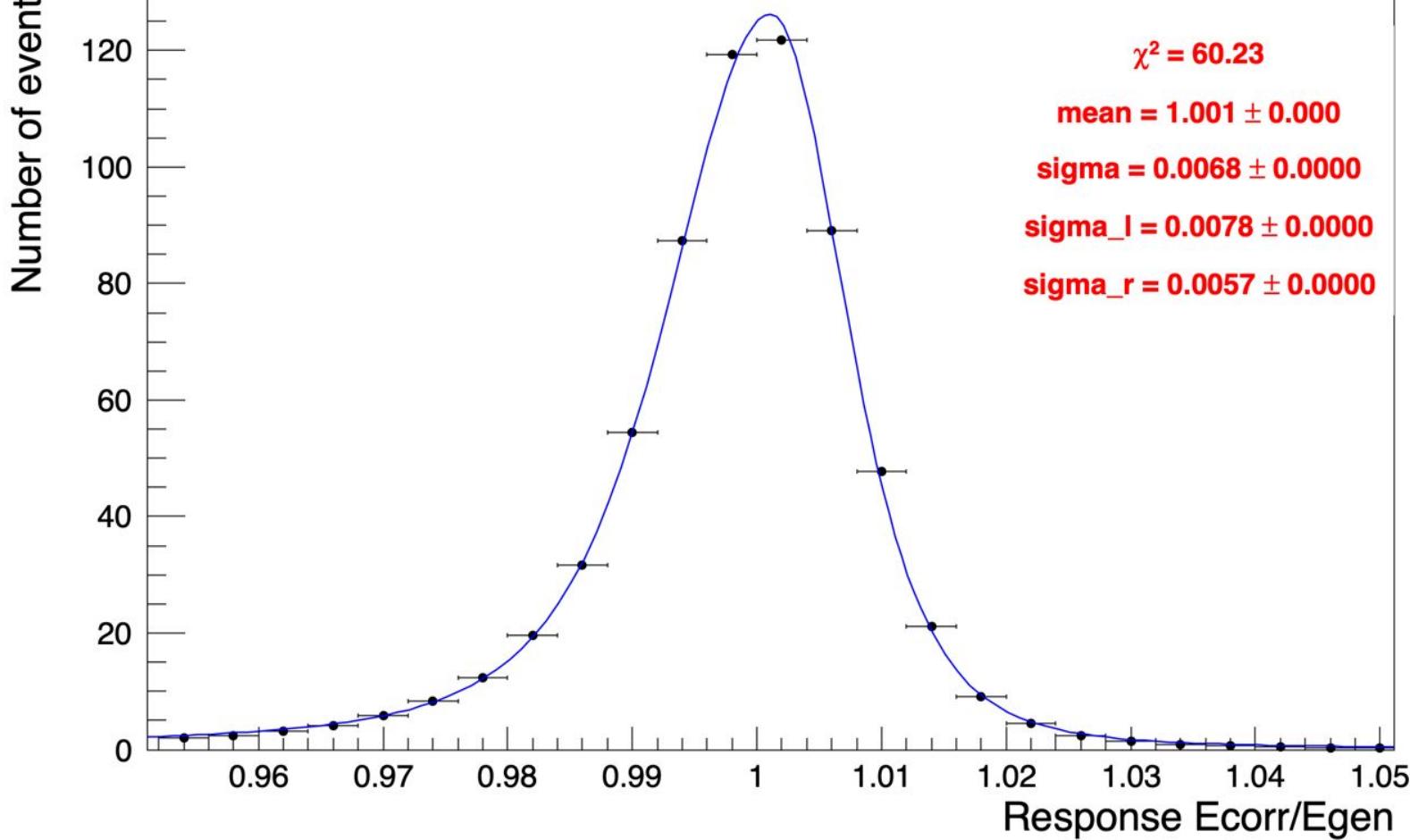




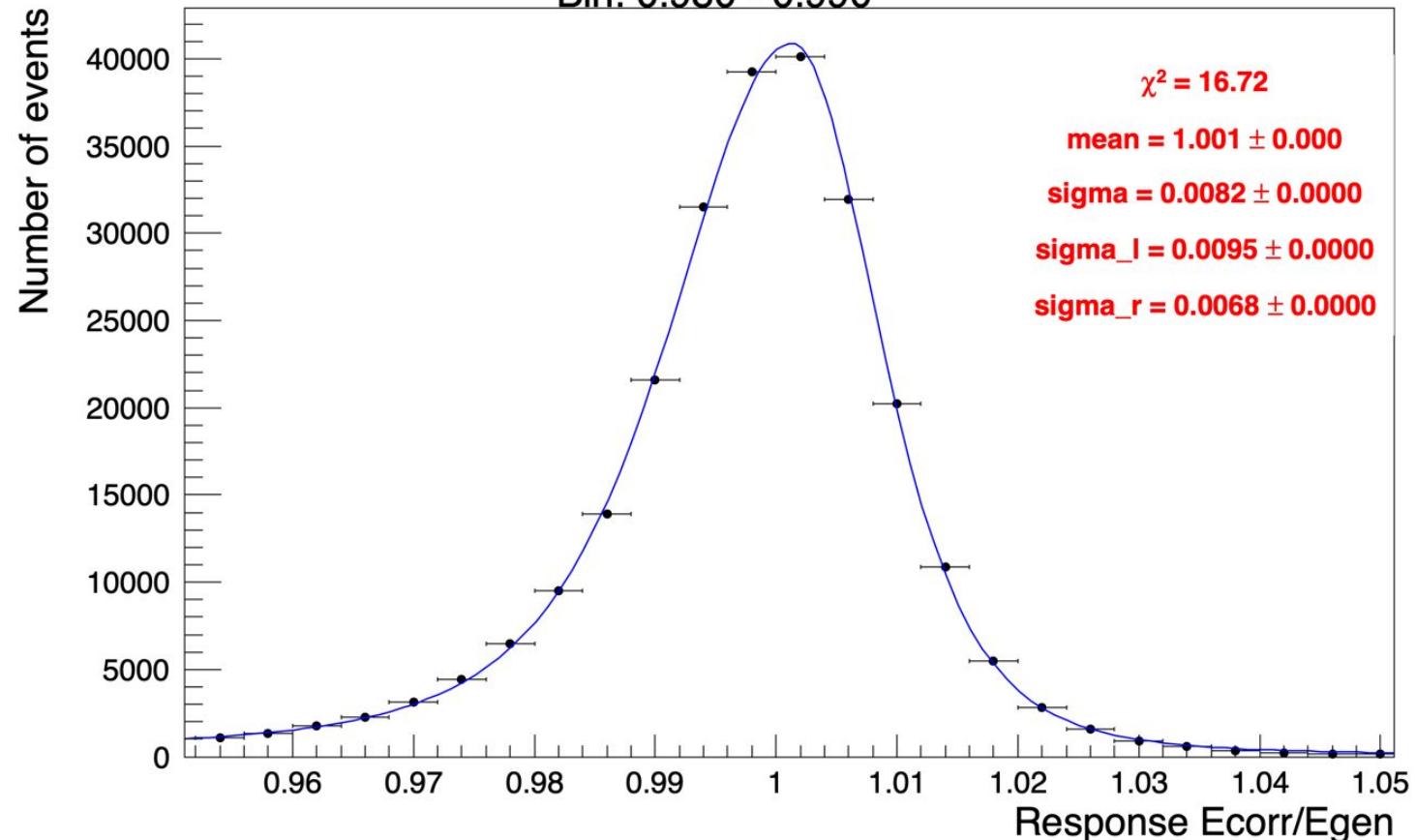




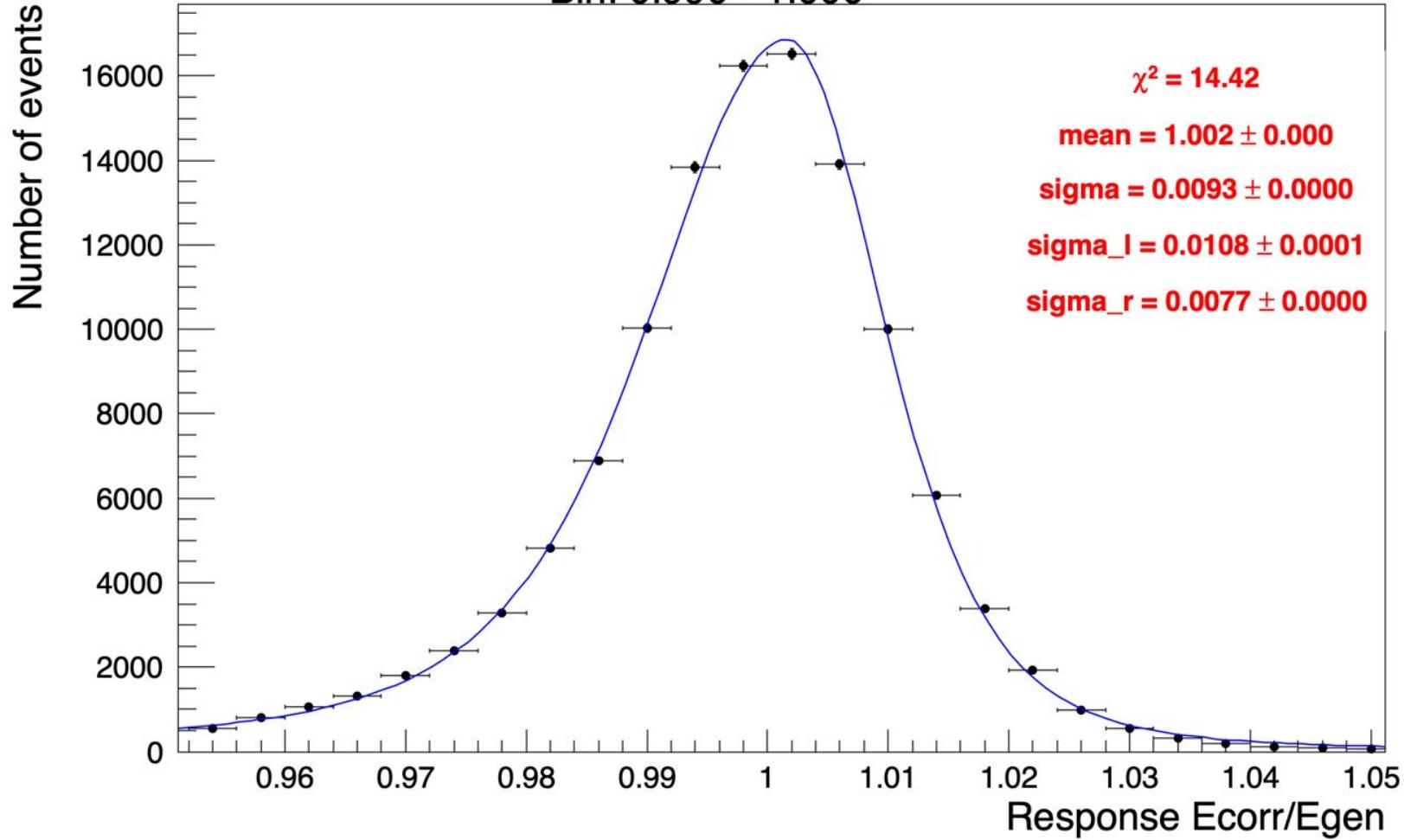


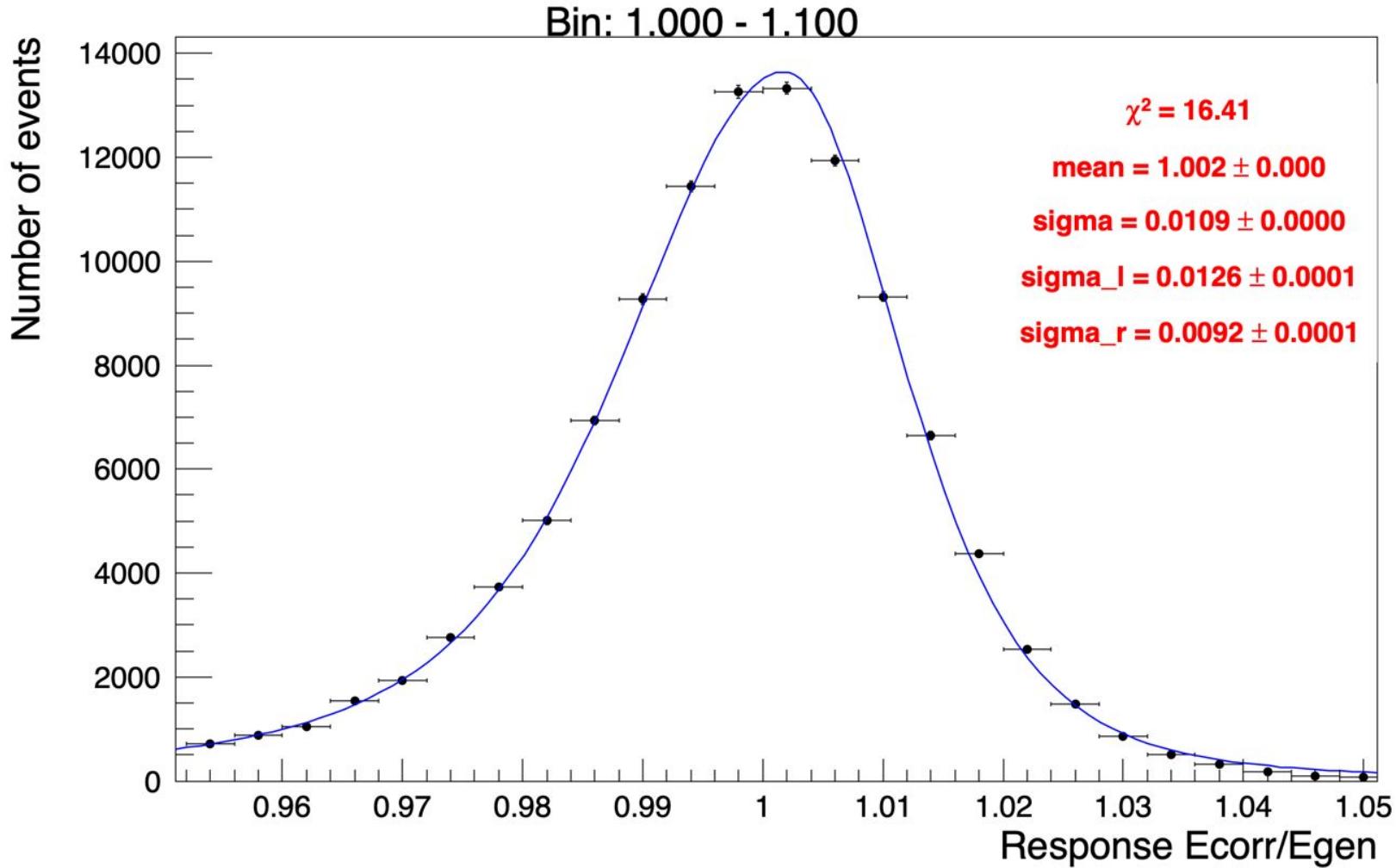


Bin: 0.980 - 0.990



