Per Event Mass Resolution of the Higgs Boson

Alison Tully

September 3, 2015

Main supervisor: Giovanni Petrucciani

Second supervisor: Michail Bachtis

Abstract

The mass resolution of the Higgs boson was studied in the decay channel $H \rightarrow ZZ \rightarrow 4l$, where $4l$ is any combination of ee and $\mu\mu$ pairs, for collisions at $\sqrt{s} = 13\,\text{TeV}$ recorded by the CMS detector at the LHC. The per event mass resolution was calculated by propagating the uncertainty of the lepton transverse momentum and was found to underestimate the actual resolution. To correct the underestimation, a scale factor, λ, was applied linearly to the estimate of the mass resolution, and hence the uncertainty of the lepton p_T. λ was derived from $Z \rightarrow ll$ decays, and measured to be 1.164 for electrons and 1.133 for muons.

1 Introduction

The discovery of the Higgs boson was announced in 2012 by the ATLAS and CMS experiments at the LHC. This was a historic moment in particle physics as the discovery of the Higgs boson was confirmation of the last piece of the Standard Model (SM) and the culmination of decades of searches. The $H \rightarrow ZZ \rightarrow 4l$ decay mode played a crucial role in this discovery as four lepton final states provide a clean and distinct experimental signature.

Measurements of the properties of the Higgs boson are vital to allow stronger tests of the SM and constrain new physics models. The mass of the Higgs is the first important measurement. This report presents an analysis of the mass resolution of the Higgs in the decay channel $H \rightarrow ZZ \rightarrow 4l$.

2 Per Event Mass Resolution

The uncertainty of the four lepton invariant mass, δm, varies considerably across the events in the data. Therefore, the estimation of the per event mass resolution is of relevance in a measurement of the Higgs mass. The spread of the mass resolution is dependent on the direction of the leptons and their transverse momentum, p_T.

For electrons, uncertainties in the momentum measurement arise from imperfect calibration of the Electromagnetic Calorimeter (ECAL) and uncertainty in the Gaussian-Sum Filter (GSF) track fit due to possible Bremsstrahlung emissions. The uncertainties in p_T are assessed from a combination of the quality of the ECAL supercluster and the GSF track fit. For muons, the uncertainties arise from the muon track due to multiple scattering of the muons in the material of the inner detector. They are assessed from the properties of the hits in the tracker and the muon system and the quality of the muon candidate fit.

In the CMS run I analysis of $H \to ZZ \to 4l$ [1], the estimate of δm was shown to underestimate the measured resolution. This is due to the lepton momentum uncertainties not being perfectly tuned and also due to non-Gaussian tails on the resolution of the individual components. When the four leptons are considered together, the non-Gaussian tails should be partially absorbed in the combined mass resolution from the central limit theorem, however, the effect still contributes to an underestimation of the uncertainty of the lepton momenta.

To correct the underestimation, a scale factor needs to be applied to the estimate of the mass resolution. This report presents a measurement of the scale factor in $H \to ZZ \to 4l$ decays at 13 TeV.

3 Estimating δm

Generally, for a decay with n particles in the final state, the δm of the parent particle can be estimated by propagating the uncertainty of the four-momenta (E, p_x, p_y, p_z) of the decay products. The covariance matrix of the i^{th} final state particle, C_i, is calculated by propagating the uncertainty on the momentum, δp_i, using a 1×4 Jacobian matrix j_i:

$$j_i = \left(\frac{\partial E_i}{\partial p_i}, \frac{\partial p_{x,i}}{\partial p_i}, \frac{\partial p_{y,i}}{\partial p_i}, \frac{\partial p_{z,i}}{\partial p_i} \right) = \left(\frac{p_i}{E_i}, \frac{p_{x,i}}{p_i}, \frac{p_{y,i}}{p_i}, \frac{p_{z,i}}{p_i} \right) \tag{1}$$

Explicitly, the entry in the p^{th} row and q^{th} column of C_i is given by:

$$C_{i,pq} = j_{i,p} j_{i,q} (\delta p_i)^2 \tag{2}$$

where $j_{i,p}$ is the p^{th} entry in j_i. From these, a 4n$\times$4n covariance matrix, C, can be formed by writing the n C_i in block diagonal form. The 1×4n Jacobian matrix to project from the full covariance matrix C to the invariant mass is given by:

$$J = \left(\frac{\partial m^2}{\partial E_1}, \frac{\partial m^2}{\partial p_{x,1}}, \frac{\partial m^2}{\partial p_{y,1}}, \frac{\partial m^2}{\partial p_{z,1}}, ..., \frac{\partial m^2}{\partial E_n}, \frac{\partial m^2}{\partial p_{x,n}}, \frac{\partial m^2}{\partial p_{y,n}}, \frac{\partial m^2}{\partial p_{z,n}} \right) \tag{3}$$

where the components are:

$$\frac{\partial m^2}{\partial E_i} = \frac{\partial}{\partial E_i} \left(\left(\sum_i E_i \right)^2 - \left(\sum_i p_i \right)^2 \right) = 2 \sum_i E_i \tag{4}$$

$$\frac{\partial m^2}{\partial p_i} = \frac{\partial}{\partial p_i} \left(\left(\sum_i E_i \right)^2 - \left(\sum_i p_i \right)^2 \right) = -2 \sum_i p_i \tag{5}$$

An estimate of the mass resolution is then given by:

$$\delta m = \frac{1}{2m}\sqrt{J^t C J} \tag{6}$$

4 The Z Boson

Initially, the per event mass resolution of the Z boson was studied using the lepton selection of the $H \to ZZ$ analysis. $Z \to ll$ decays, where $l = e, \mu$, are high statistics and have been studied extensively such that the mass and width of the Z are tightly constrained. Additionally, since both decays rely on lepton isolation the detector response is expected to be similar. Therefore, $Z \to ll$ decays provide a good proof of concept for the Higgs mass resolution calibration and can also accommodate a comparison between data and MC at 13 TeV.

4.1 MC Study

The mass resolution, δm, is dependent on the direction of the leptons, parameterised by the pseudorapidity η and the azimuthal angle ϕ, and the lepton transverse momentum p_T. The contribution to δm from the uncertainty of each of these individual parameters was assessed using MC data. The difference in the Z boson mass, M_Z, when evaluated using the recorded values of the parameters and when evaluated from the generated values gives each parameter's contribution to δm. For example, $M_Z(p_T^{rec}, \eta^{rec}, \phi^{rec}) - M_Z(p_T^{gen}, \eta^{rec}, \phi^{rec})$ gives the contribution to δm from the uncertainty of p_T. The contribution to the uncertainty from p_T, η and ϕ is shown in Fig. 1 for $Z \to ee$ events and Fig. 2 for $Z \to \mu\mu$ events. The spread of the mass difference for p_T, of order a few GeV/c^2, is much greater than for η or ϕ, which are of order 20 MeV/c^2 and 6 MeV/c^2 respectively. Therefore, the resolution on the angles can be neglected and only the uncertainty on p_T is taken into account when calculating δm.

