Scale and Smearing corrections for ECAL Run 3 @ low energy

EGM General Meeting - Low energy electrons - 29/11/2024

Elena De Santis - Università di Roma ‘La Sapienza’

Scale and smearing

Scale for data from:

$$1 - \frac{m_{data}}{m_{MC}}$$

Smearing for Monte Carlo from:

$$\frac{\sigma_{data}}{\sigma_{MC}}$$

- Needed to fully characterize ECAL low-energy response (last step of calibration)
- ***Goal:*** derive single electron corrections to match scale in data & smearing in Monte-Carlo
- Right now **not available for** $p_T < 20$ **GeV** as they are extracted from the Z peak

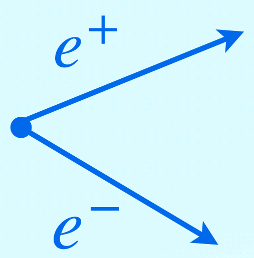
$\downarrow$

Use J/ψ **peak** to extract corrections down to $p_T = 4$ GeV

Ref. Albert M Sirunyan et al. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: JINST 16.05(2021), P05014. doi: 10.1088/1748-0221/16/05/P05014. arXiv: 2012.06888 [hep-ex].

Dataset

- B-Parking Double Electron Low Mass (*/ParkingDoubleElectronLowMass%d/Run2022F-22Sep2023-v1/MINIAOD*)
- 17.78 fb^{-1} 2022 part F data
- Double electron trigger:
  - Vertexed electrons
  - Several symmetric p_T thresholds (lowest at 4 GeV)
  - $\eta < 1.22$ (barrel only)
  - $m(e^+e^-) < 6$ GeV



- Ntuples for the present analysis obtained through a customization of the tool used for central scale and smearing on the Z peak

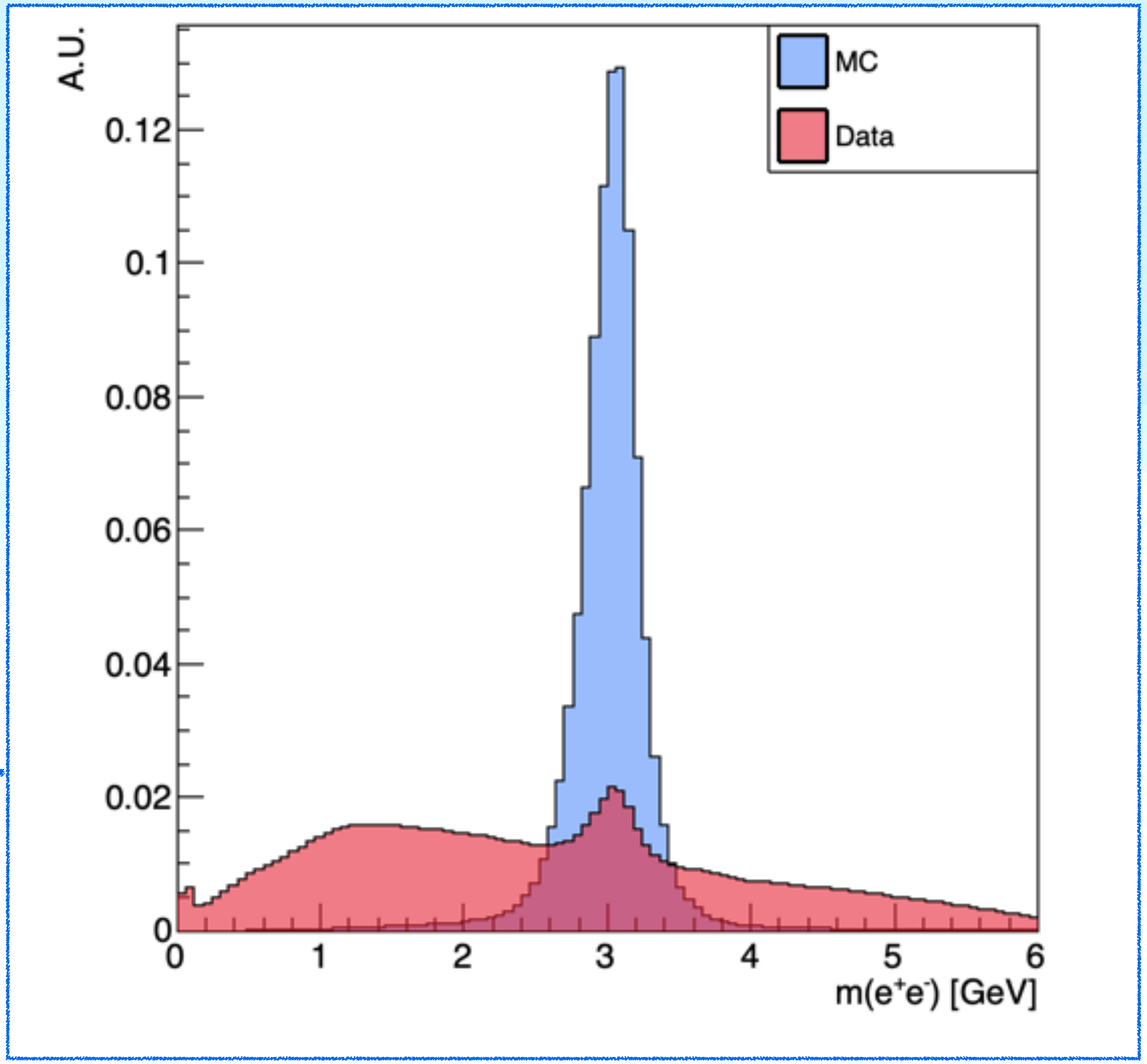
https://github.com/Raffaella07/ScaleAndSmearingTools/tree/jpsi

Monte Carlo

- $B \to K\,J/\psi$ sample (*/BuToKJPsi_JPsiToEE_SoftQCD_TuneCP5_13p6TeV_pythia8-evtgen/Run3Summer22EEMiniAODv4-130X_mcRun3_2022_realistic_postEE_v6-v2/MINIAODSIM*)
- Double electron trigger:
  - Parking thresholds not simulated
  - $\eta < 1.22$ (barrel only)
  - $m(e^+e^-) < 6$ GeV

Invariant mass distribution

- Use invariant mass reconstructed using only ECAL regressed energy estimate (no track info)
- Plot is inclusive in electrons p_T
- Need to **filter background** in data to properly model the signal



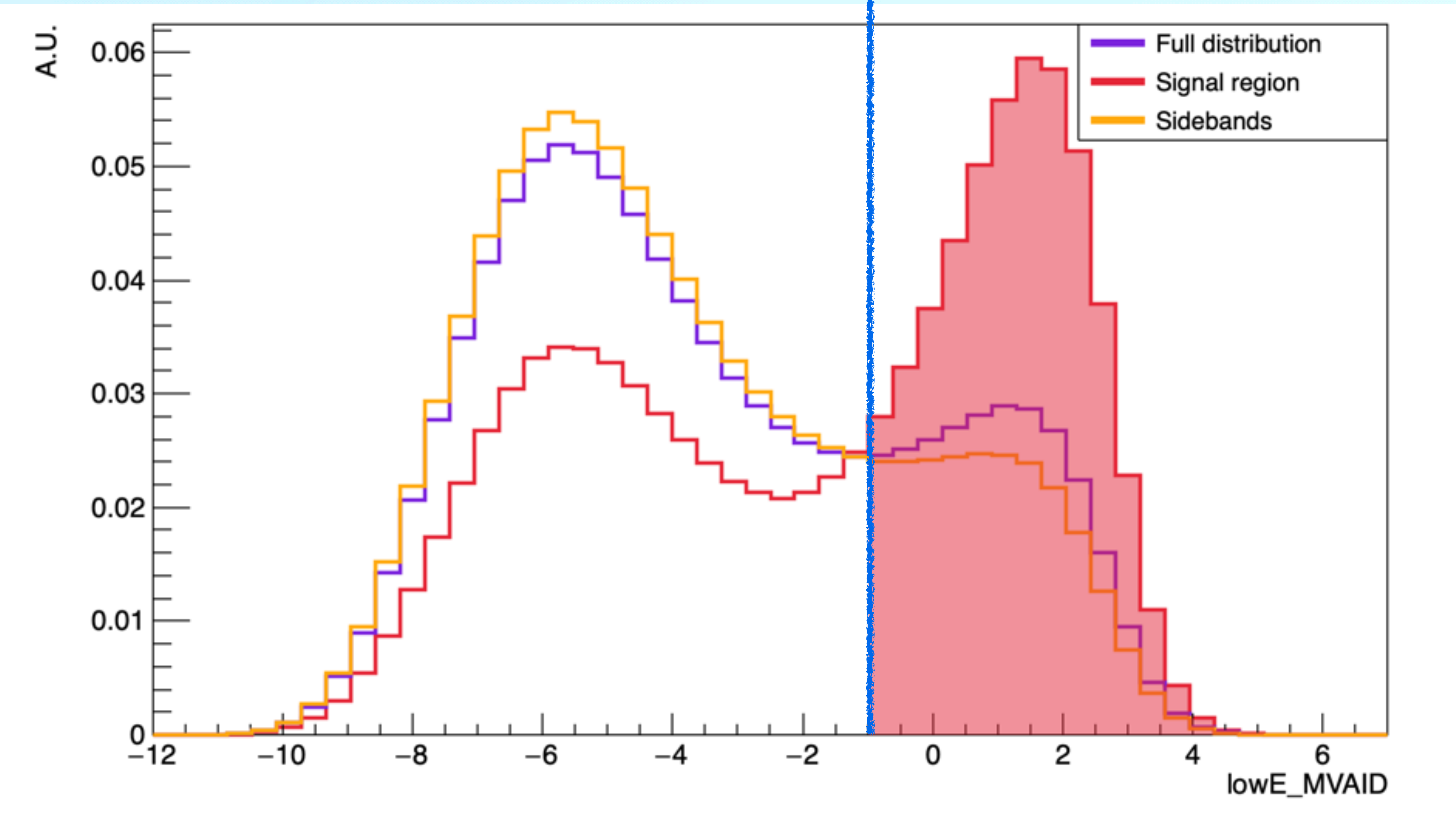
Normalized distribution of ECAL invariant mass