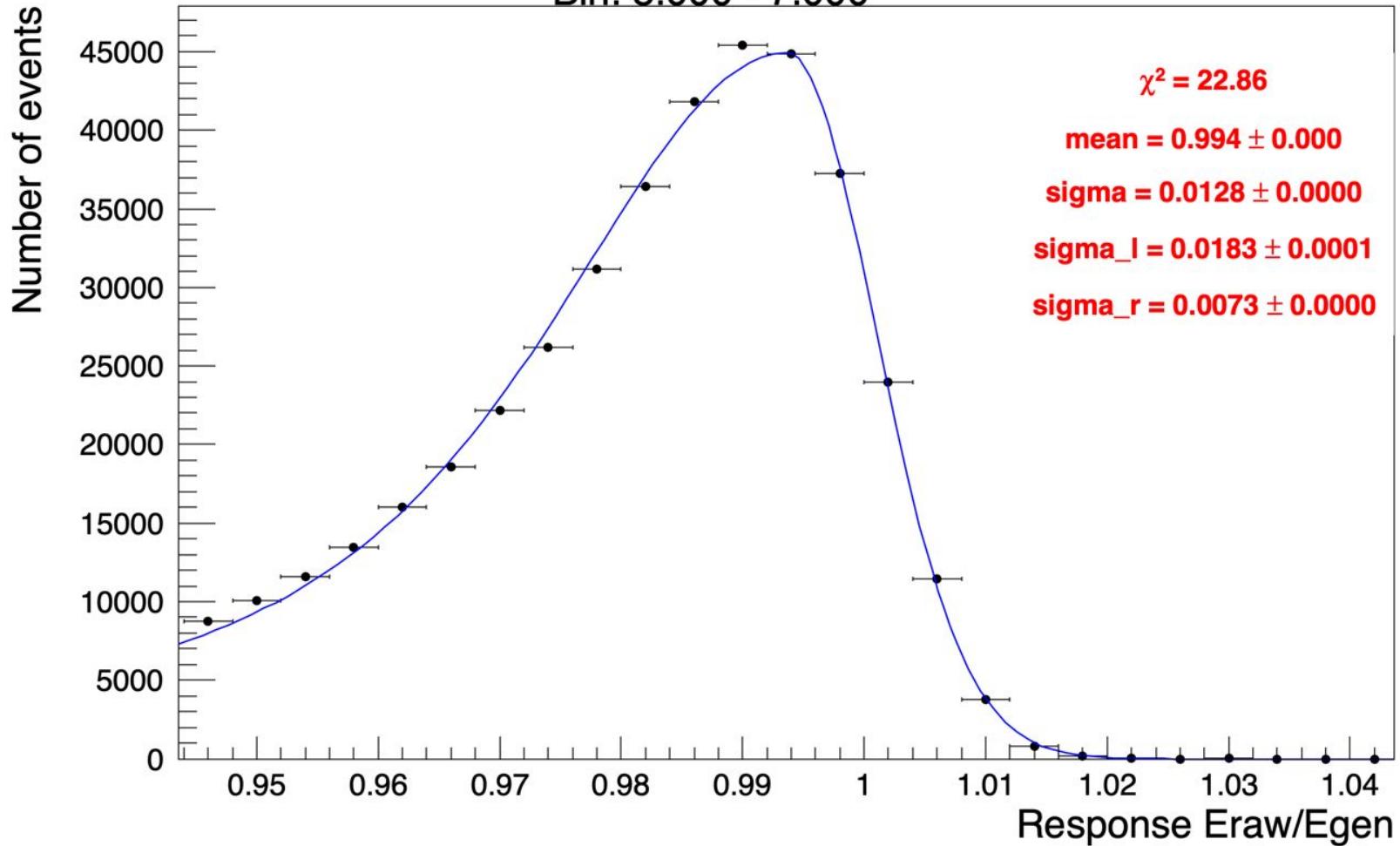
Bin: 0.990 - 1.000

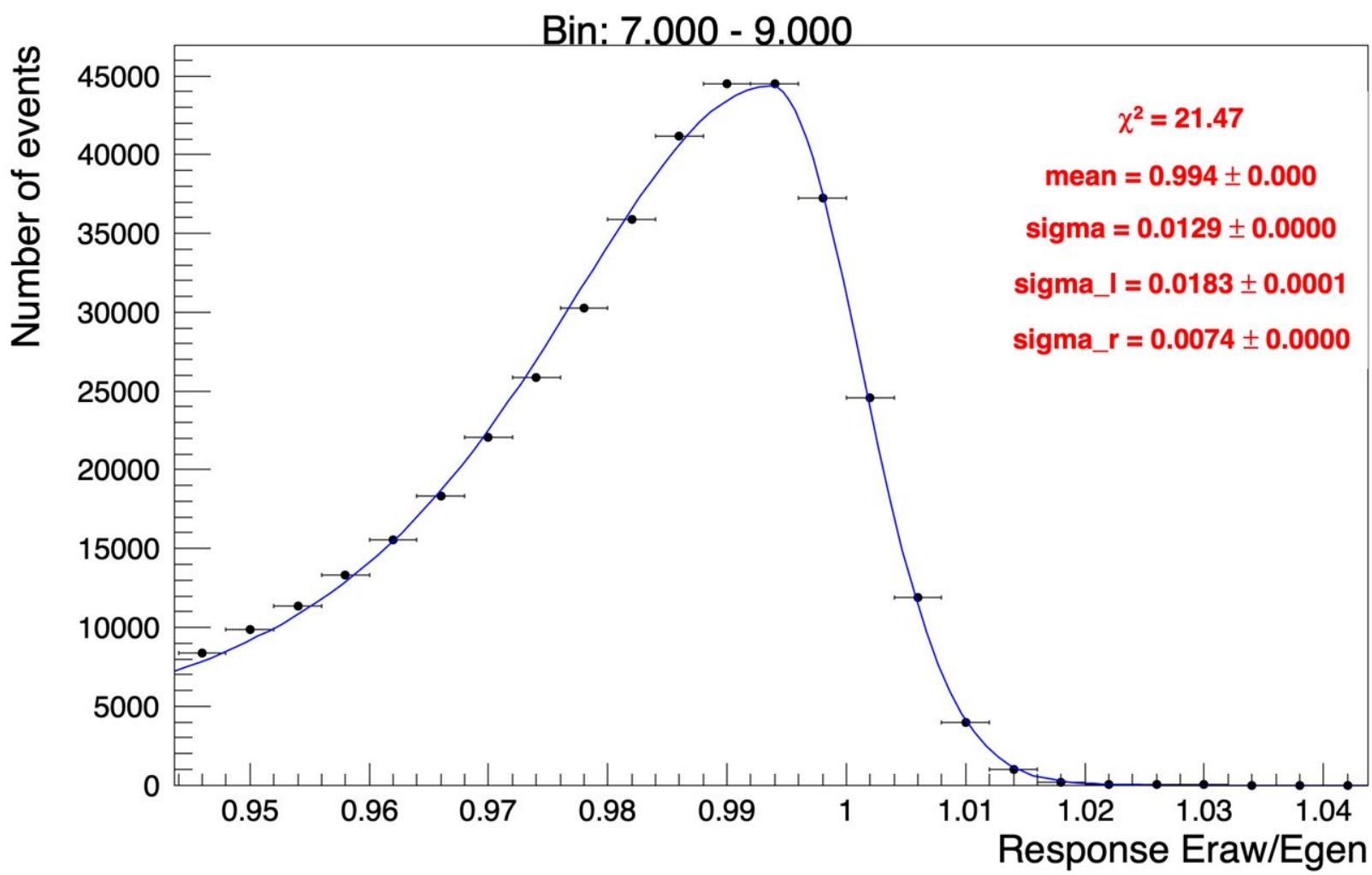




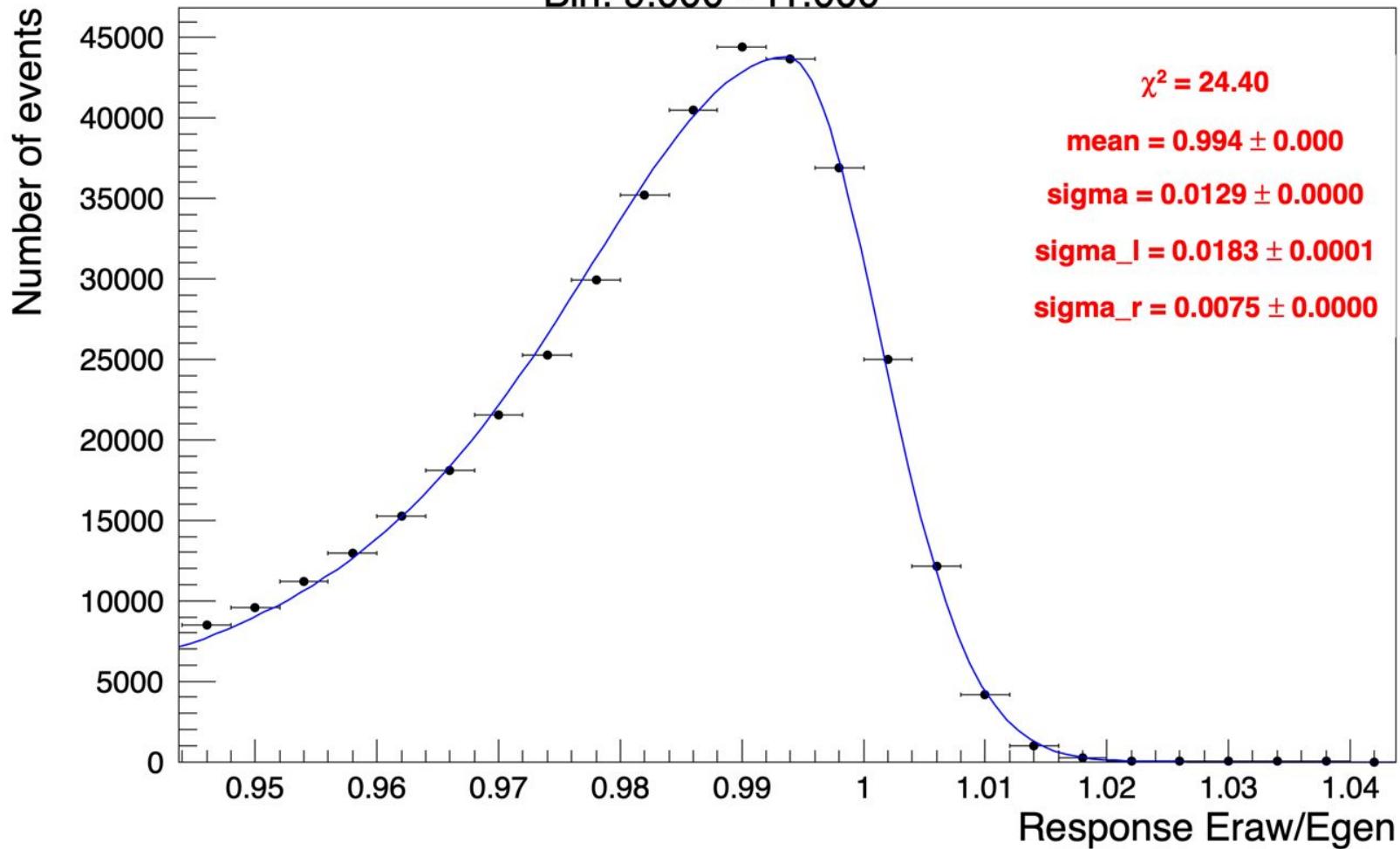
# Rho raw

Bin: 5.000 - 7.000