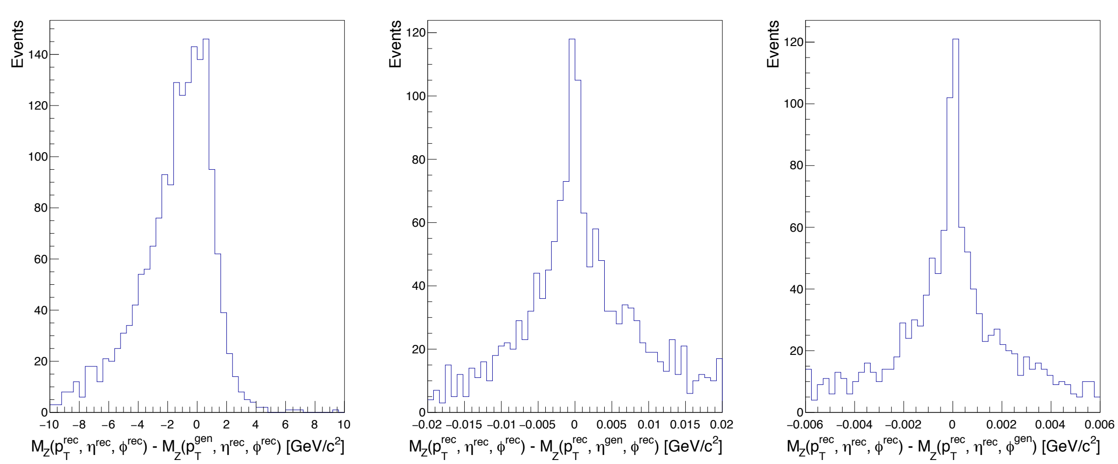

Figure 1: The difference in M_Z when evaluated from recorded and generated values of p_T (*left*), η (*right*) and ϕ (*right*) for $Z \to ee$ MC events.

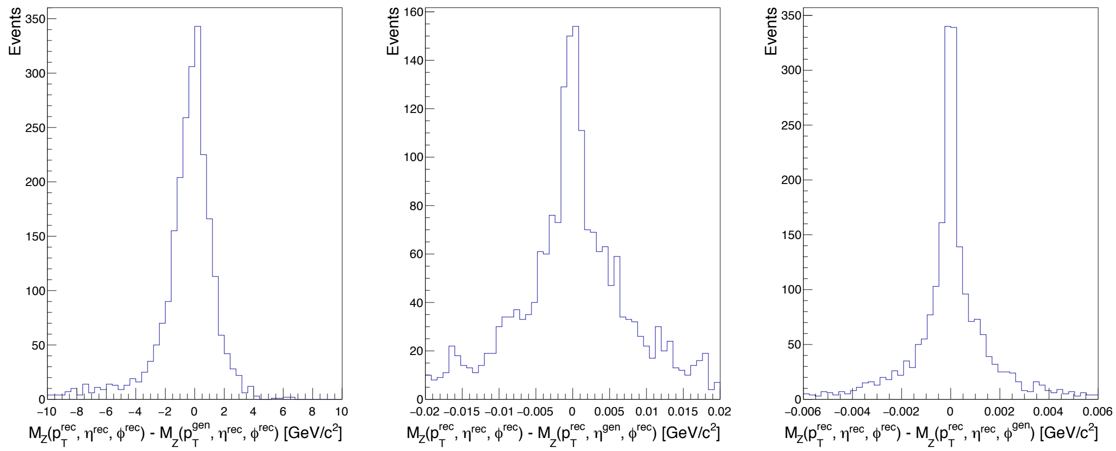

Figure 2: The difference in M_Z when evaluated from recorded and generated values of p_T (*left*), η (*right*) and ϕ (*right*) for $Z \to \mu\mu$ MC events.

4.2 δm Distribution

The mass resolution of $Z \to ll$ events, where $l = e, \mu$, was calculated using Eq. (6). There were around 60,000 data events used in this analysis, taken by the CMS detector during run II of the LHC at 13 TeV, and around 400,000 MC events with a 50 ns pile-up. There were two sets of MC events used with varying versions of the ECAL calibration to demonstrate the effect the new calibration has on the mass resolution. Normalised histograms of δm are shown in Fig. 3. The shape of the δm histogram is shown to change with the new calibration, however, there remain minor discrepancies between data and MC. The measurement of the electron energy is made using a combination of information from the ECAL and the tracker and so some disagreement between data and MC could remain if the combination is not yet fully calibrated.

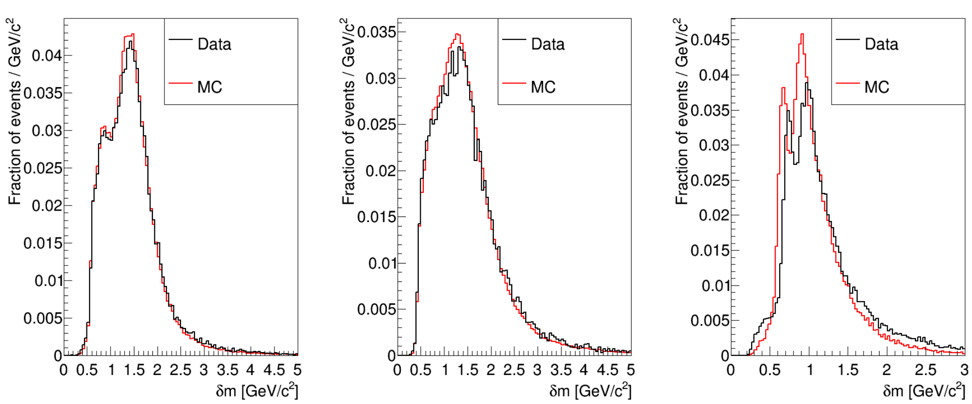

Figure 3: Normalised δm distributions of data (black points) and MC (red points) using $Z \to ee$ events (*left*), $Z \to ee$ events with the new ECAL calibration (*middle*) and $Z \to \mu\mu$ events (*right*).

4.3 Z Mass Model

The Z boson mass signal was modeled using a Breit-Wigner (BW) distribution convoluted with a Crystal Ball (CB) distribution, which consists of a Gaussian with an exponential tail towards the lower mass side. This is an empirical choice where the BW distribution models the decay resonance and the CB takes into account the detector resolution. The mean and width of the BW distribution were fixed to the PDG values [2] when fitting the m_{2l} invariant mass. The background was modeled using a decaying exponential.

The Z mass fit in the first δm bin is shown in Fig. 4 for $Z \to \mu\mu$ decays. This bin is below the p_T limit at which the Z mass can be formed and so doesn't contain true Z decays. This is because low δm corresponds to low values of the uncertainty of the lepton p_T, and hence small values of p_T. As decays with muons in the final state leave a more distinct experimental signature, the p_T of the muons can be measured more accurately than for the decays with electrons in the final state. Hence, this effect only appears for $Z \to \mu\mu$ decays. The low δm bins were excluded from the measurement of the calibration constant in $Z \to \mu\mu$ decays.

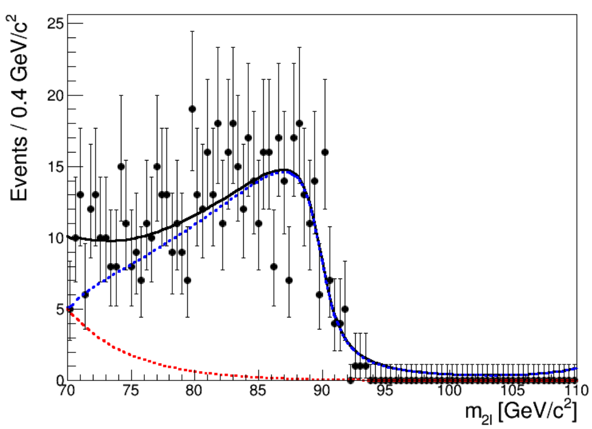

Figure 4: The two lepton invariant mass m_{2l} using $Z \to \mu\mu$ decays for the first δm bin.