Selection with low energy multivariate ID

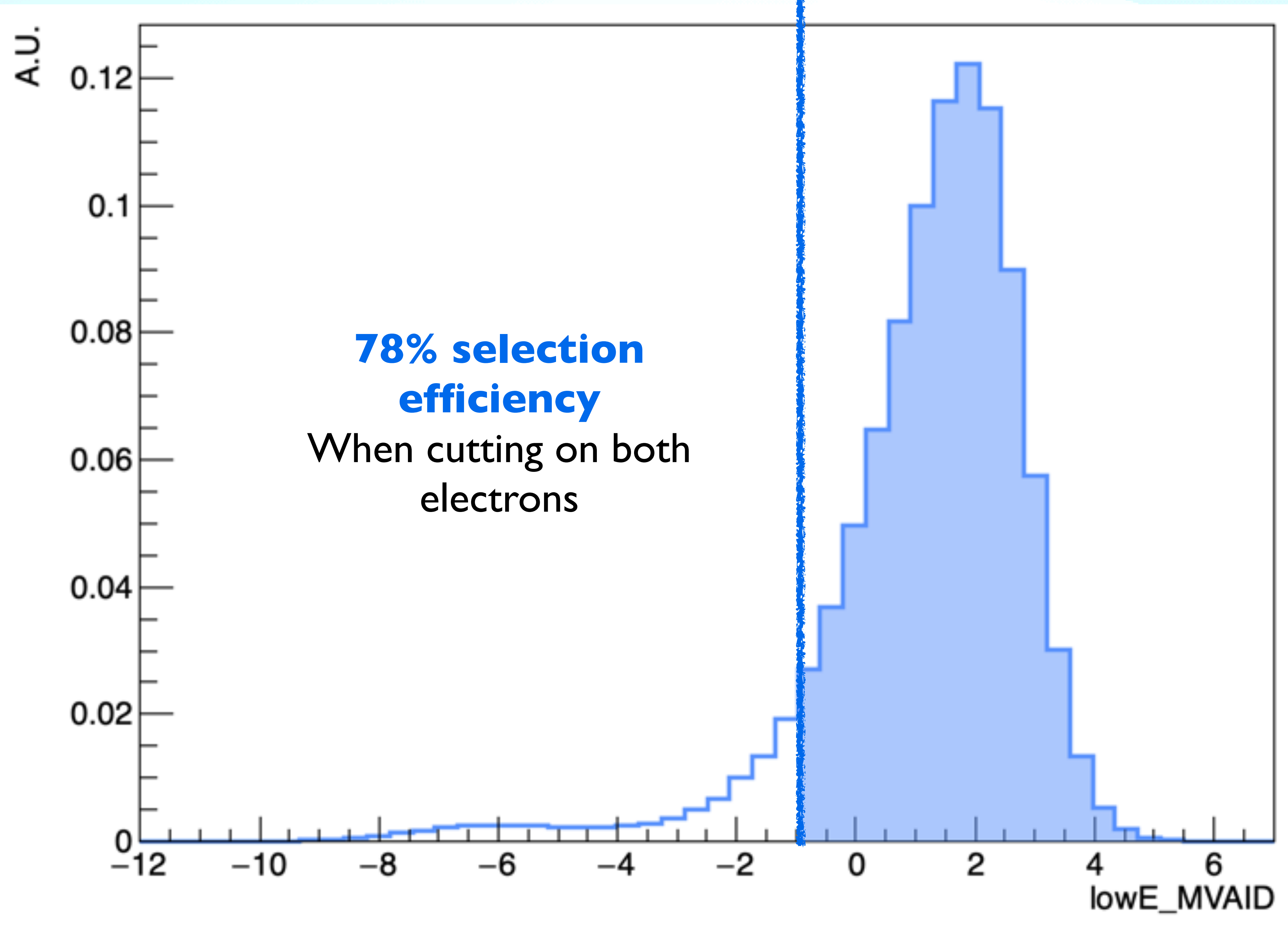
- Using 2018 Low-energy retrained MVA ID
- Plan to move to 2022 central EGM ID or novel PF low-energy retrain (*see Peter's talk later*)

Working point

ID distribution in data

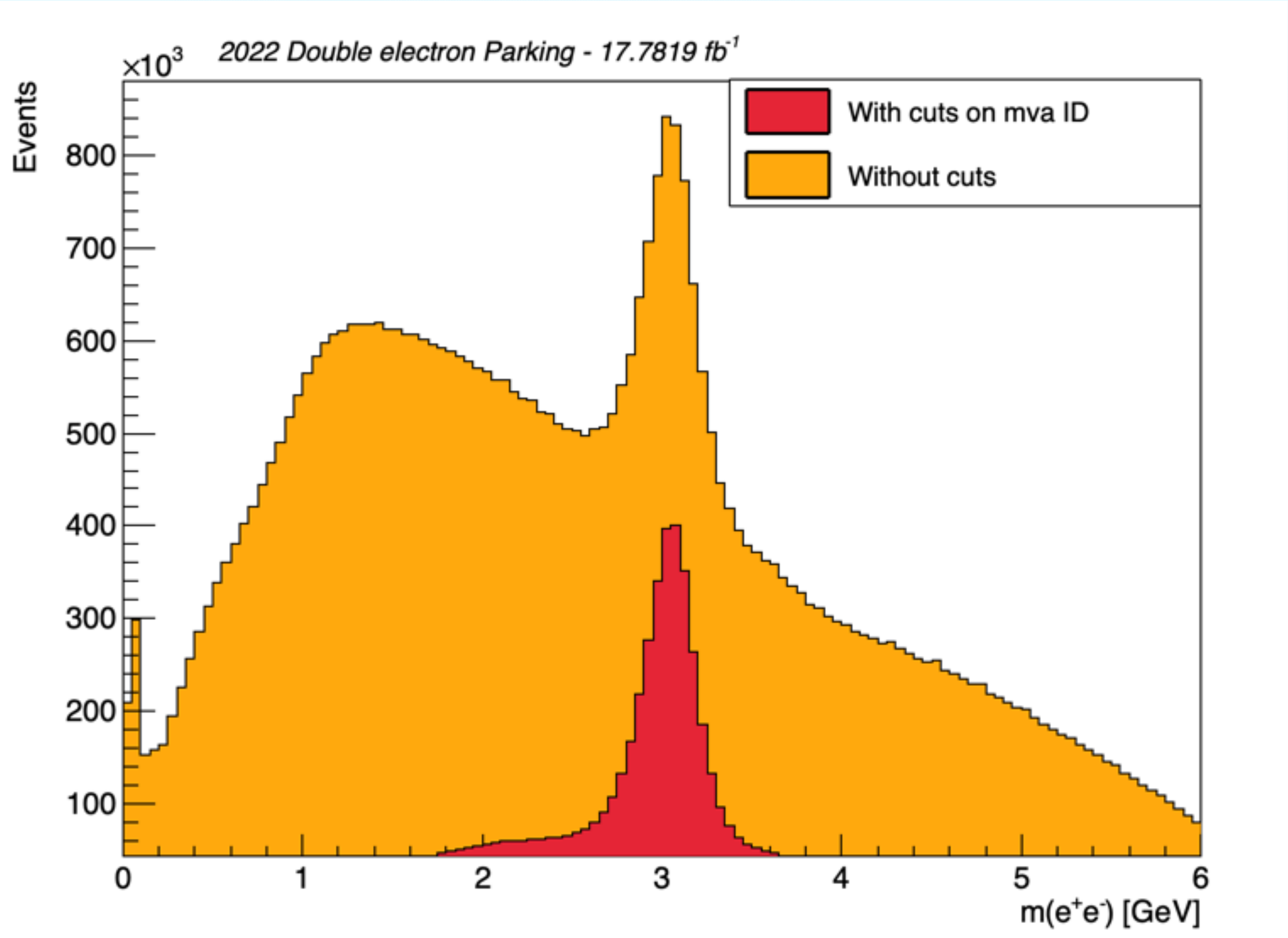


ID distribution in MC



Signal region: 2.7 GeV < $m(e^+e^-)$ ***< 3.3 GeV***

Invariant mass shape after selection



Inclusive in p_T

Selection cut:

- MVA ID > -1.0
- $\eta < 1.22,\ p_T > 4$ GeV
- Triggering electrons
- $\Delta R > 0.1$
- Opposite charge

On both electrons

Reweighting on $p_T^{(J/\psi)}$

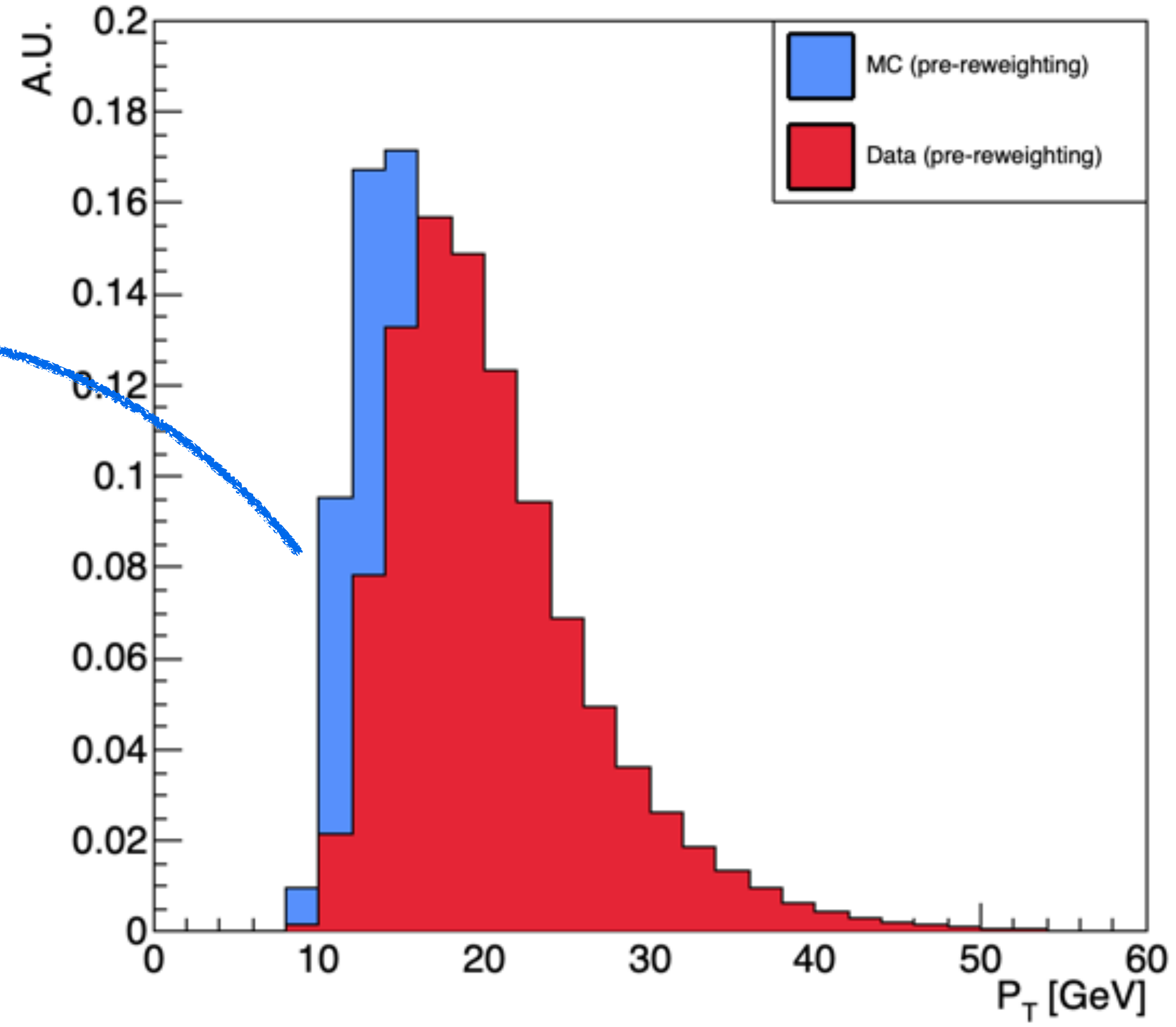
- In data different triggers with different p_T tresholds and prescales are opened when needed
- This is not simulated in MC

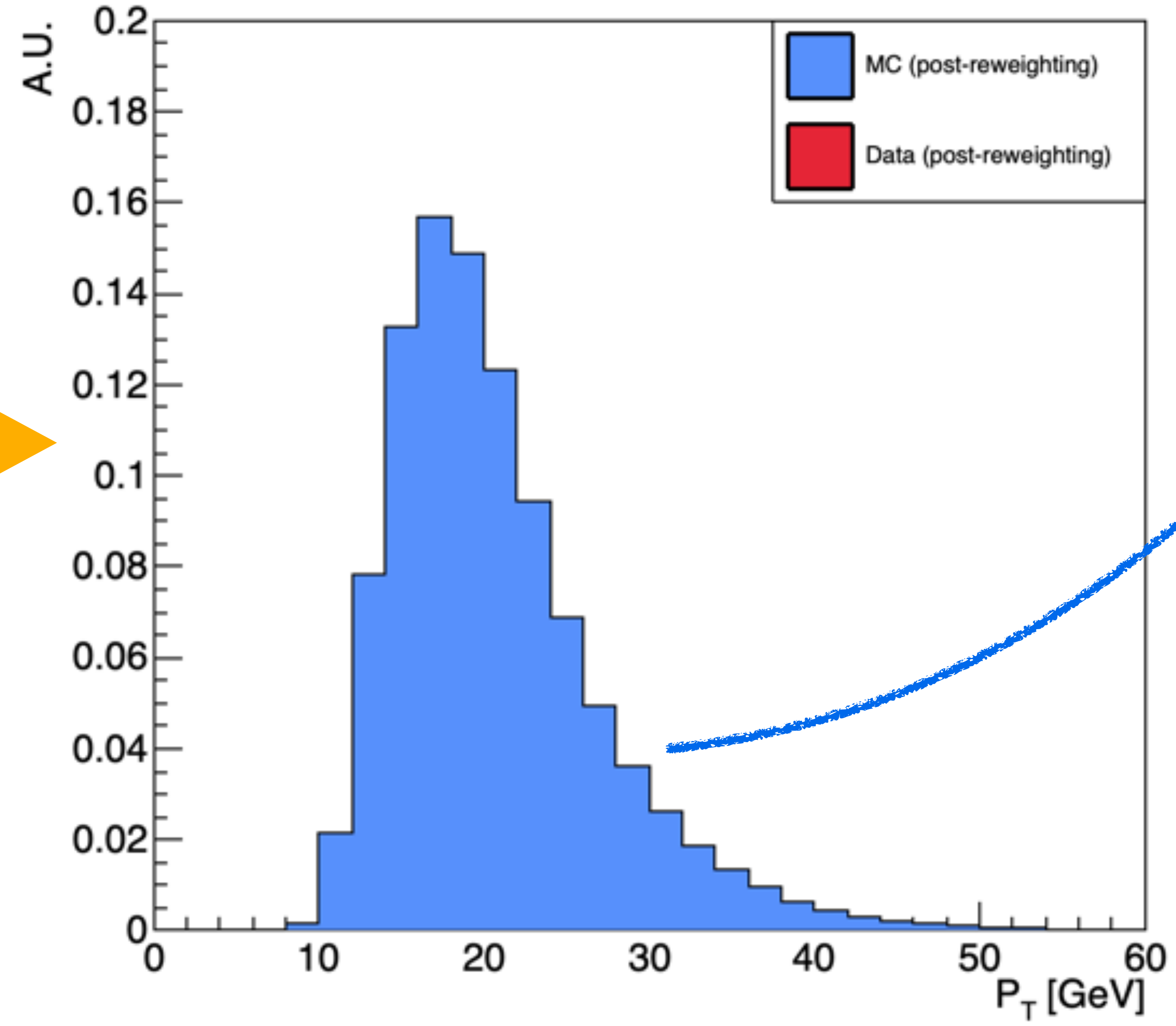
Reweight MC distribution of $p_T^{(J/\psi)}$ match the data sample

$p_T^{(J/\Psi)}$ = transverse momentum of the J/ψ

The ratio $\frac{Data}{MC}$ is computed binwise

A **weight coefficient** is obtained for each bin





If weights are reapplied to MC the two distributions coincide

For each event, weights are applied to the MC invariant mass distribution depending on the value of $p_T^{(J/\Psi)}$ in that event