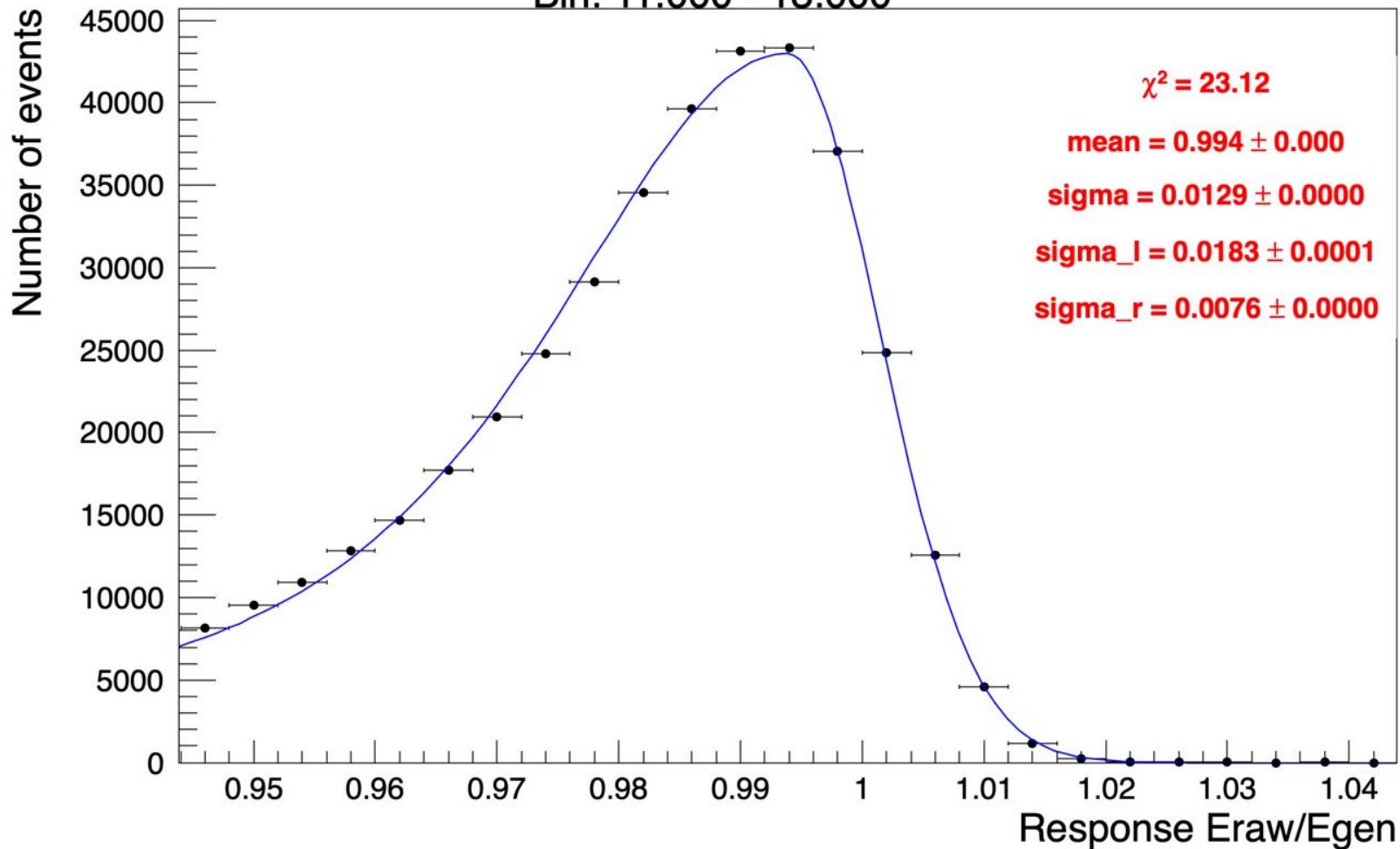


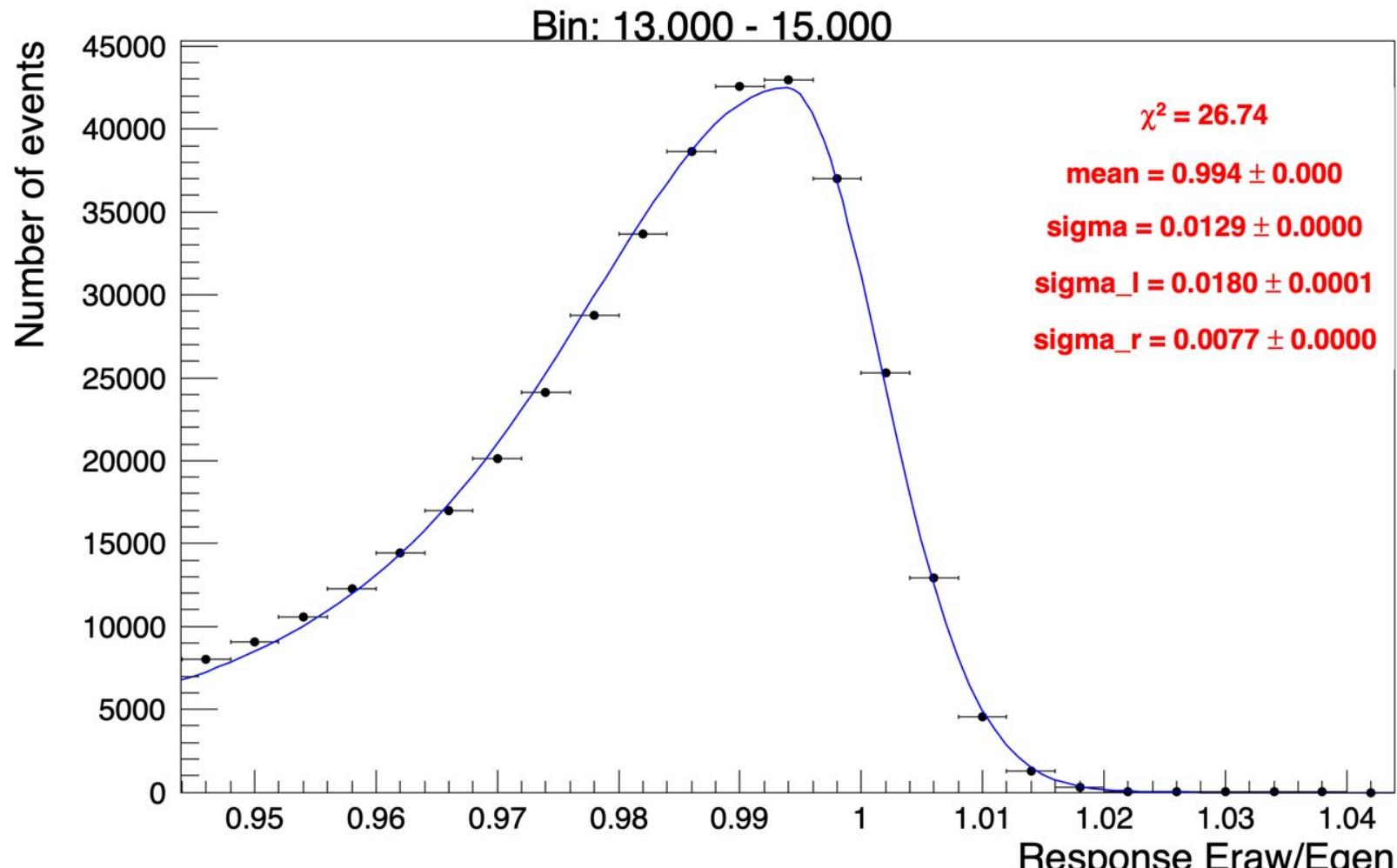


Bin: 9.000 - 11.000

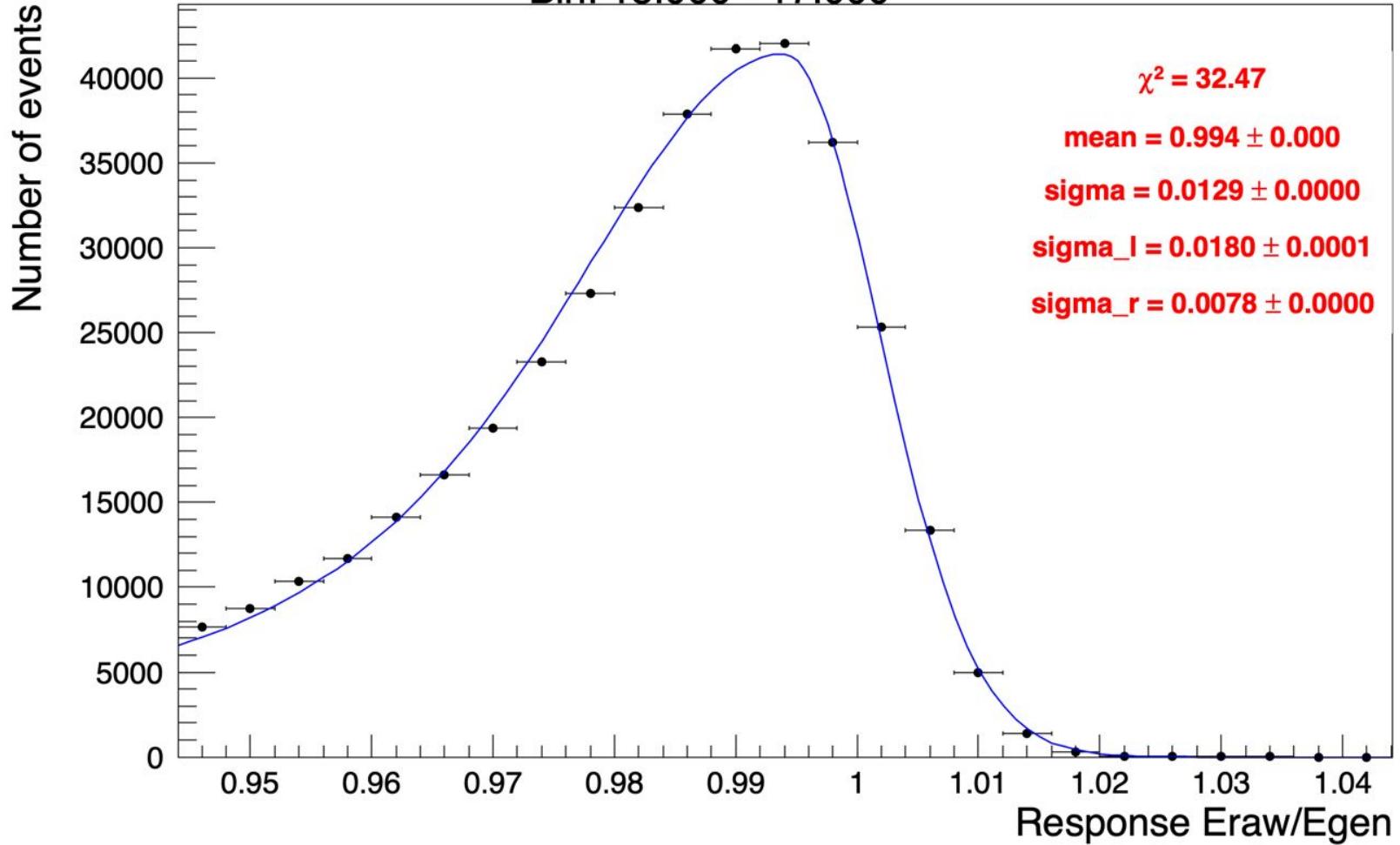


Bin: 11.000 - 13.000

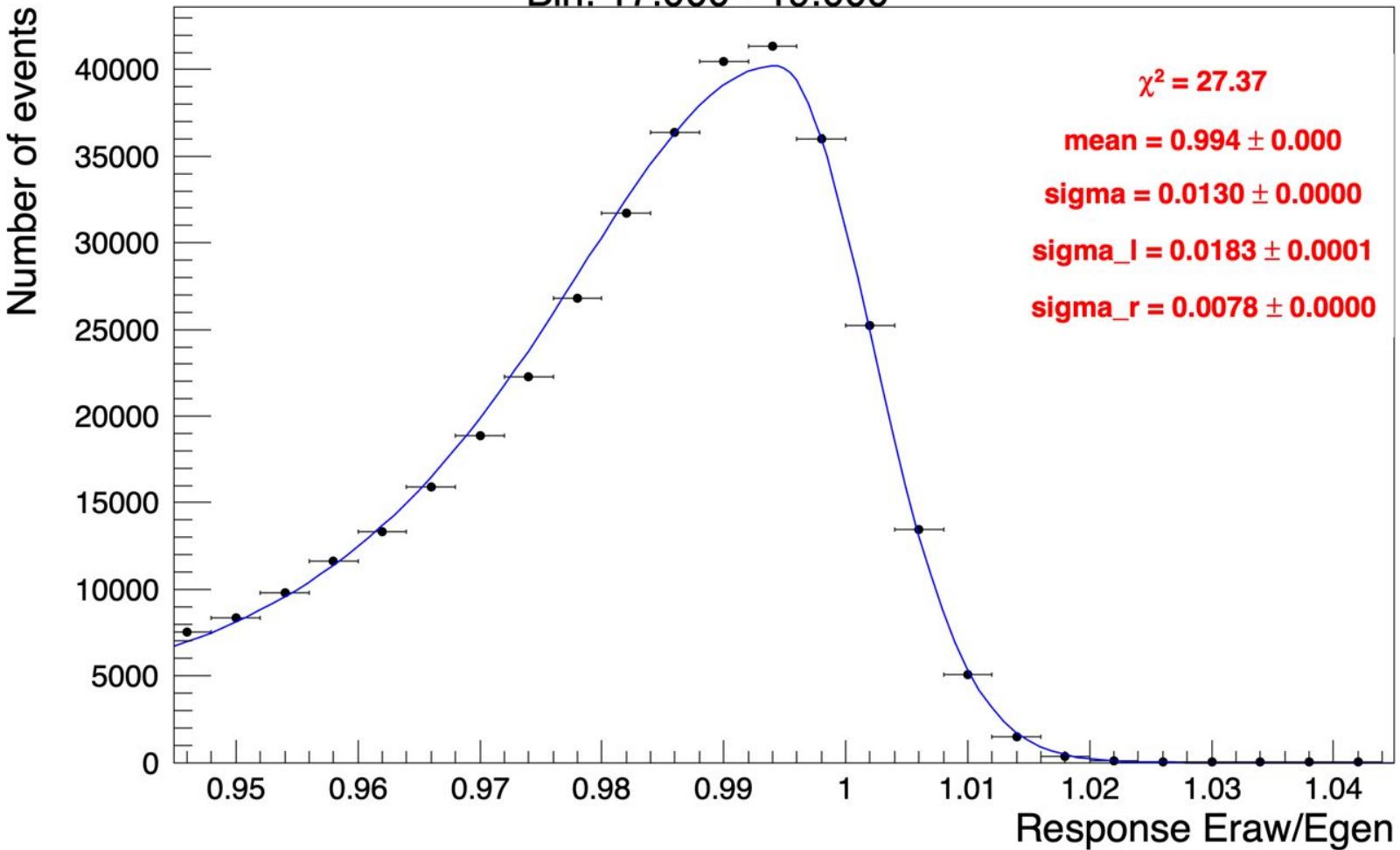