4.4 Calibration Test

The dataset was divided into roughly equally populated bins of δm; for MC, 50 bins were used and for the lower statistics run II data, 20 bins were used. In each bin, the Z mass was fitted and the parameters of the CB were extracted. The width of the CB, σ_{CB}, corresponds to the measured mass resolution and so by comparing this to the predicted mass resolution from Eq. (6), the calibration can be tested. Plots of the measured mass resolution against the predicted mass resolution are shown in Fig 4. Both data and MC lie above the line for perfect calibration, showing that the prediction results in an underestimation of the mass resolution. The $Z \to ee$ MC moves within the 20% envelope with the new ECAL calibration. The $Z \to \mu\mu$ data lies within the 20% envelope, in agreement with the results from the 8 TeV analysis [1].

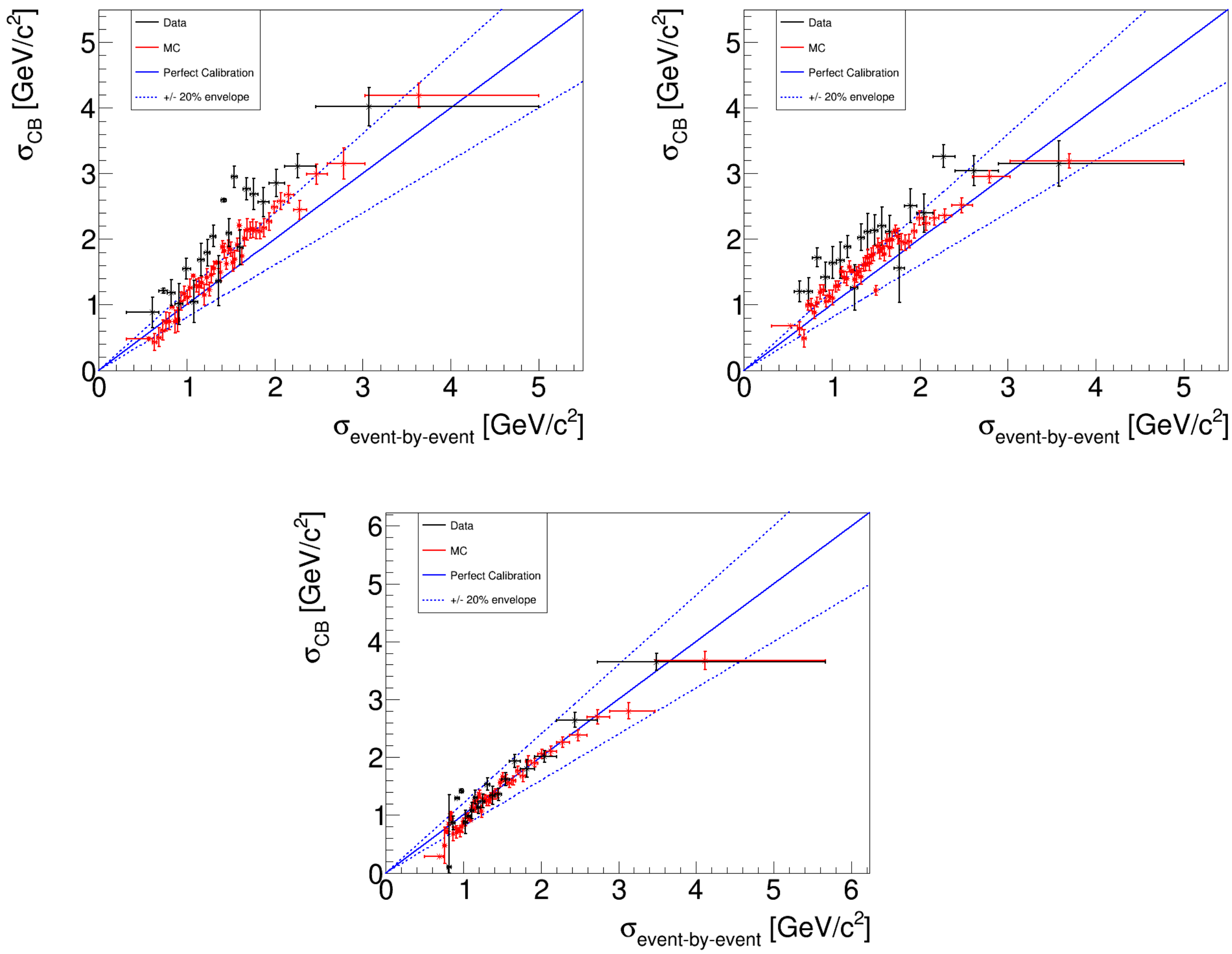

Figure 5: The measured mass resolution, σ_{CB}, plotted against the predicted mass resolution, $\sigma_{event-by-event}$, for data (black points) and MC (red points) using $Z \to ee$ events (*top left*), $Z \to ee$ events with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ events (*bottom*). The blue line is the line of perfect calibration and the dotted blue lines are the $\pm 20\%$ envelope.

To verify the choice of model and the MC results, σ_{CB} was also extracted by fitting a CB distribution to the Z mass resolution, $M_Z(p_T^{rec}, \eta^{rec}, \phi^{rec}) - M_Z(p_T^{gen}, \eta^{rec}, \phi^{rec})$. By fitting directly to the Z mass resolution, the CB parameters were measured with higher accuracy and the empirical choice of using a CB to measure the detector resolution was verified. The plots of σ_{CB} against $\sigma_{event-by-event}$ are shown in Fig. 6. As expected, the two methods are in agreement with one another.