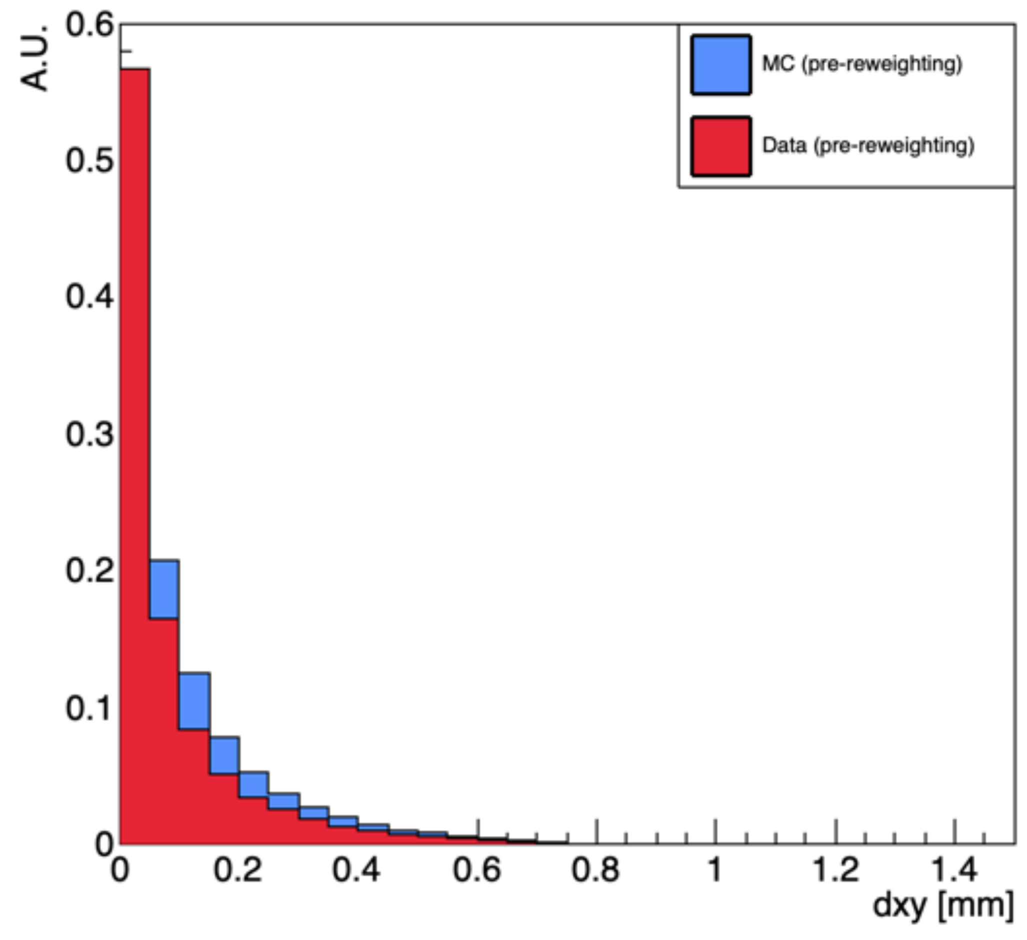
Additional reweightings

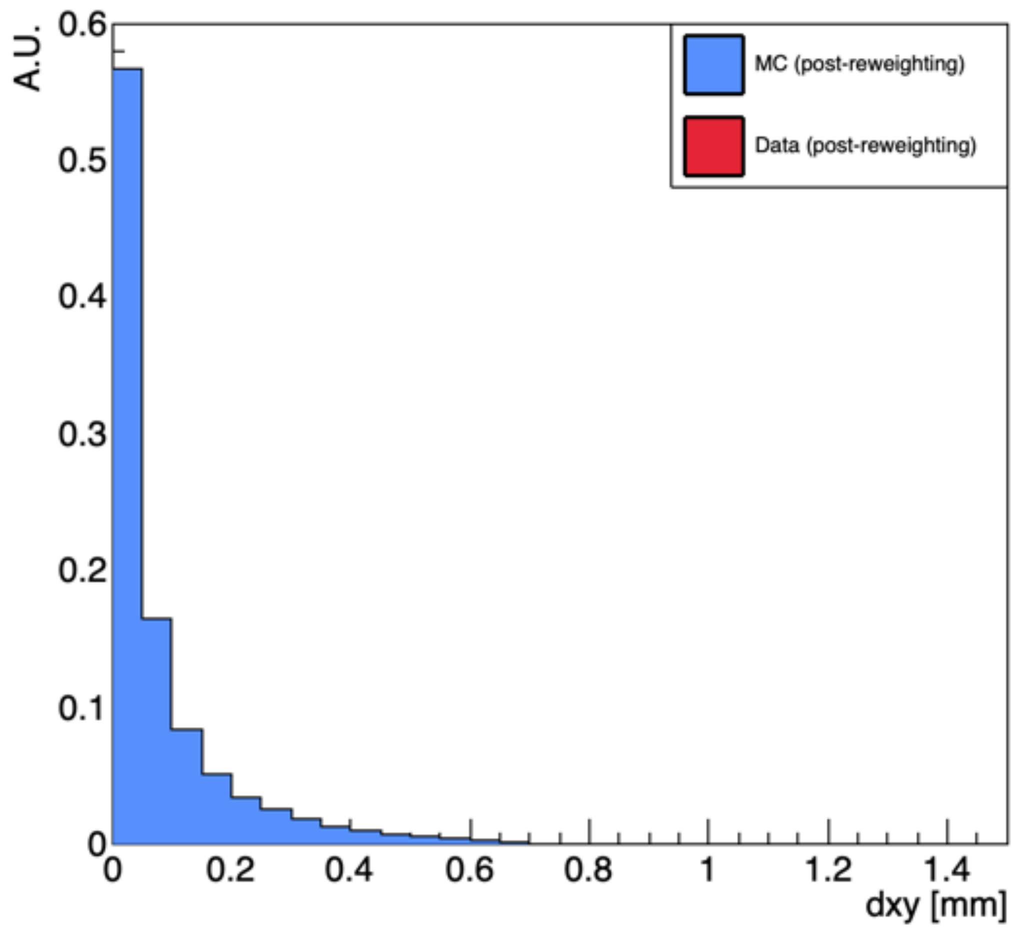
Same procedure applied to other distributions that exhibit differences between data and MC

Transverse impact parameter (d_{xy})

- MC only includes the process $B \to K\, J/\psi$
- Data sample also includes prompt J/ψ