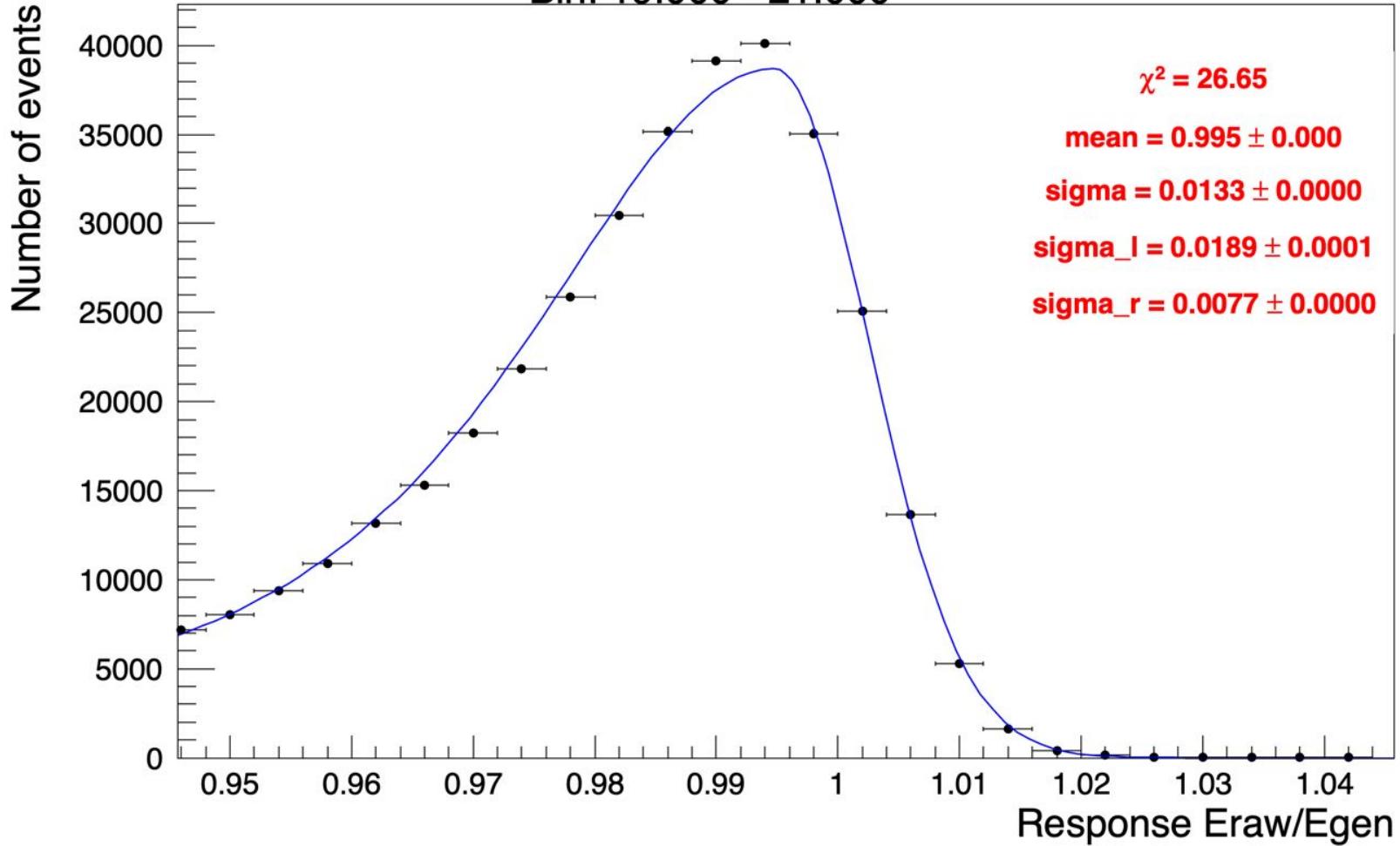
Bin: 15.000 - 17.000

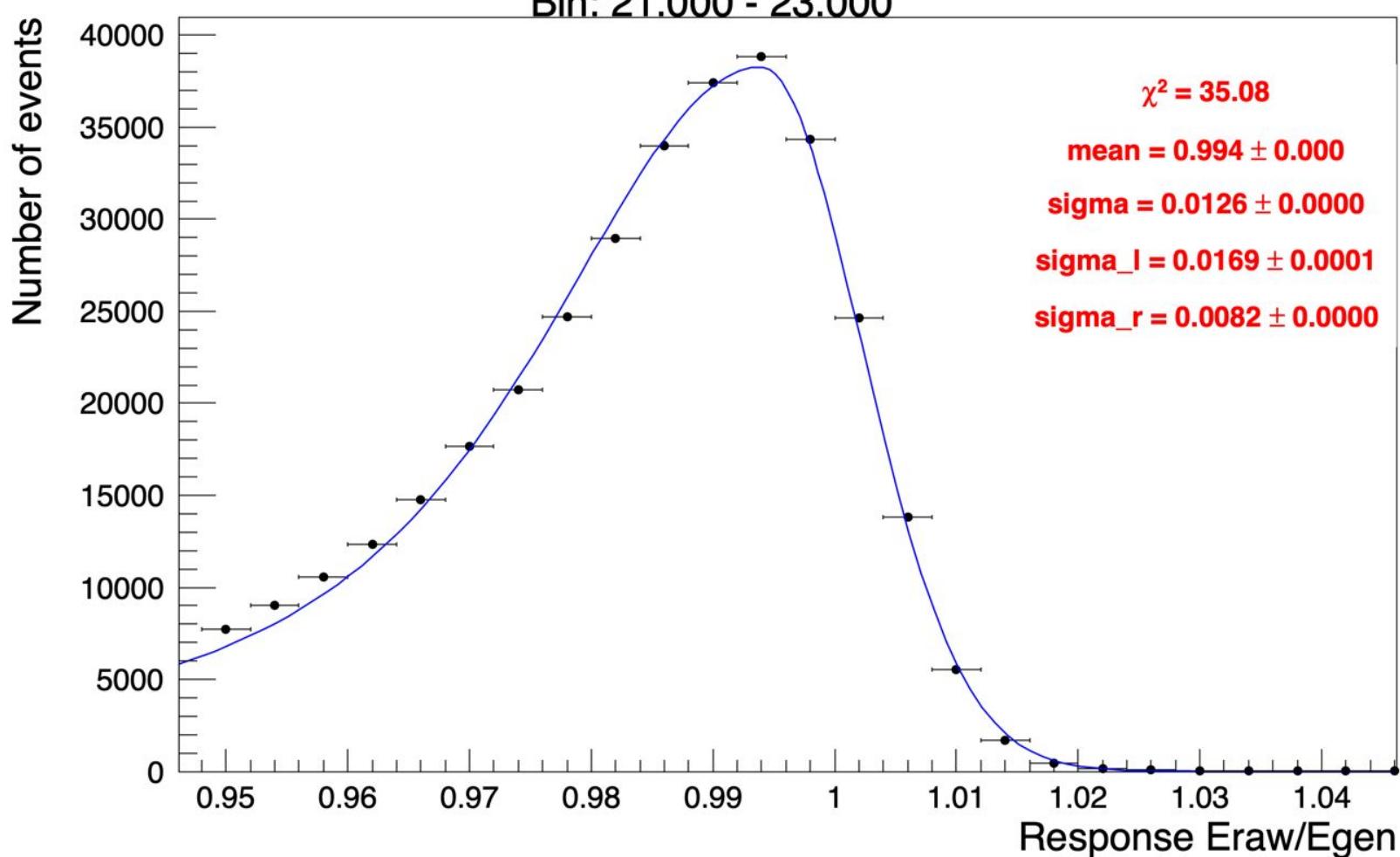


Bin: 17.000 - 19.000

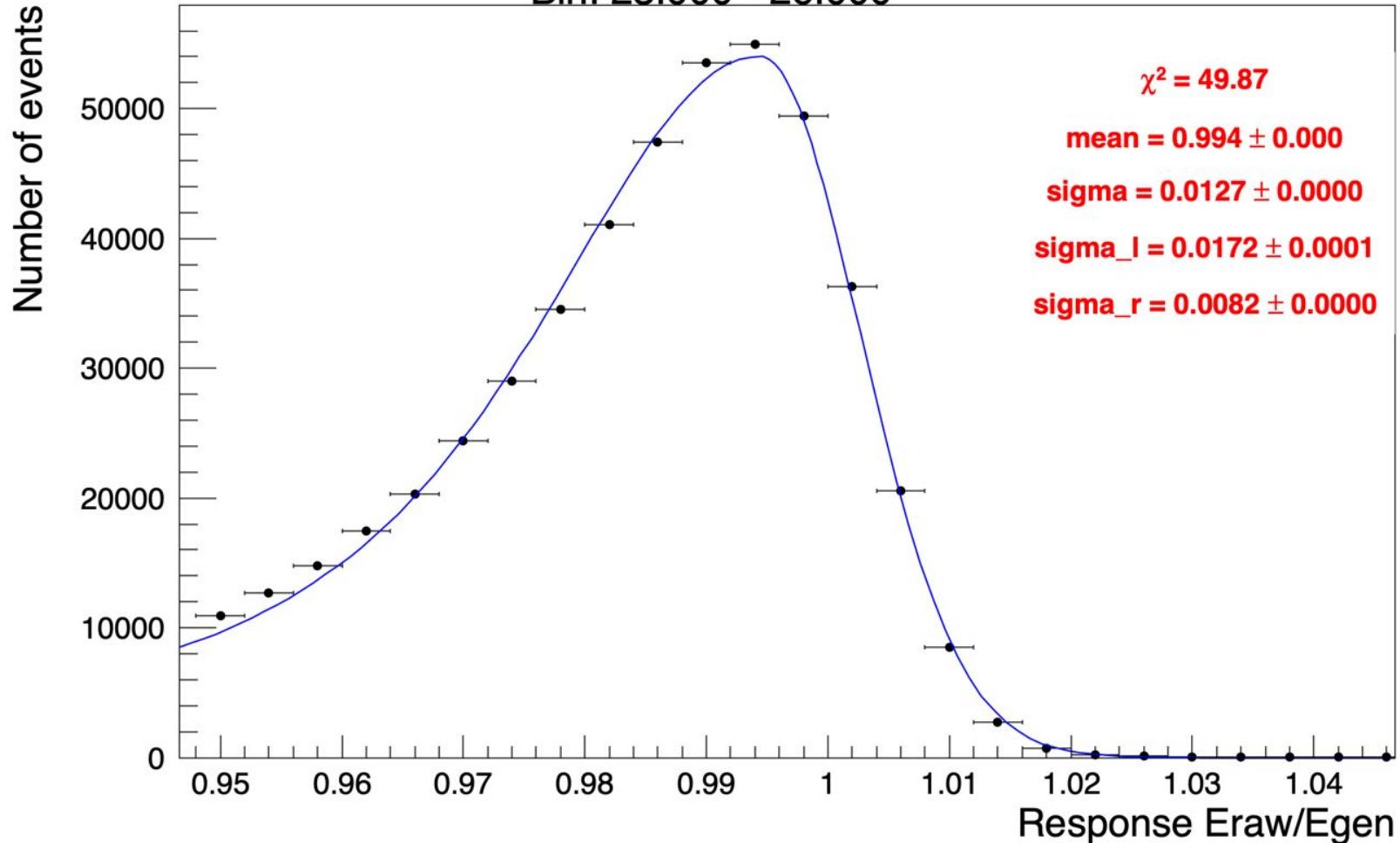


Bin: 19.000 - 21.000