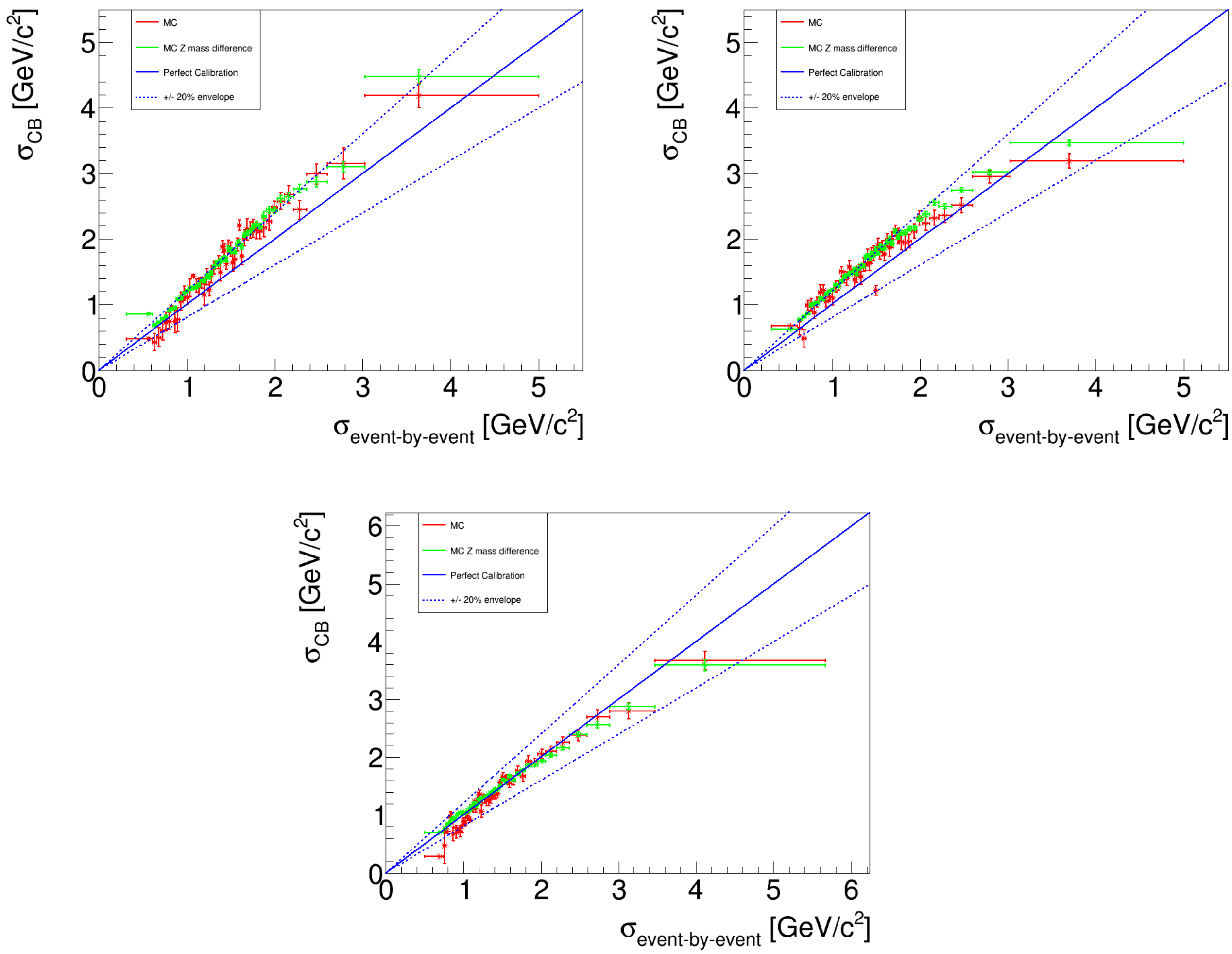

Figure 6: σ_{CB} plotted against $\sigma_{event-by-event}$ using $Z \to ee$ events (*top left*), $Z \to ee$ events with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ events (*bottom*). The red points are extracted by fitting directly to the Z mass and the green points are extracted by fitting to the Z mass resolution. The solid blue line is the line of perfect calibration and the dotted blue lines are the $\pm 20\%$ envelope.

4.5 Coverage Test

In addition, the coverage of σ_{CB} was investigated by taking the ratio of σ_{CB} with the effective width, $\sigma_{effective}$, which is the width of the CB that contains 68.3% of the distribution. For perfect coverage, $\sigma_{effective}/\sigma_{CB} = 1$. As shown in Fig. 6, the ratio of $\sigma_{effective}/\sigma_{CB}$ is greater than one, which means that the measurement of σ_{CB} results in an under-coverage of the confidence interval. The ratio decreases as $\sigma_{event-by-event}$ increases due to the Gaussian shape becoming more prominent in the mass model.

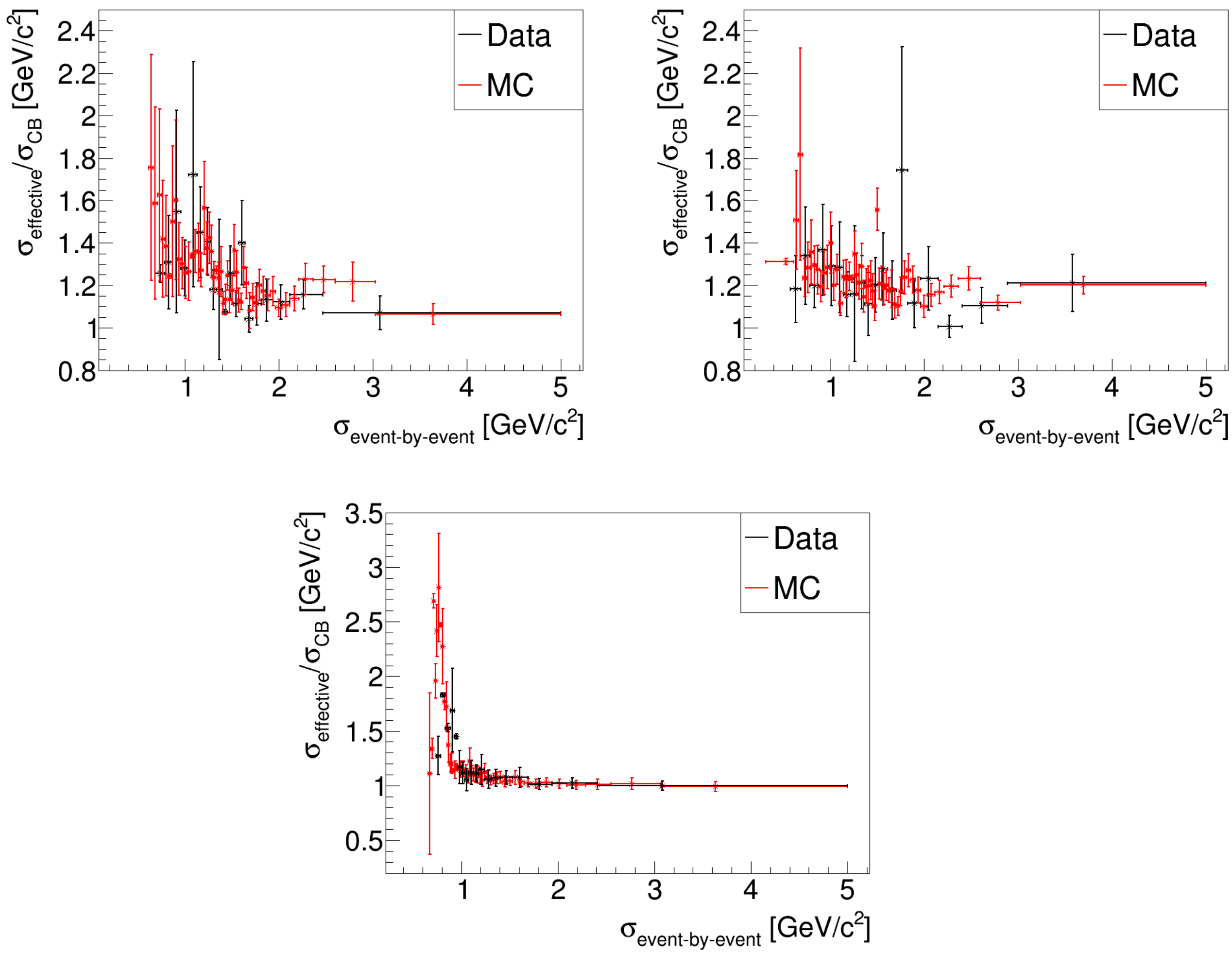

Figure 7: $\sigma_{effective}/\sigma_{CB}$ plotted as a function of $\sigma_{event-by-event}$ for data (black points) and MC (red points) using $Z \rightarrow ee$ events (*top left*), $Z \rightarrow ee$ events with the new ECAL calibration (*top right*) and $Z \rightarrow \mu\mu$ events (*bottom*).