Reweight MC distribution of d_{xy} (more displaced) to match the data sample

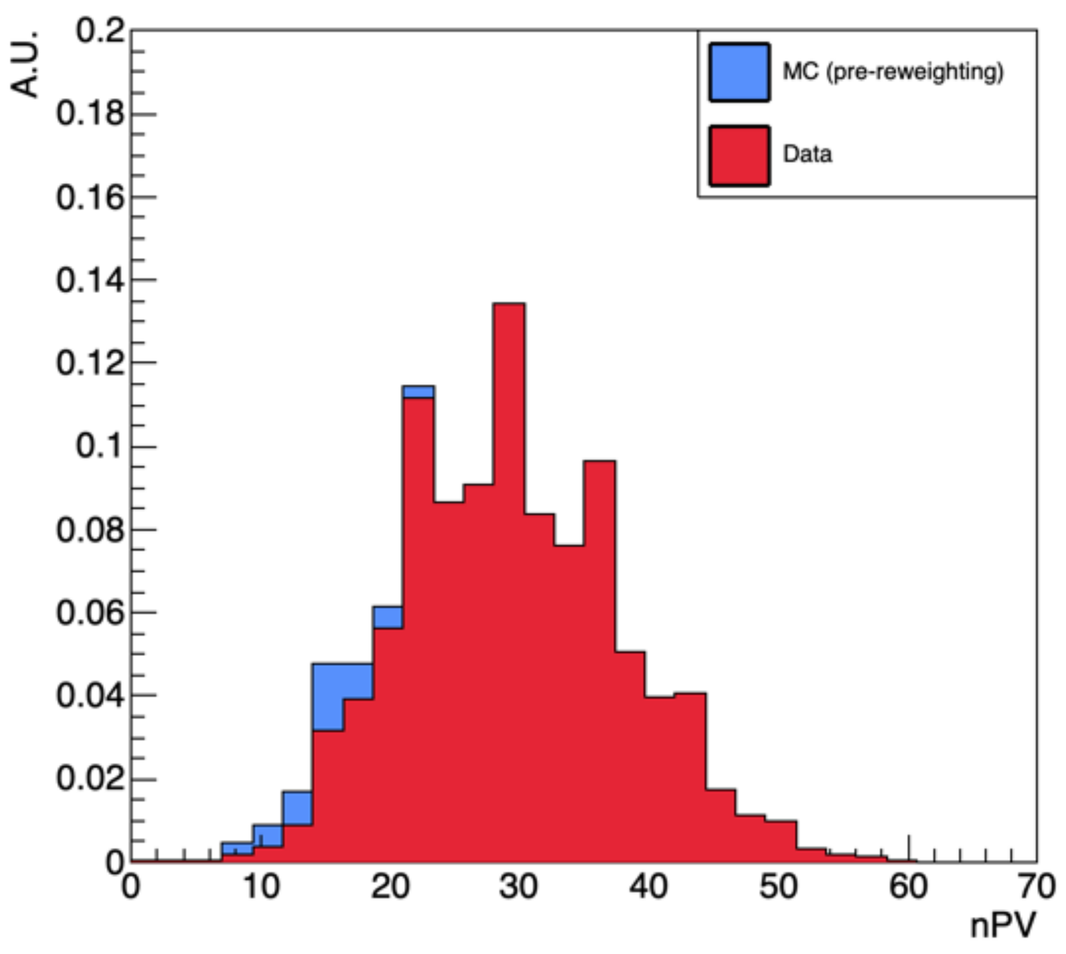


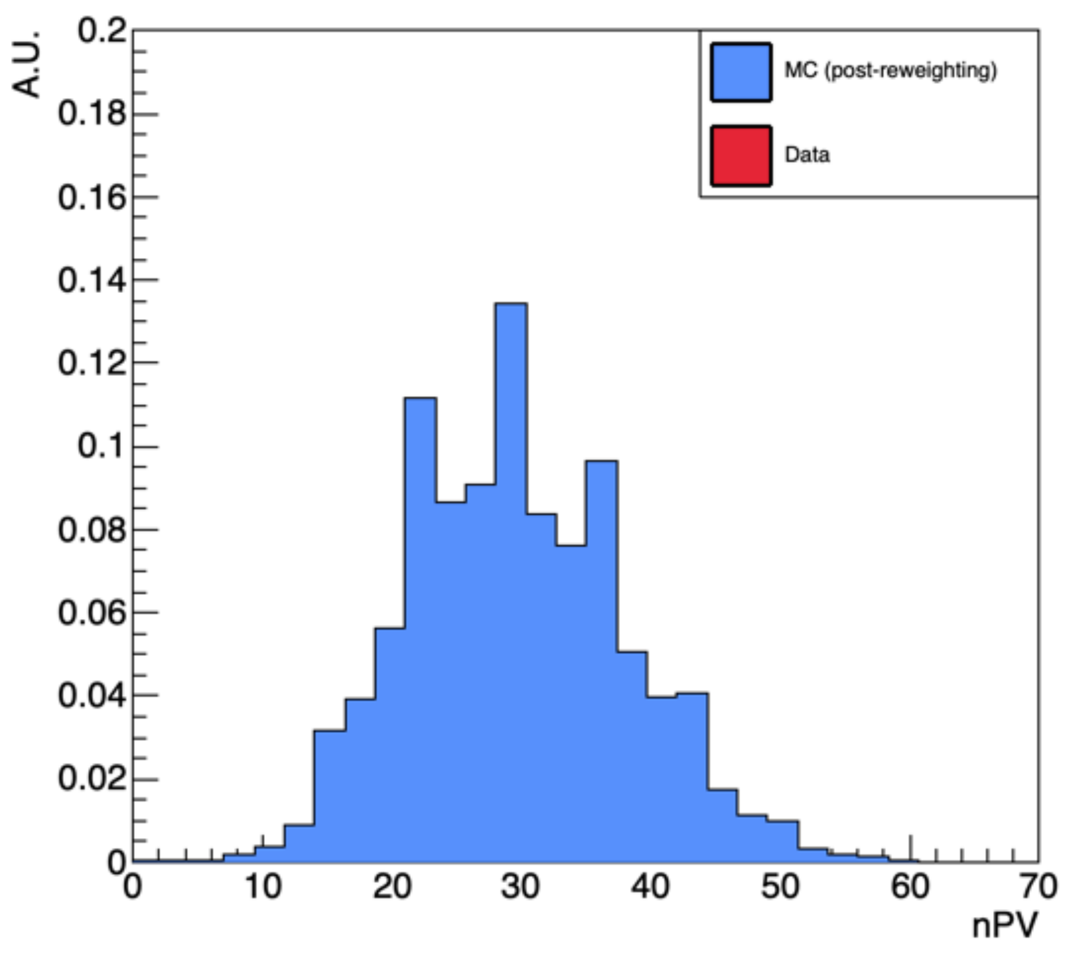


N. of primary vertices

Slightly different pileup between data and MC

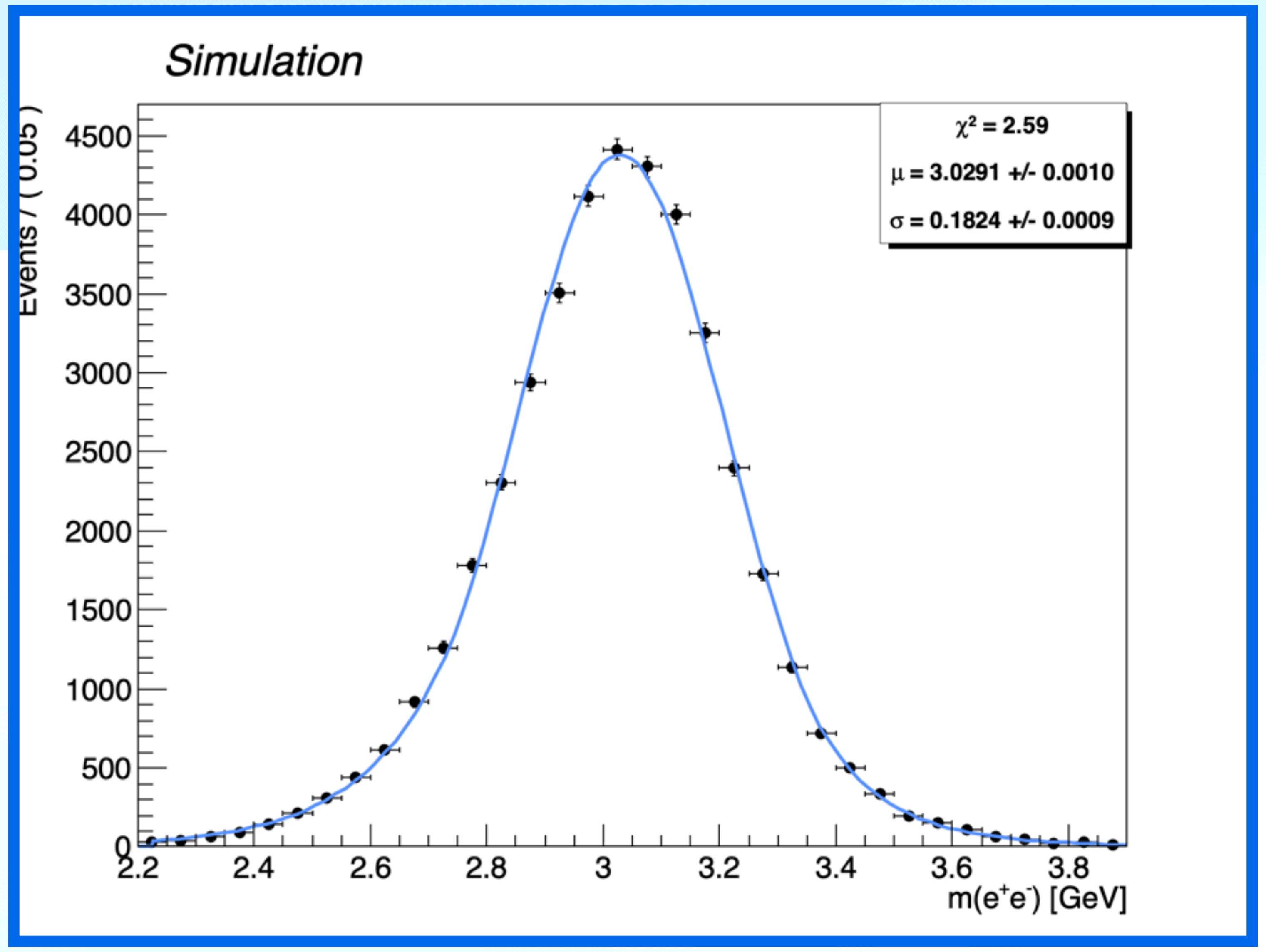
Again, MC is reweighted on data