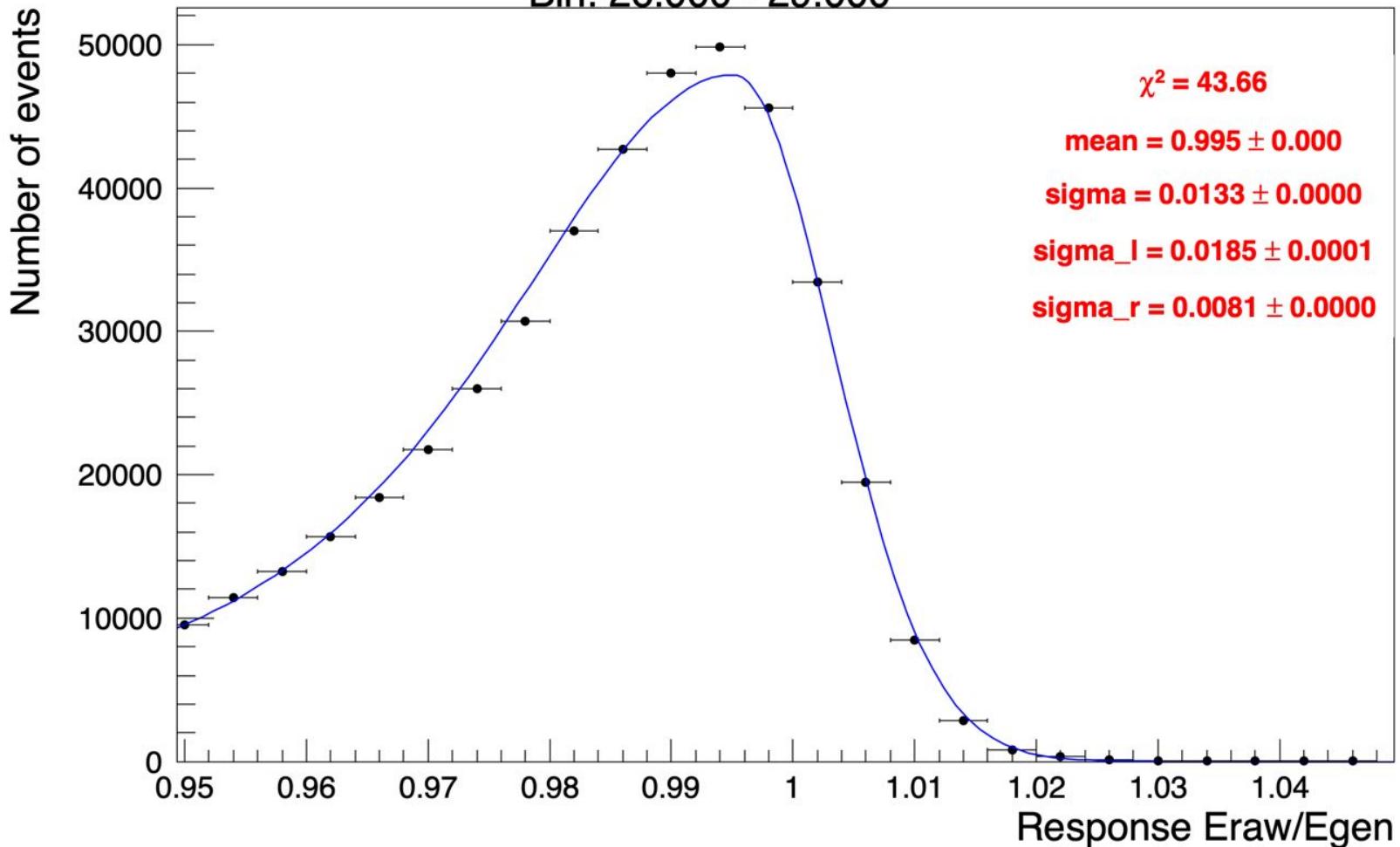




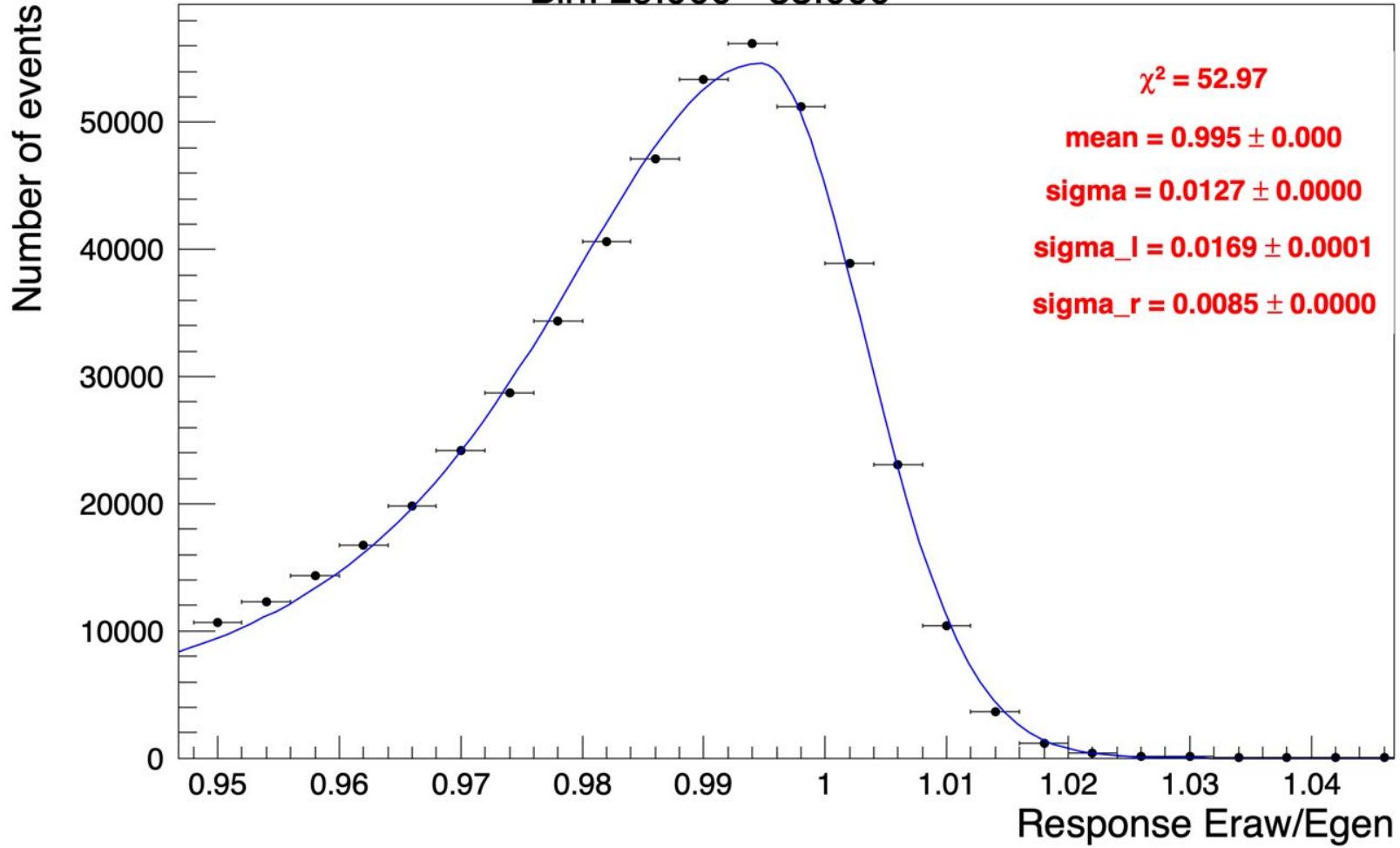
Bin: 23.000 - 26.000



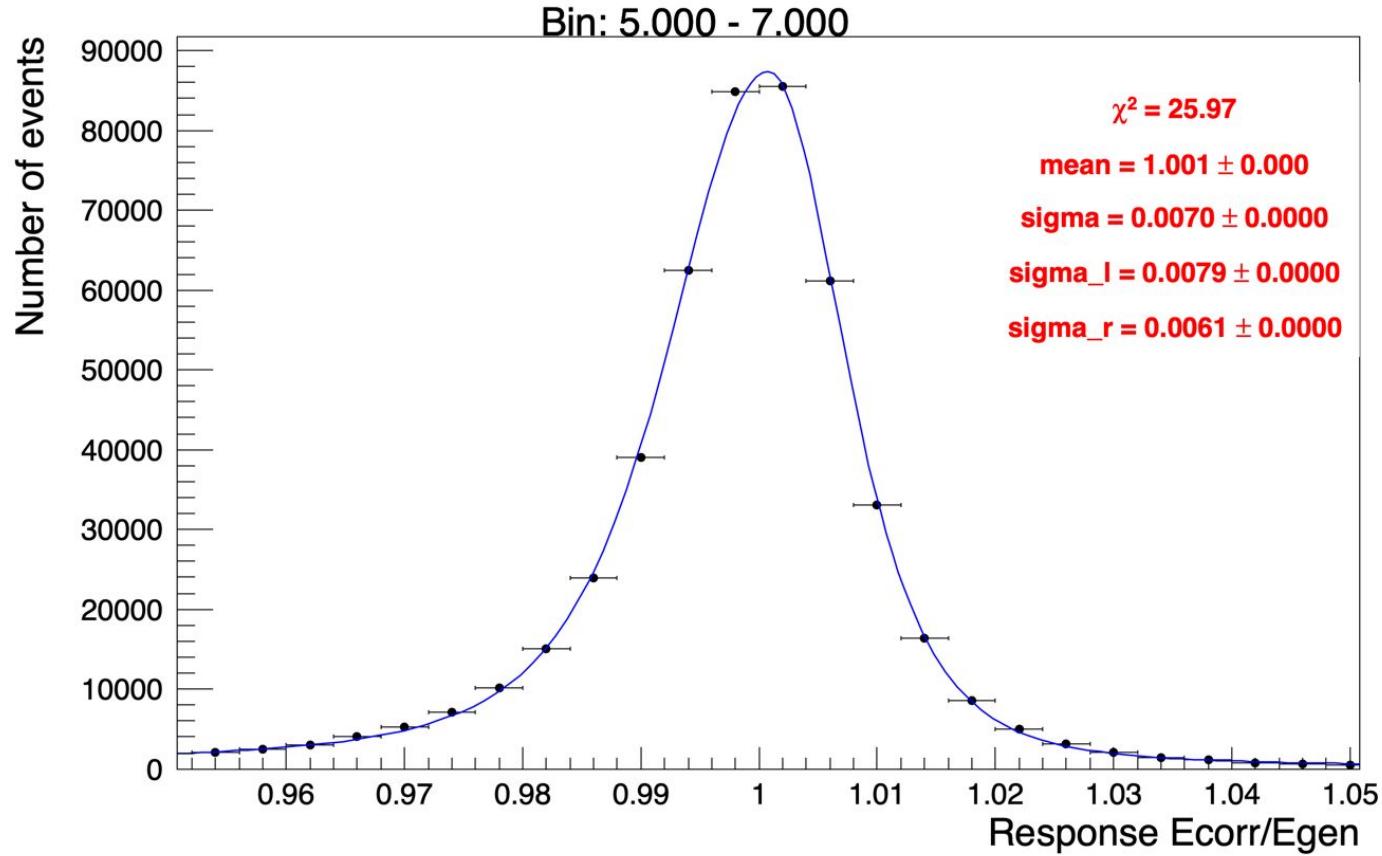
Bin: 26.000 - 29.000