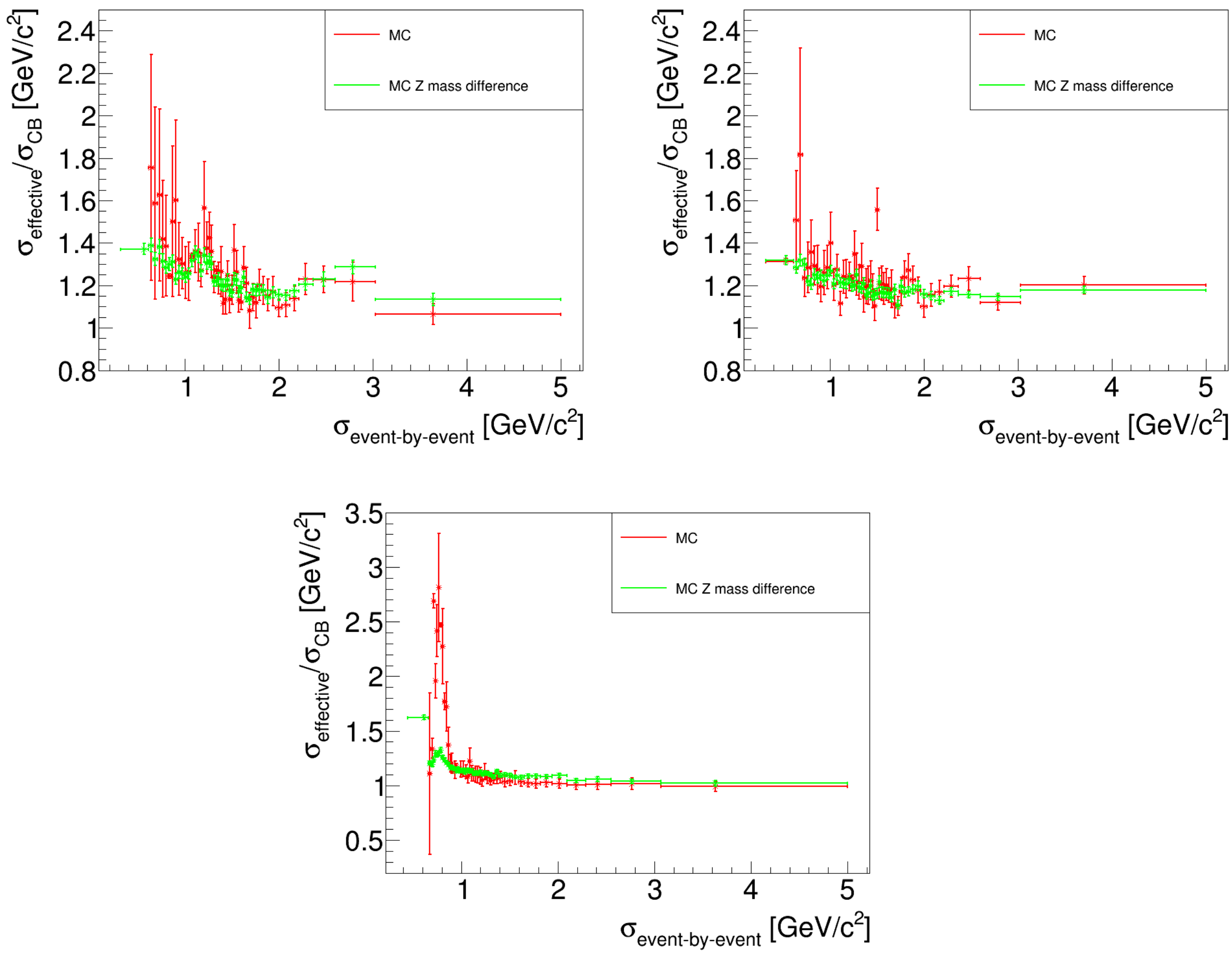

Figure 8: $\sigma_{effective}/\sigma_{CB}$ plotted as a function of $\sigma_{event-by-event}$ for MC using $Z \to ee$ events (*top left*), $Z \to ee$ events with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ events (*bottom*). The red points are extracted by fitting directly to the Z mass and the green points are extracted by fitting to the Z mass resolution.

4.6 Performing the Calibration

To correct the prediction of the mass resolution, a calibration constant λ is linearly applied to the mass resolution such that $\delta m^{calibrated} = \lambda \times \delta m^{estimated}$. λ is derived by fitting linearly to the MC points in the plots of σ_{CB} against $\sigma_{event-by-event}$ in Fig. 4. The MC extracted by fitting to the Z mass is used as the method can be used for both data and MC, whereas fitting to the Z mass resolution is only possible with MC. The linear fits are shown in Fig. 8. The low δm $Z \to \mu\mu$ points were excluded from the fit as these points are not from genuine Z decays, as described in Sec. 4.3. For $Z \to ee$ decays, λ decreased with the new ECAL calibration, from $\lambda = 1.179$ to 1.164. This decrease in λ quantitatively shows that the ECAL calibration improved the estimation of the mass resolution. For $Z \to \mu\mu$ decays, the value of λ was measured as 1.148.

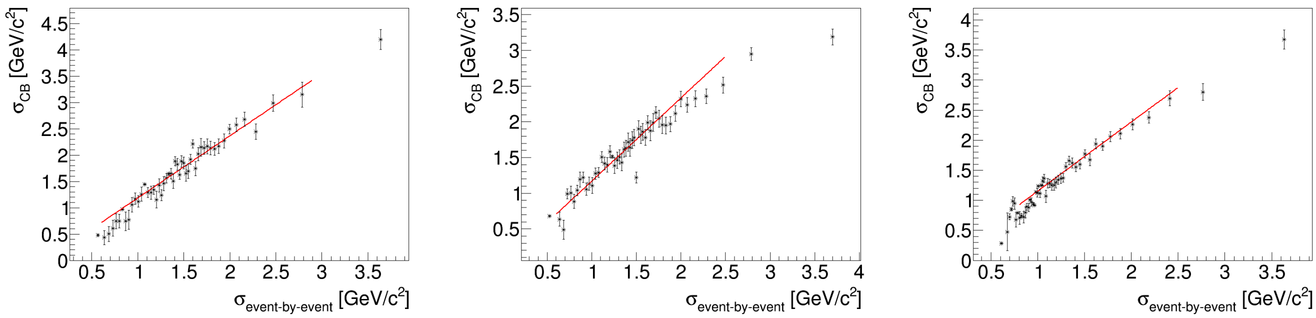

Figure 9: Linear fits to σ_{CB} as a function of $\sigma_{event-by-event}$ using $Z \to ee$ (*top left*), $Z \to ee$ with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ (*bottom*) MC events.

After applying the calibration constant to the predicted mass resolution, σ_{CB} was again plotted against $\sigma_{event-by-event}$, as shown in Fig. 9. The MC points now lie on the perfect calibration line and the calibration for the data has also improved, particularly for $Z \to \mu\mu$ events, which now lie well within the 20% envelope.