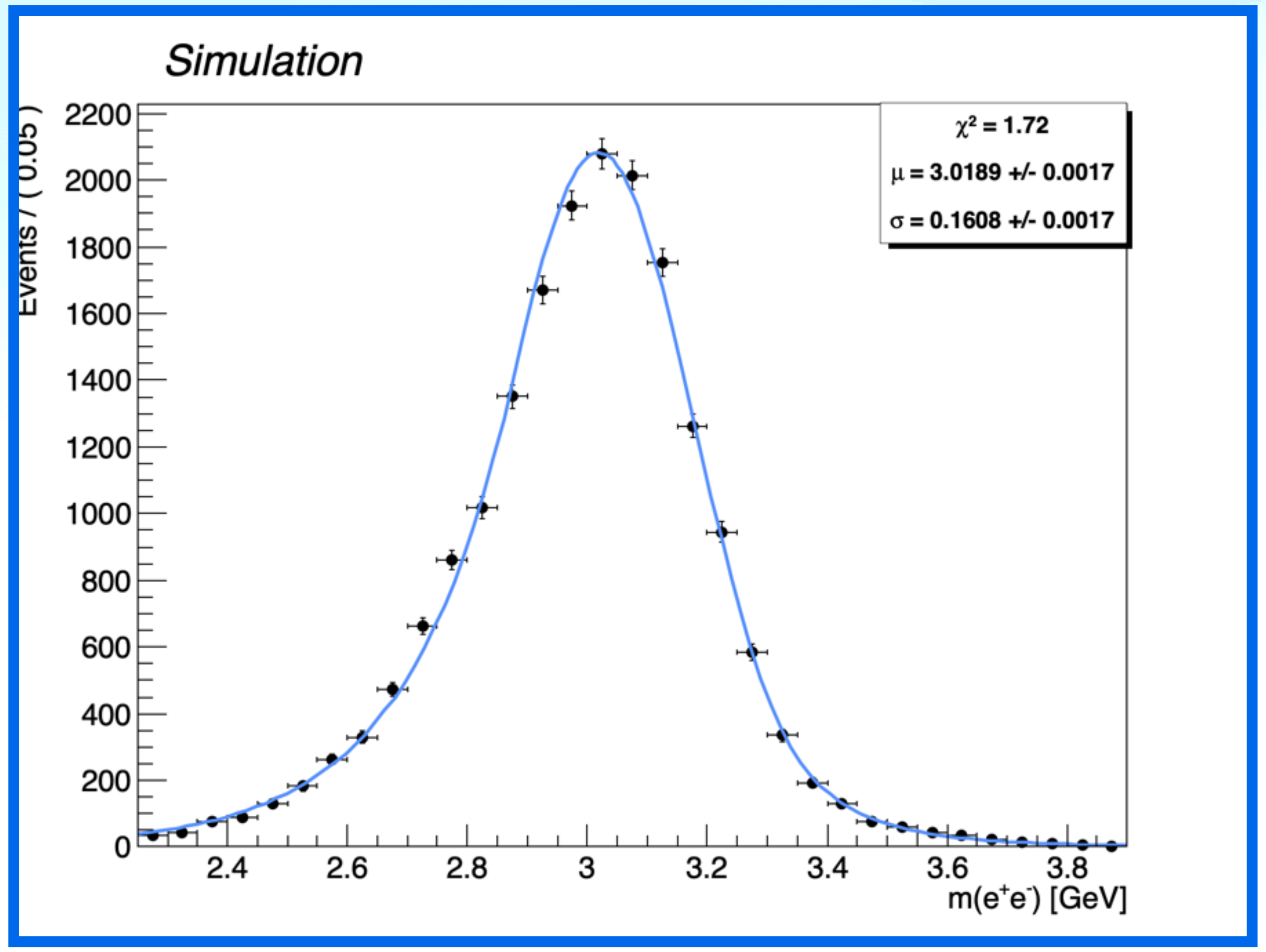


Fit on Monte Carlo

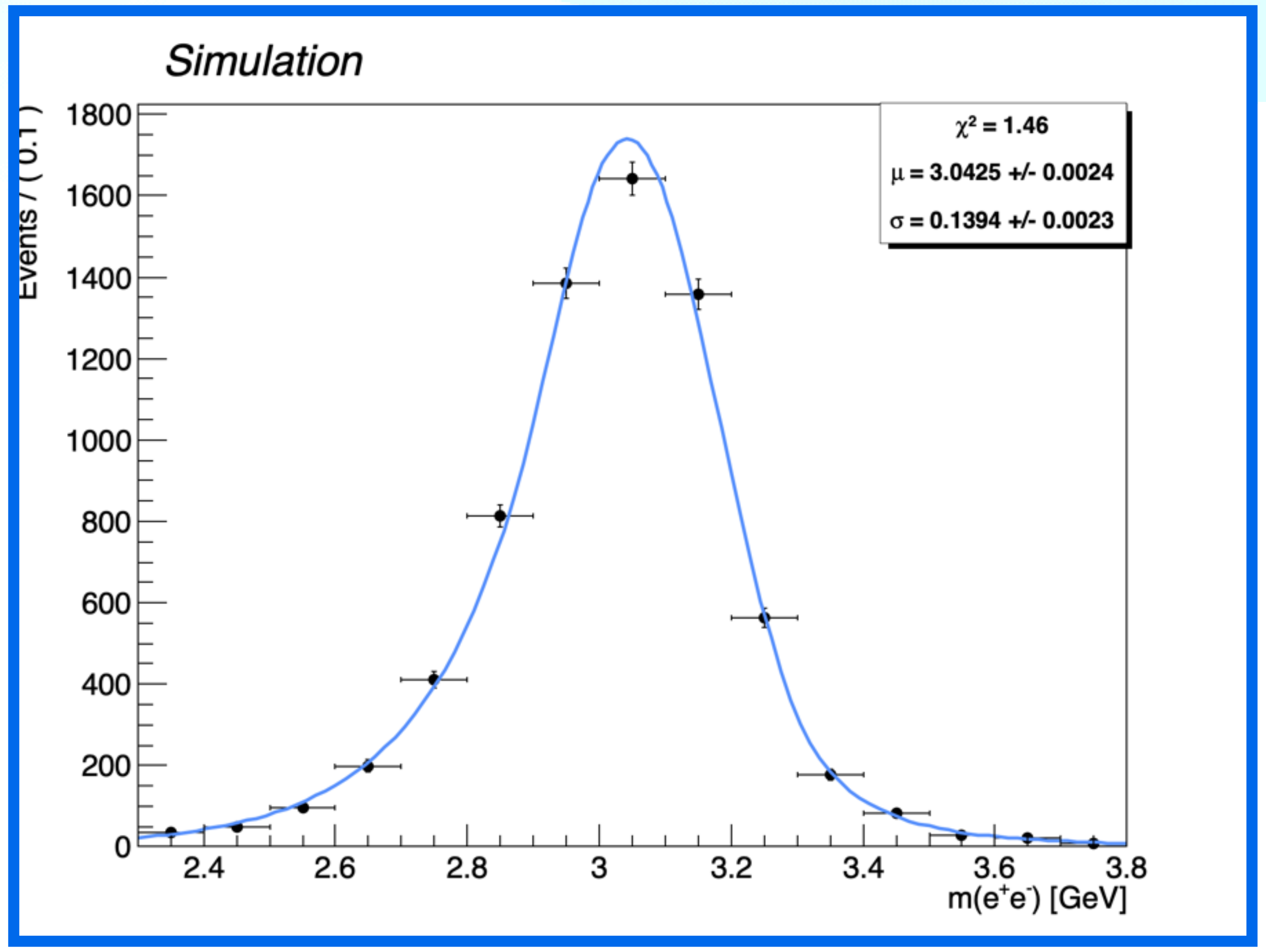
- Model: two-tailed Crystal Ball
- 6 Bins in p_T (transverse momentum of the single electron)
- Only events with both electrons in the same bin ⟶ no bias from other p_T bins