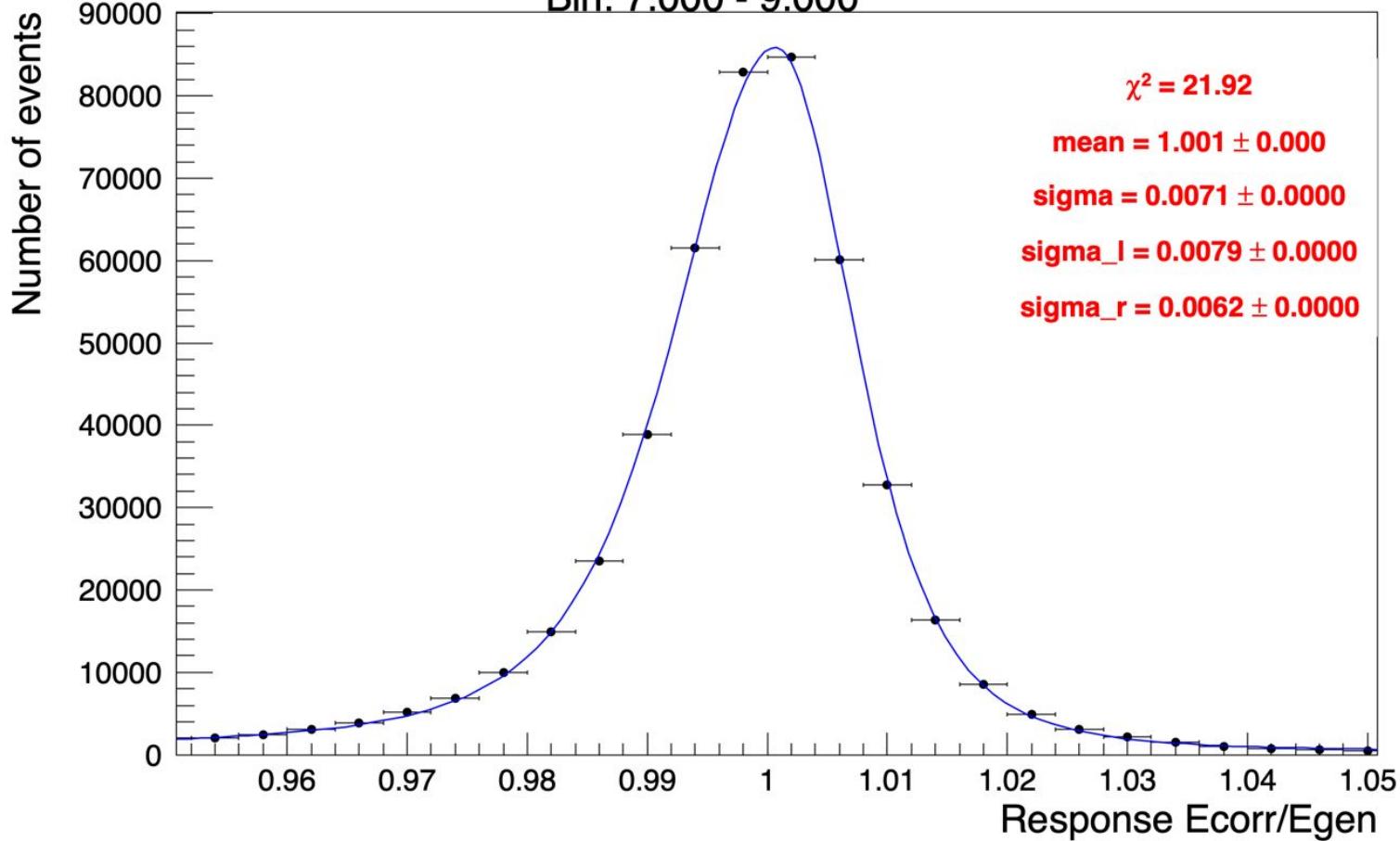
Bin: 29.000 - 33.000



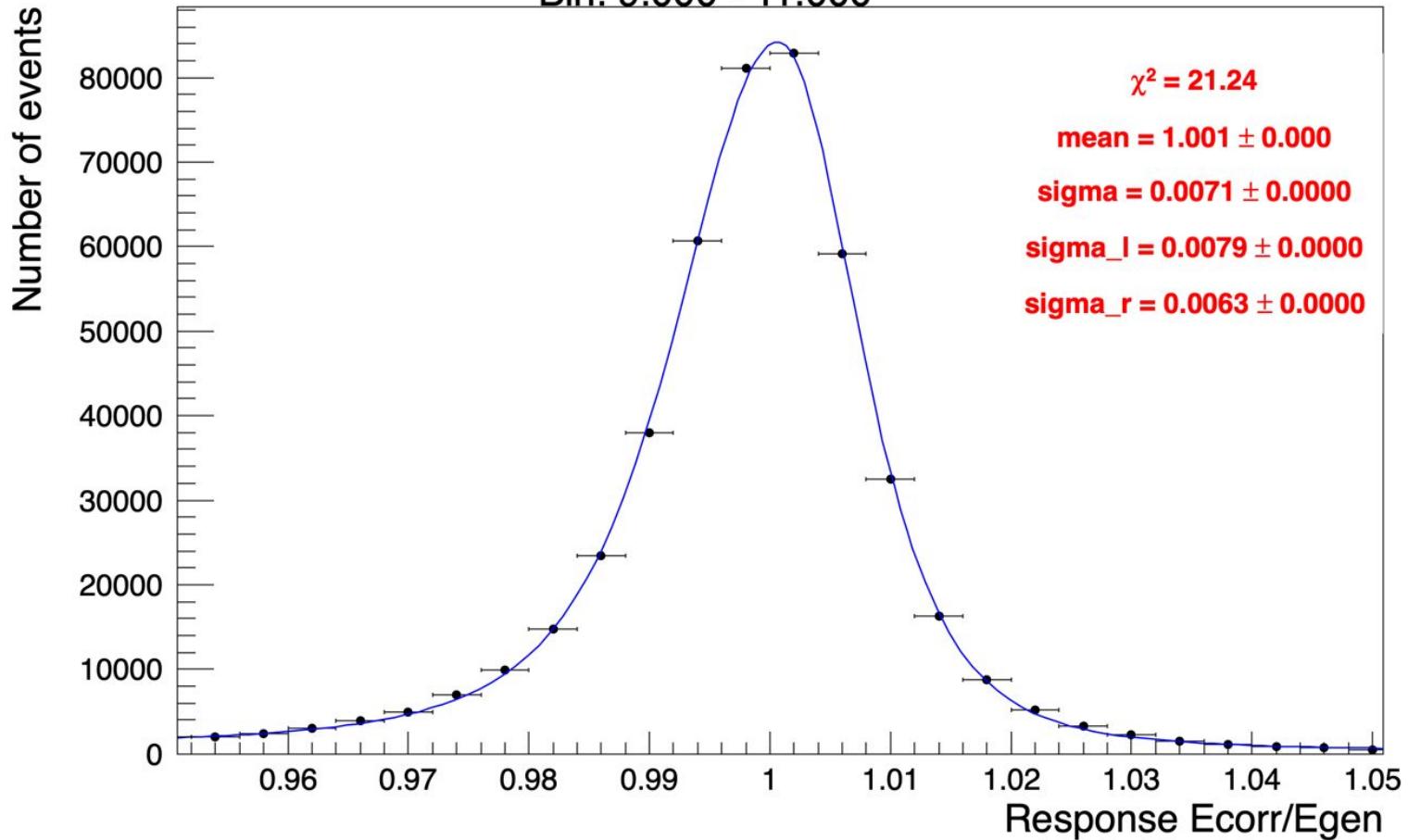
# Rho corrected



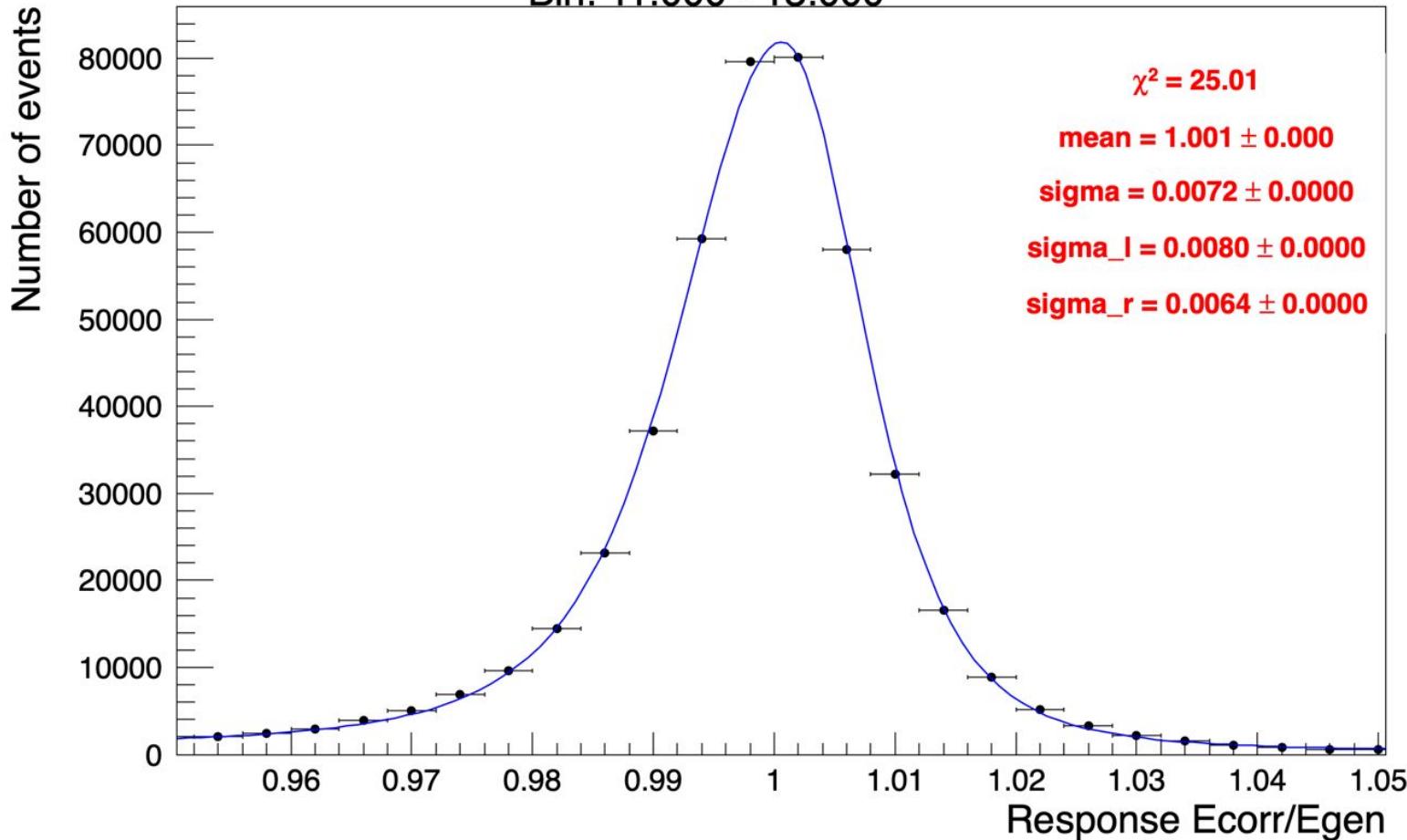
Bin: 7.000 - 9.000



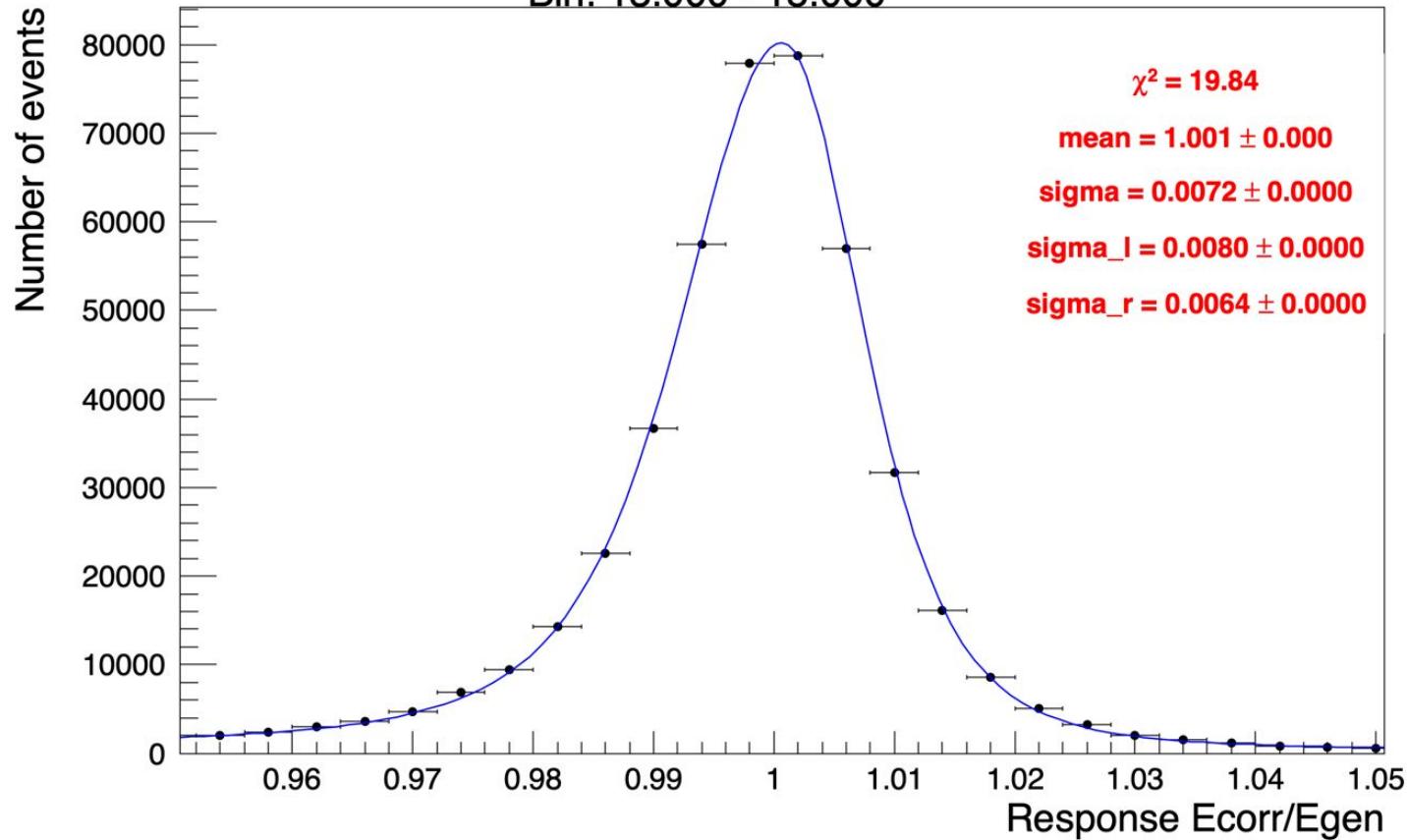