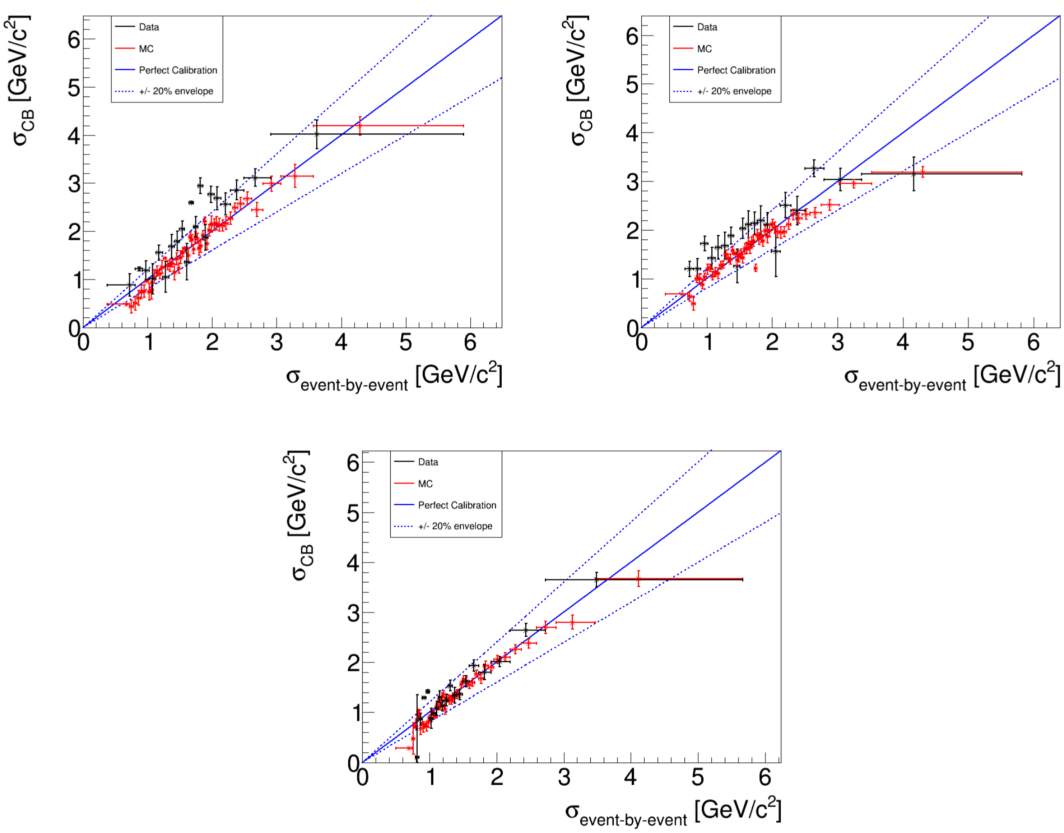

Figure 10: σ_{CB} plotted against the calibrated predicted mass resolution, $\sigma_{event-by-event}$, for data (black points) and MC (red points) using $Z \to ee$ events (*top left*), $Z \to ee$ events with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ events (*bottom*). The blue line is the line of perfect calibration and the dotted blue lines are the $\pm 20\%$ envelope.

5 The Higgs Boson

The calibration of the mass resolution was then applied to the Higgs boson for the decay mode $H \to ZZ \to 4l$ where $4l$ is any combination of ee and $\mu\mu$ pairs. There is not yet enough data collected by CMS during run II to conduct this analysis on data and so MC events were used.

5.1 δm Distribution

The mass resolution of $H \to ZZ \to 4l$ events, where $l = e, \mu$, was calculated using Eq. (6). There were around 80,000 MC events used in this analysis, the normalised histograms of which are shown in Fig. 11.

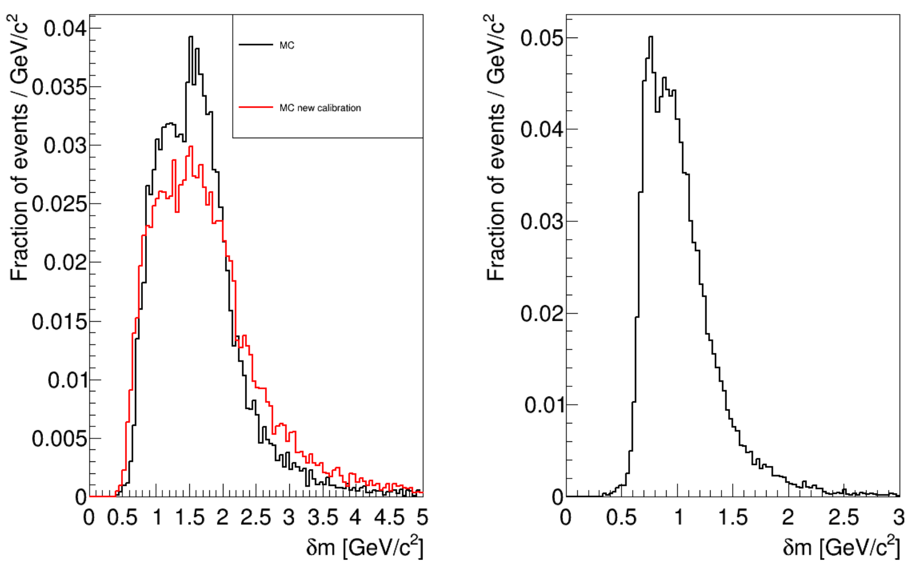

Figure 11: Normalised δm distributions of MC using $H \to ZZ \to 4e$ events (*left*) for the old (black points) and new (red points) versions of the ECAL calibration and $H \to ZZ \to 4\mu$ events (*right*).

5.2 Higgs Mass Model

The invariant mass of the four leptons, m_{4l}, was modeled using a CB distribution. This model differs from the one used for the Z boson since the Higgs boson width is estimated to be much smaller than the width of the Z boson. Hence, the width of the BW distribution is expected to be small and doesn't contribute to the mass peak. As MC events were used, there was no background model included. Plots of m_{4l} are shown in Fig. 11 for $H \to ZZ \to 4e$ decays. The new ECAL calibration improves the mass resolution, which is evident since the peak of the CB increases in height. For the Higgs analysis, only the new ECAL calibration was used.

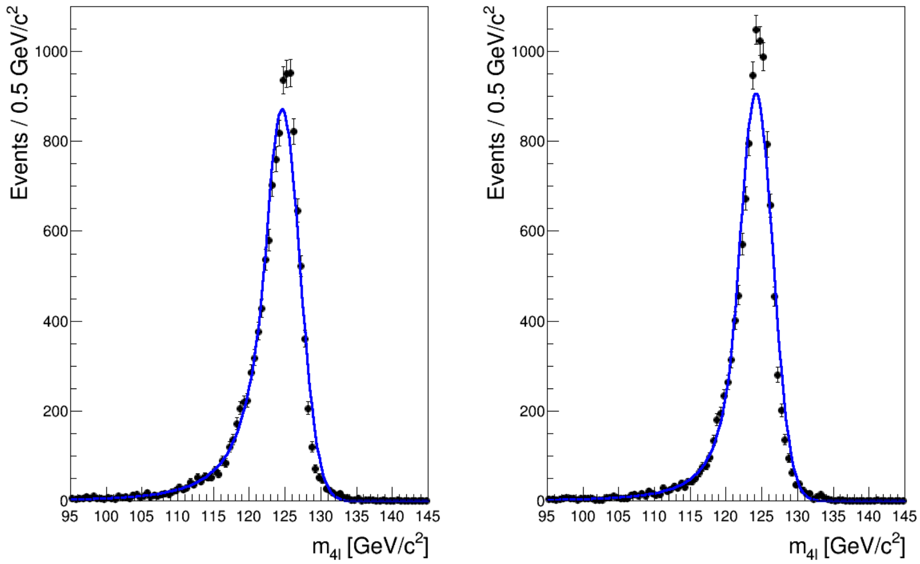

Figure 12: The four lepton invariant mass m_{4l} using $H \to ZZ \to 4e$ decays for the old (*right*) and newer (*right*) ECAL calibrations.