4 GeV $< p_T < 7$ GeV



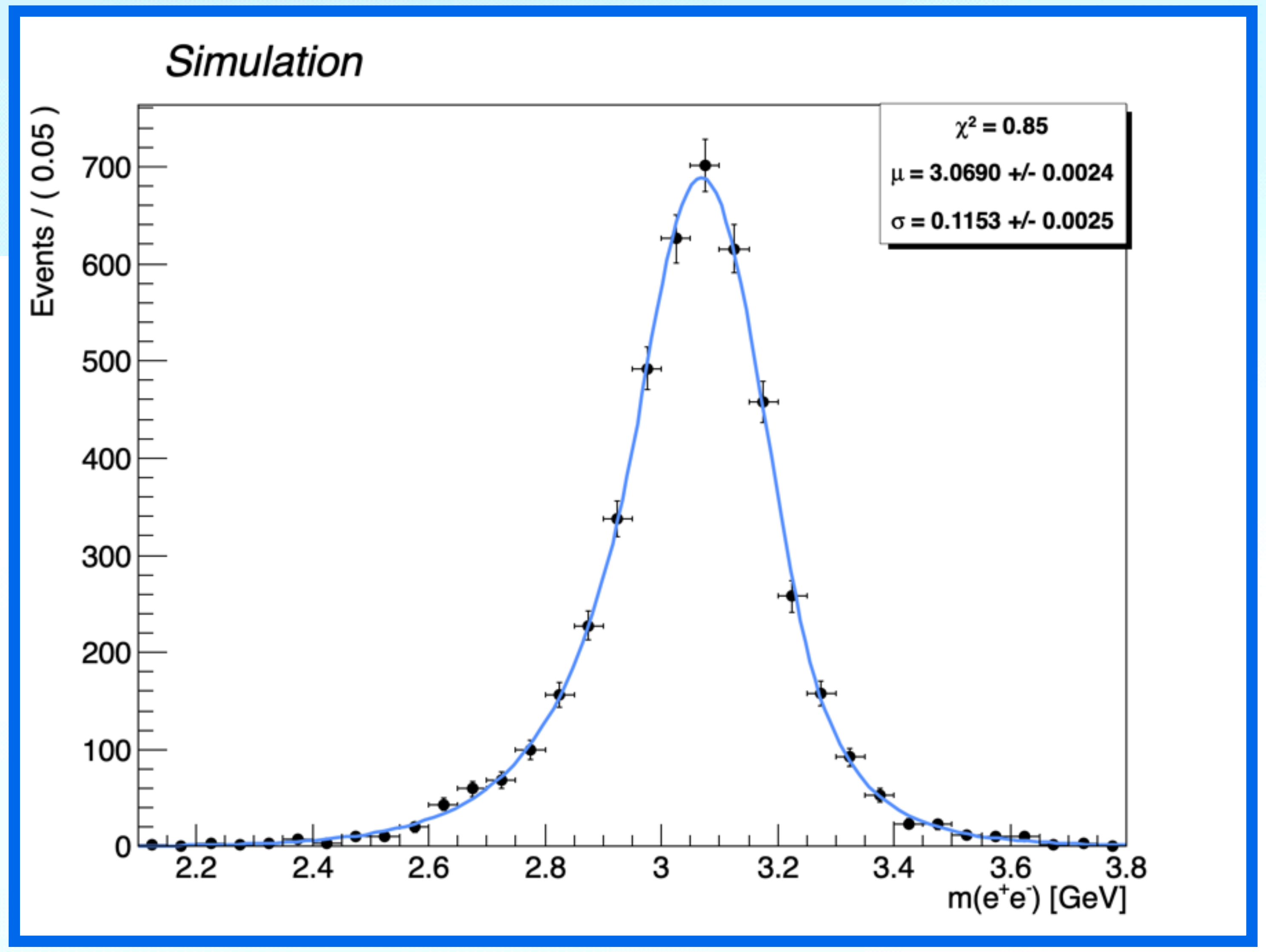
7 GeV $< p_T < 9$ GeV



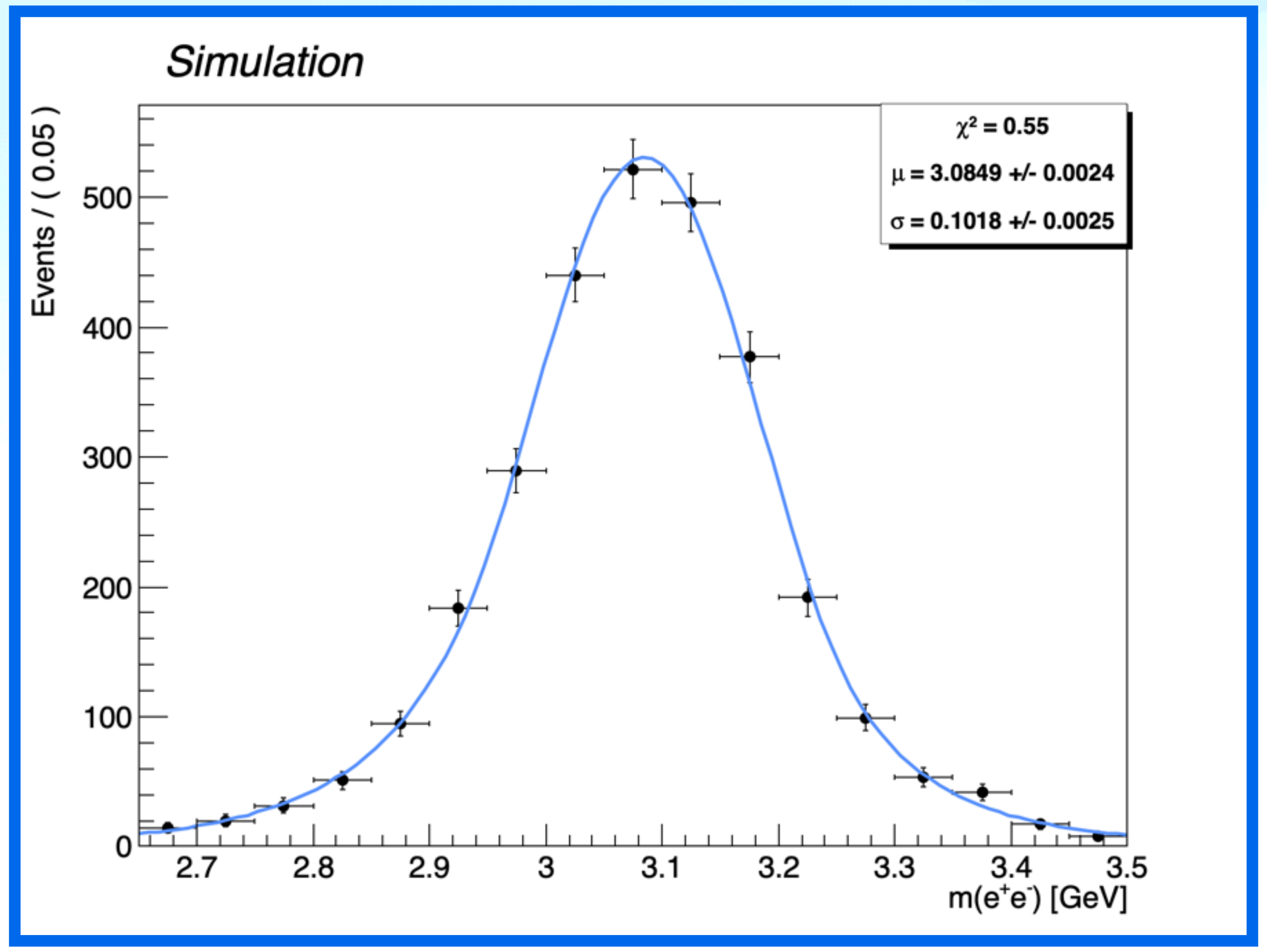
9 GeV $< p_T < 11$ GeV

Fit on Monte Carlo

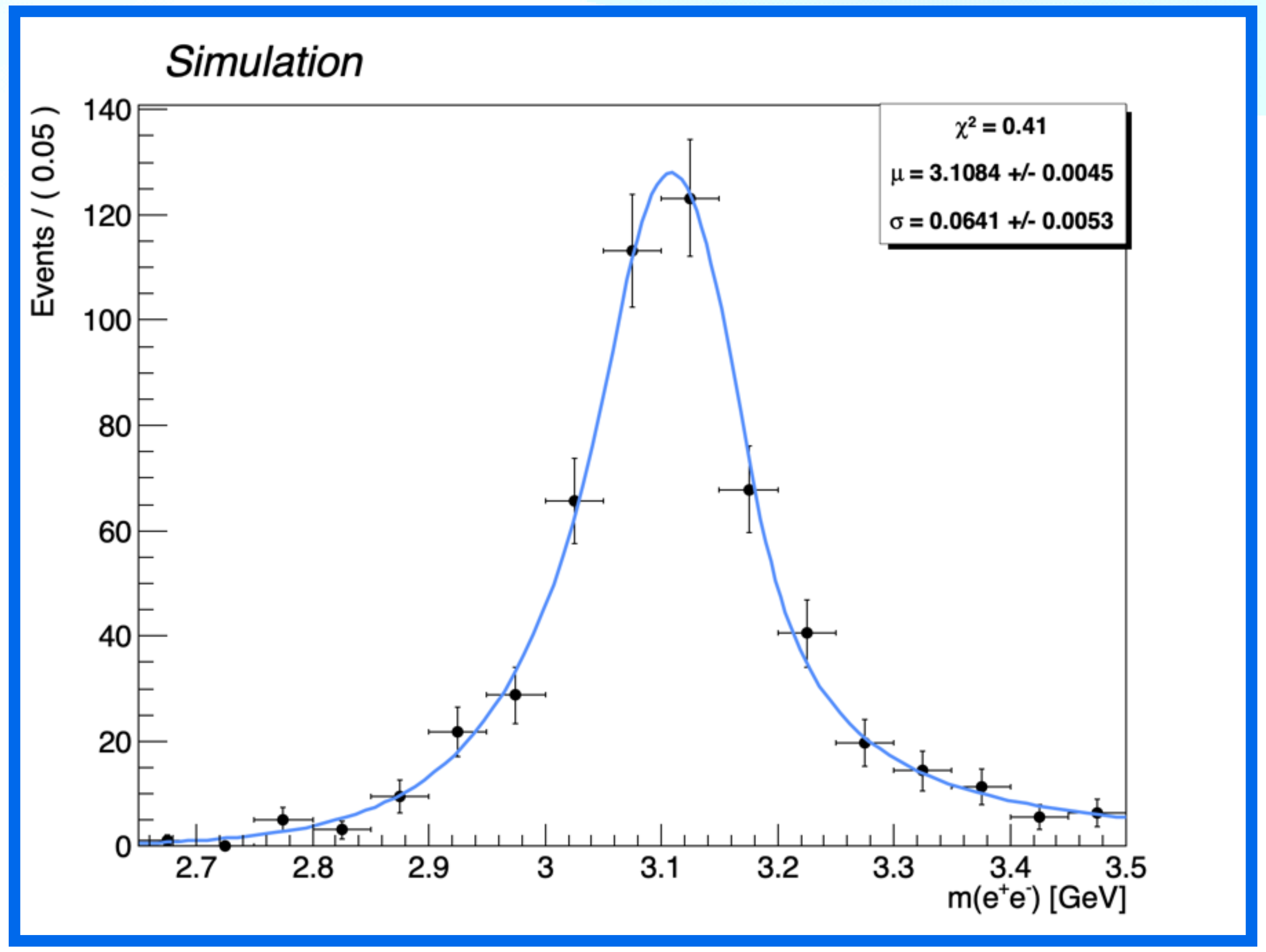
- Model: two-tailed Crystal Ball
- 6 Bins in p_T (transverse momentum of the single electron)
- Only events with both electrons in the same bin → no bias from other p_T bins