Bin: 9.000 - 11.000

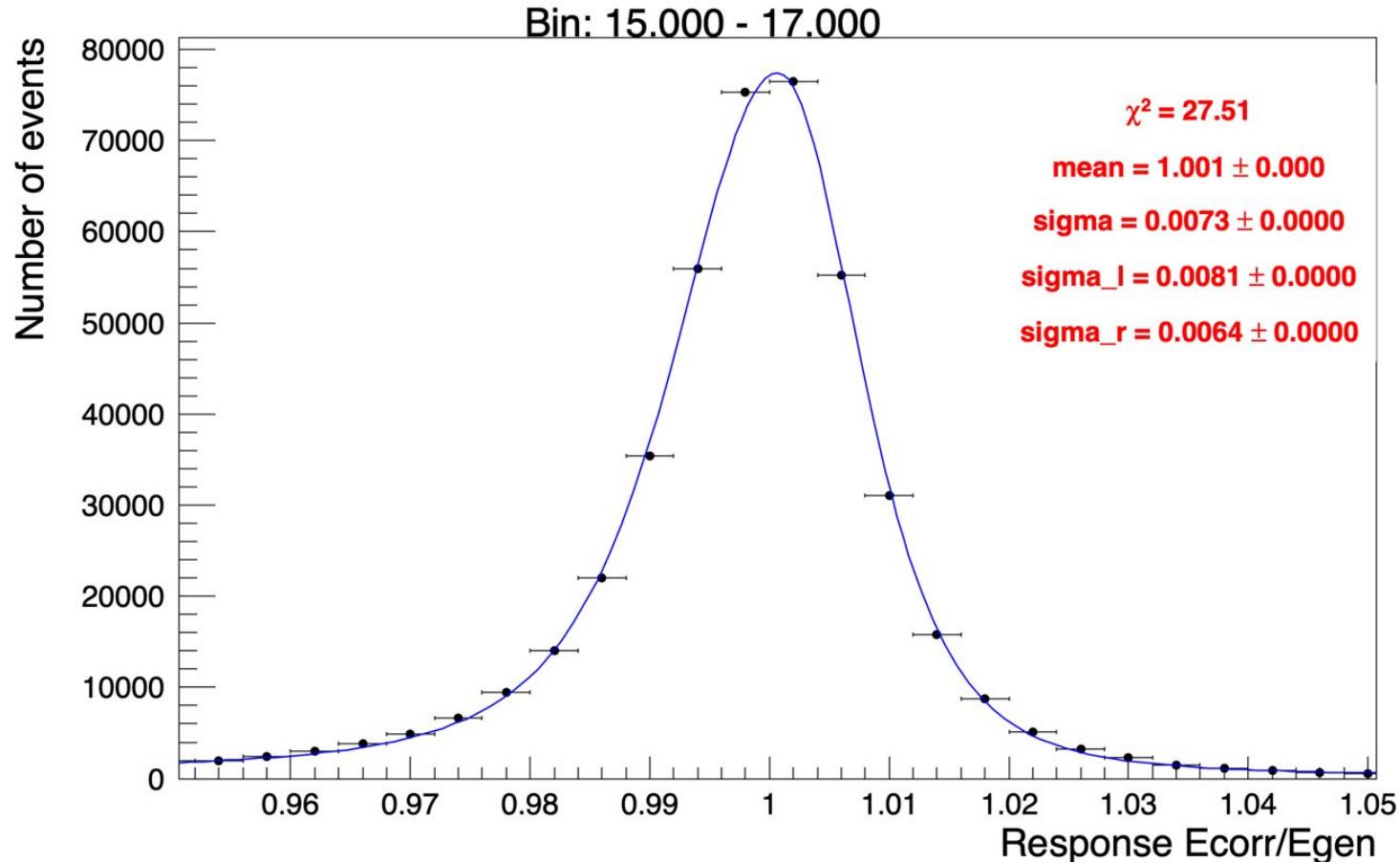


Bin: 11.000 - 13.000

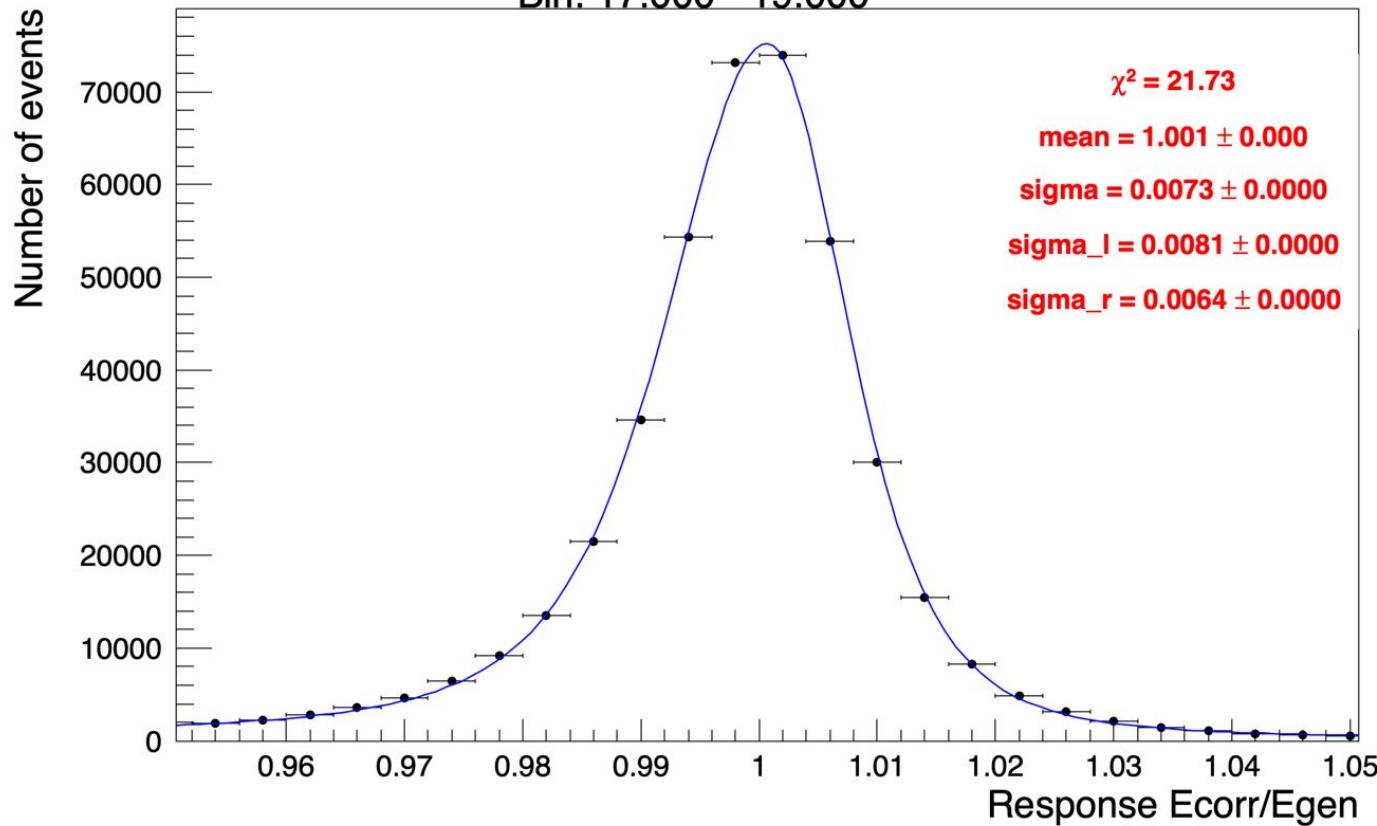


Bin: 13.000 - 15.000

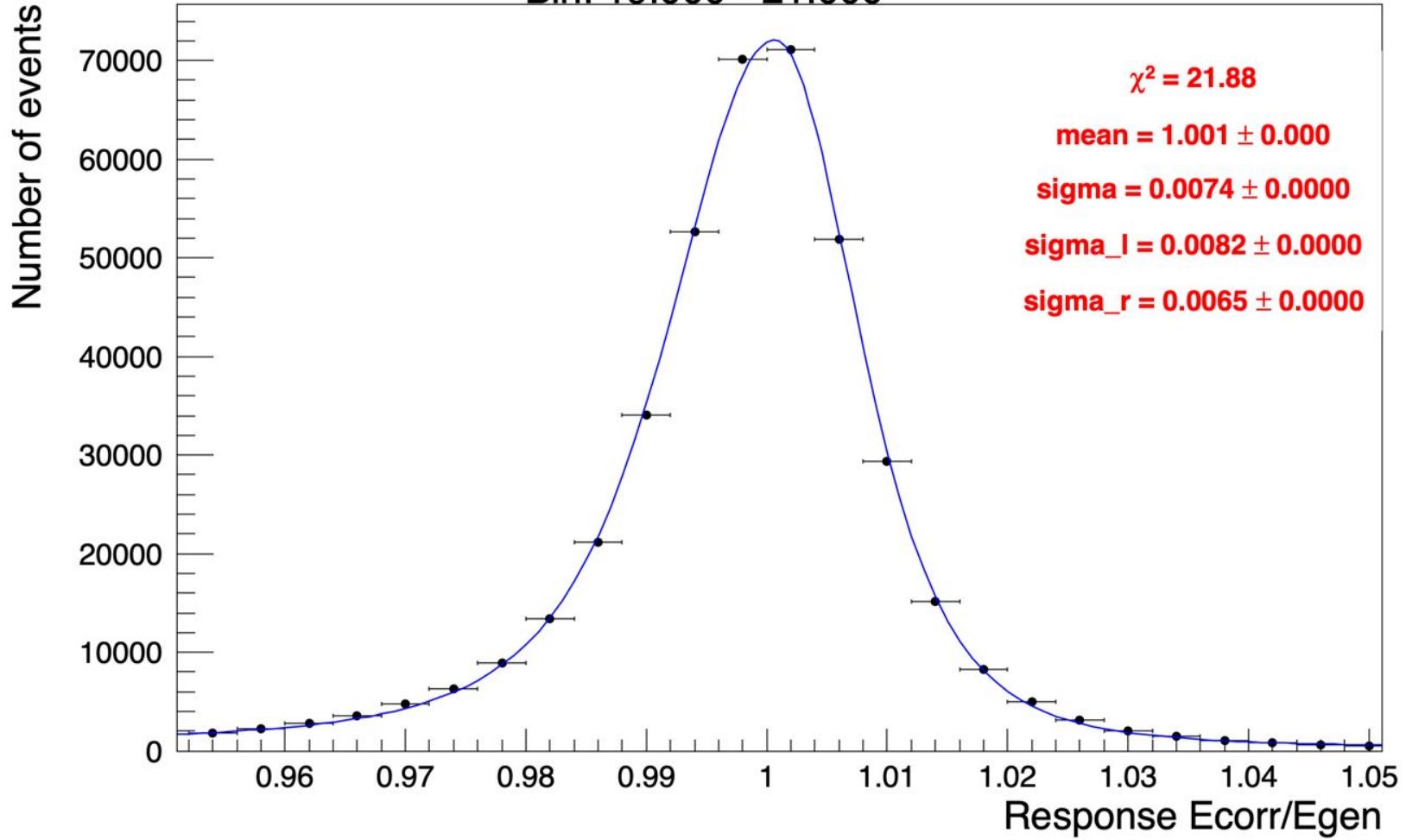