5.3 Applying the Calibration

The calibration constants derived from the $Z \to ll$ analysis were used to calibrate the $H \to ZZ \to 4l$ decays. This was done since the Higgs MC is heavily theory based and no point of comparison can be made between the MC and data at 13TeV. Additionally, since both decays rely on lepton isolation, the detector response is expected to be similar. The effect of applying λ on the mass resolution is shown in Figs. 11-12. The slope of the line for $H \to ZZ \to 4e$ events decreases from 1.300 to 1.117, bringing the MC within the $\pm 20\%$ envelope. For the $H \to ZZ \to 4\mu$ events, the slope decreases from 1.375 to 1.214, bringing the MC to the edge of the envelope.

To propagate the calibration on δm to the uncertainty on the lepton four momenta, the individual lepton p_T need to be scaled. The mass resolution is directly proportional to the uncertainty on p_T, $\delta m \propto \delta p_T$, so the calibration constant λ is applied linearly. The correction factors are summarised in Tab. 1.

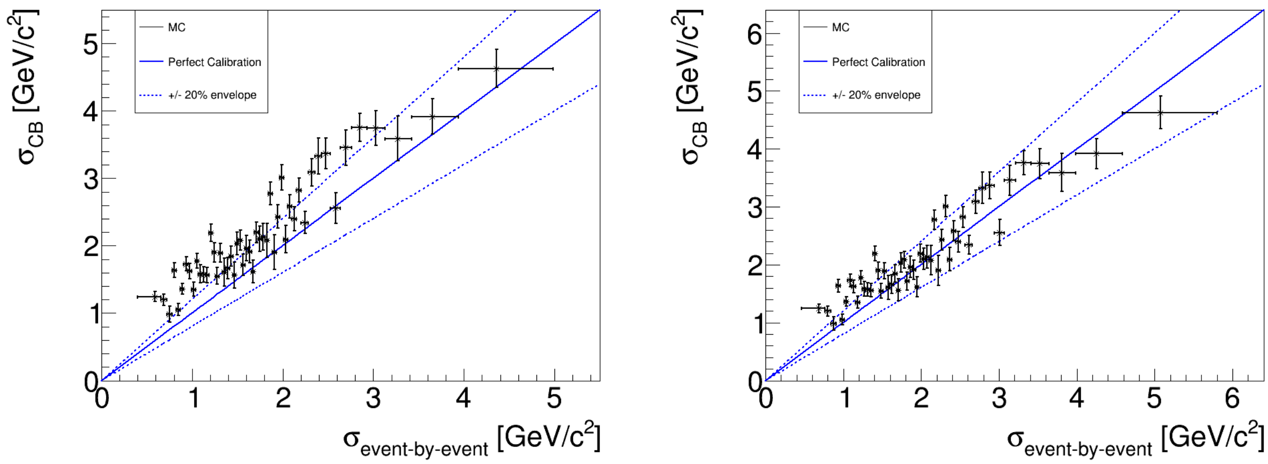

Figure 13: σ_{CB} plotted against $\sigma_{event-by-event}$ for $H \rightarrow ZZ \rightarrow 4e$ MC events before (*left*) and after (*right*) the calibration.

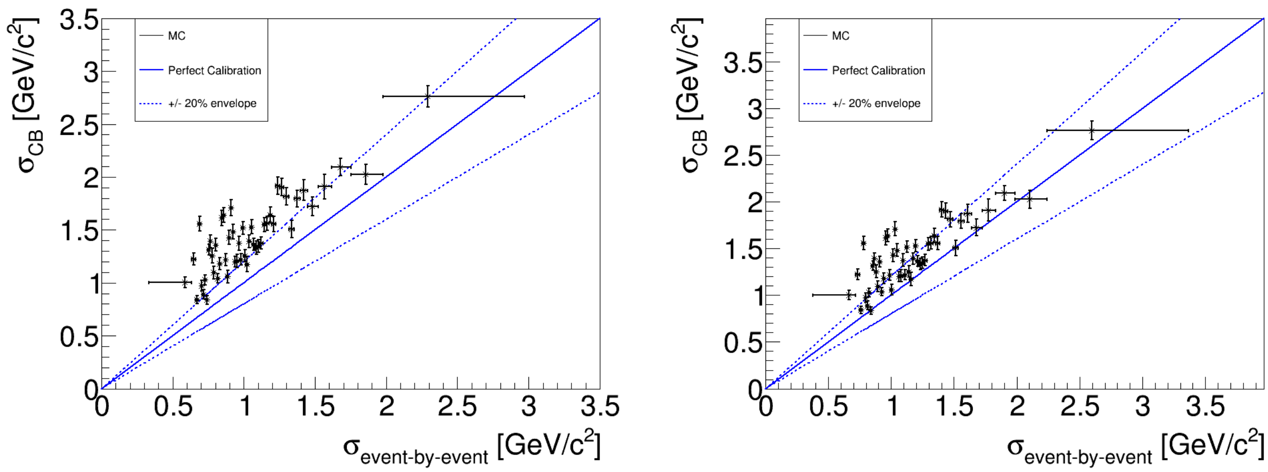

Figure 14: σ_{CB} plotted against $\sigma_{event-by-event}$ for $H \rightarrow ZZ \rightarrow 4\mu$ MC events before (*left*) and after (*right*) the calibration.

	Electron	Muon
λ	1.164	1.133

Table 1: A summary of the correction factors that need to be applied to the uncertainty on the lepton p_T.

6 Conclusion

In conclusion, the spread of the mass resolution across the events is large, making it of interest in a measurement of the Higgs boson mass. The per event mass resolution was calculated by propagating the uncertainty of the lepton p_T. This prediction, when compared to the measured mass resolution, was found to result in an underestimation that required a scale factor λ for correction, $\delta m^{calibrated} = \lambda \times \delta m^{estimated}$. λ was derived from $Z \to ll$ decays. To propagate the correction to the leptons, the scale factors in Tab. 1 should be applied directly to the lepton p_T uncertainty.

References

[1] CMS Collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, Phys. Rev. D 89 (2014) 092007, arXiv:1312.5353.

[2] K.A. Olive *et al.* (Particle Data Group), Chin. Phys. C38, 090001 (2014)

Appendix

Supplementary Plots of CB parameters against $\sigma_{event-by-event}$

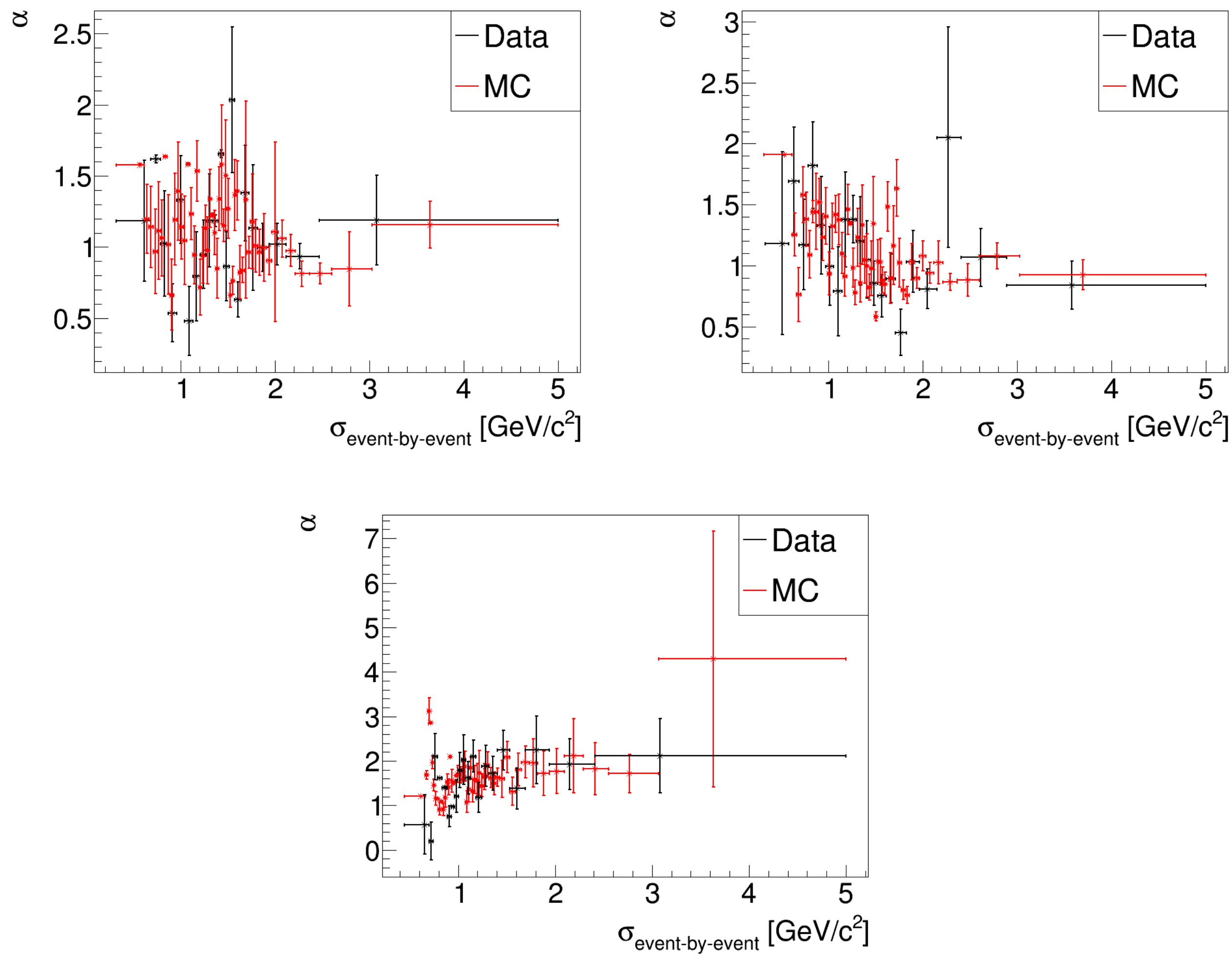

Figure 15: α plotted as a function of $\sigma_{event-by-event}$ using data (black points) and MC (red points) using $Z \to ee$ events (*top left*), $Z \to ee$ events with the new ECAL calibration (*top right*) and $Z \to \mu\mu$ events (*bottom*)

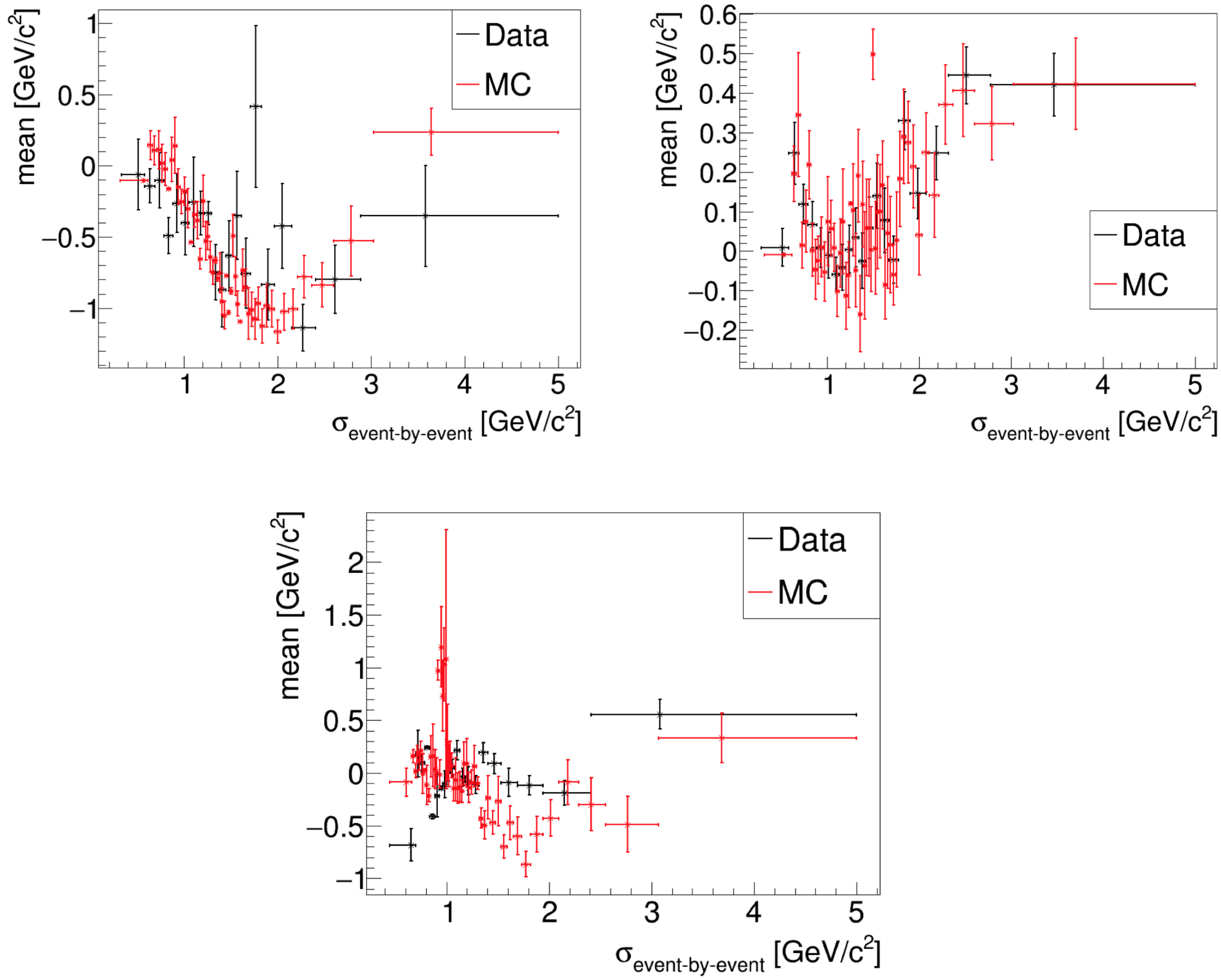

Figure 16: The mean of the CB plotted as a function of $\sigma_{event-by-event}$ for data (black points) and MC (red points) using $Z \rightarrow ee$ events (*top left*), $Z \rightarrow ee$ events with the new ECAL calibration (*top right*) and $Z \rightarrow \mu\mu$ events (*bottom*)