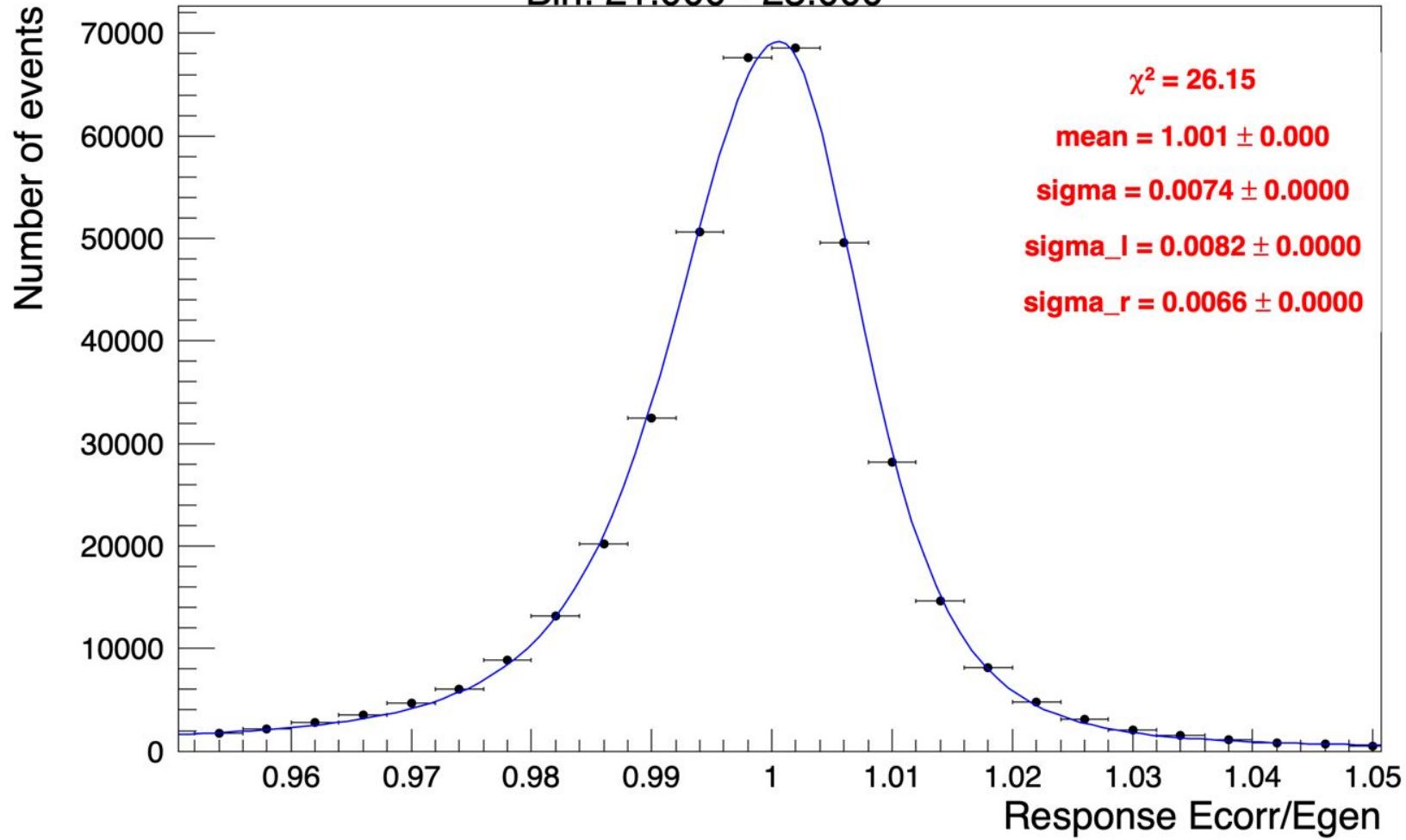
Bin: 17.000 - 19.000

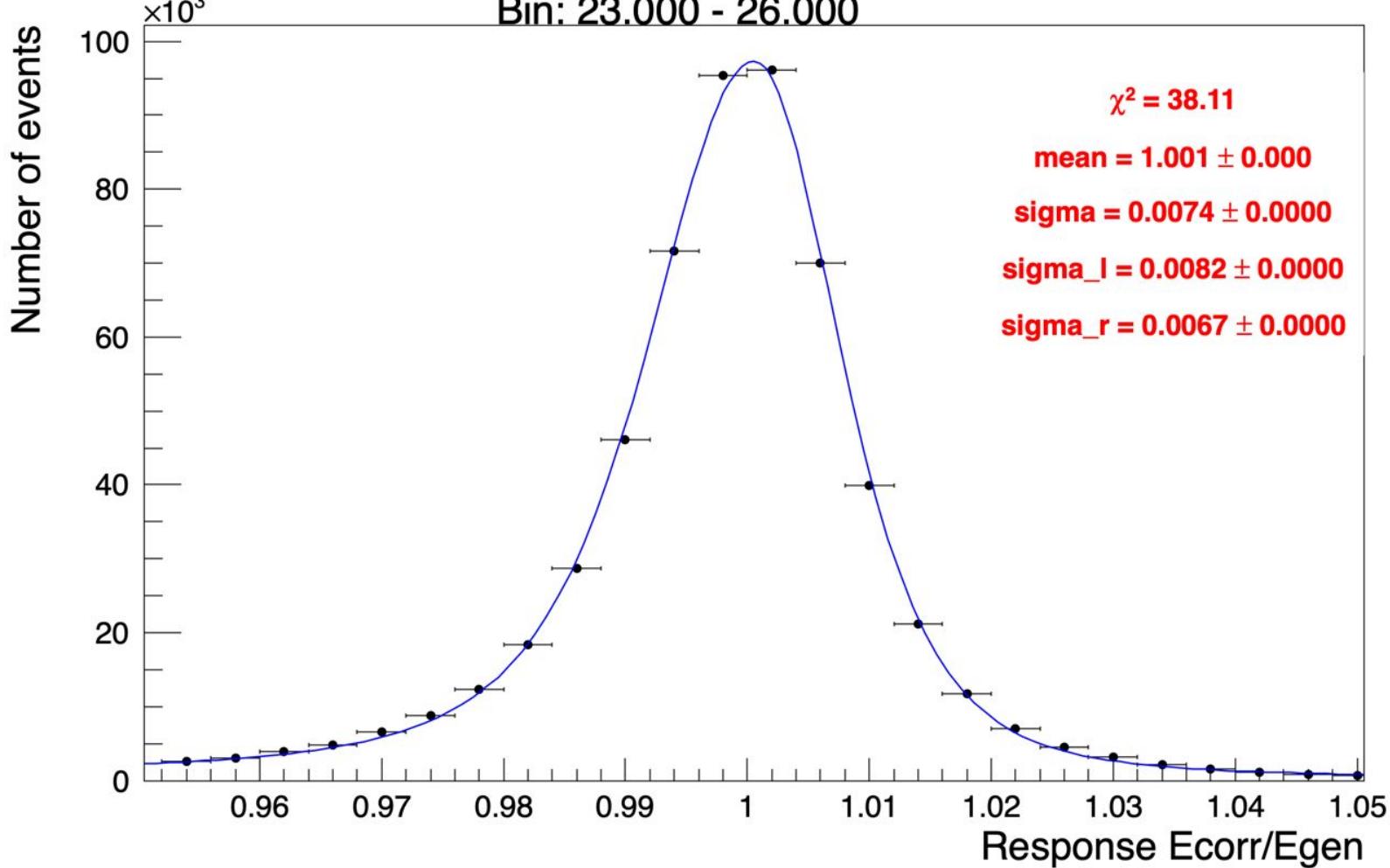


Bin: 19.000 - 21.000



Bin: 21.000 - 23.000





Bin: 26.000 - 29.000