11 GeV $< p_T <$ 14 GeV



14 GeV $< p_T <$ 20 GeV



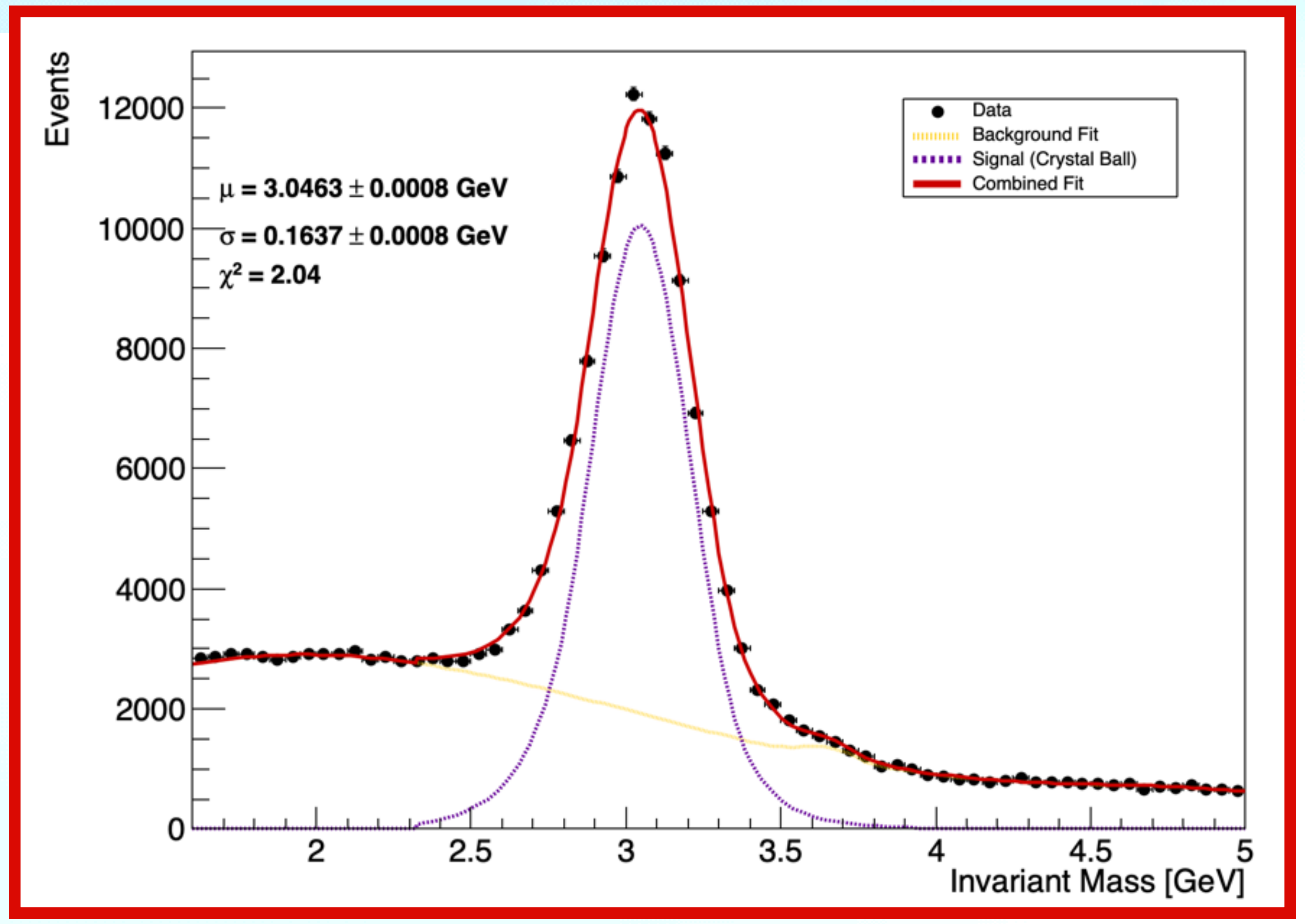
20 GeV $< p_T <$ 40 GeV

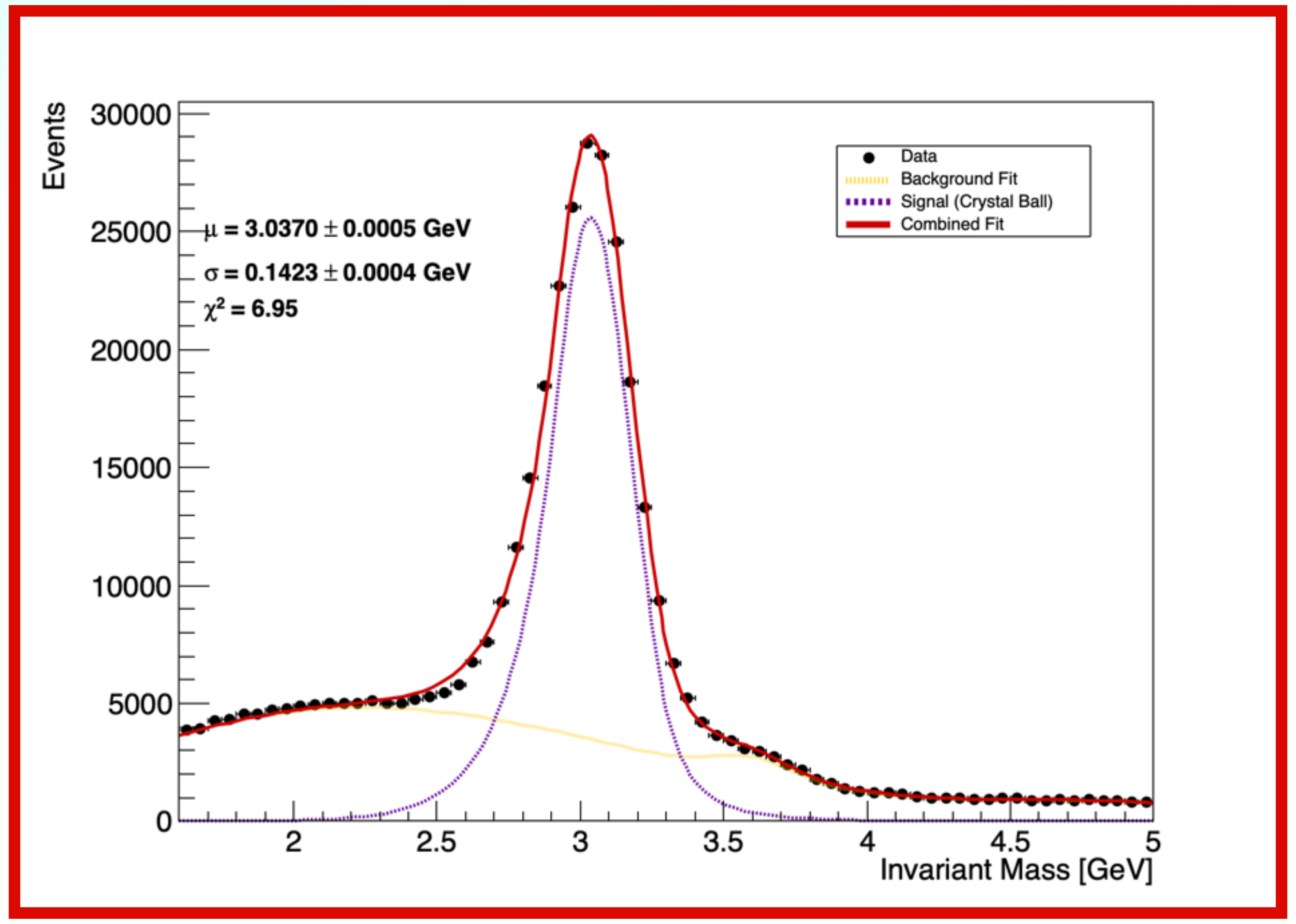
Fit on Data

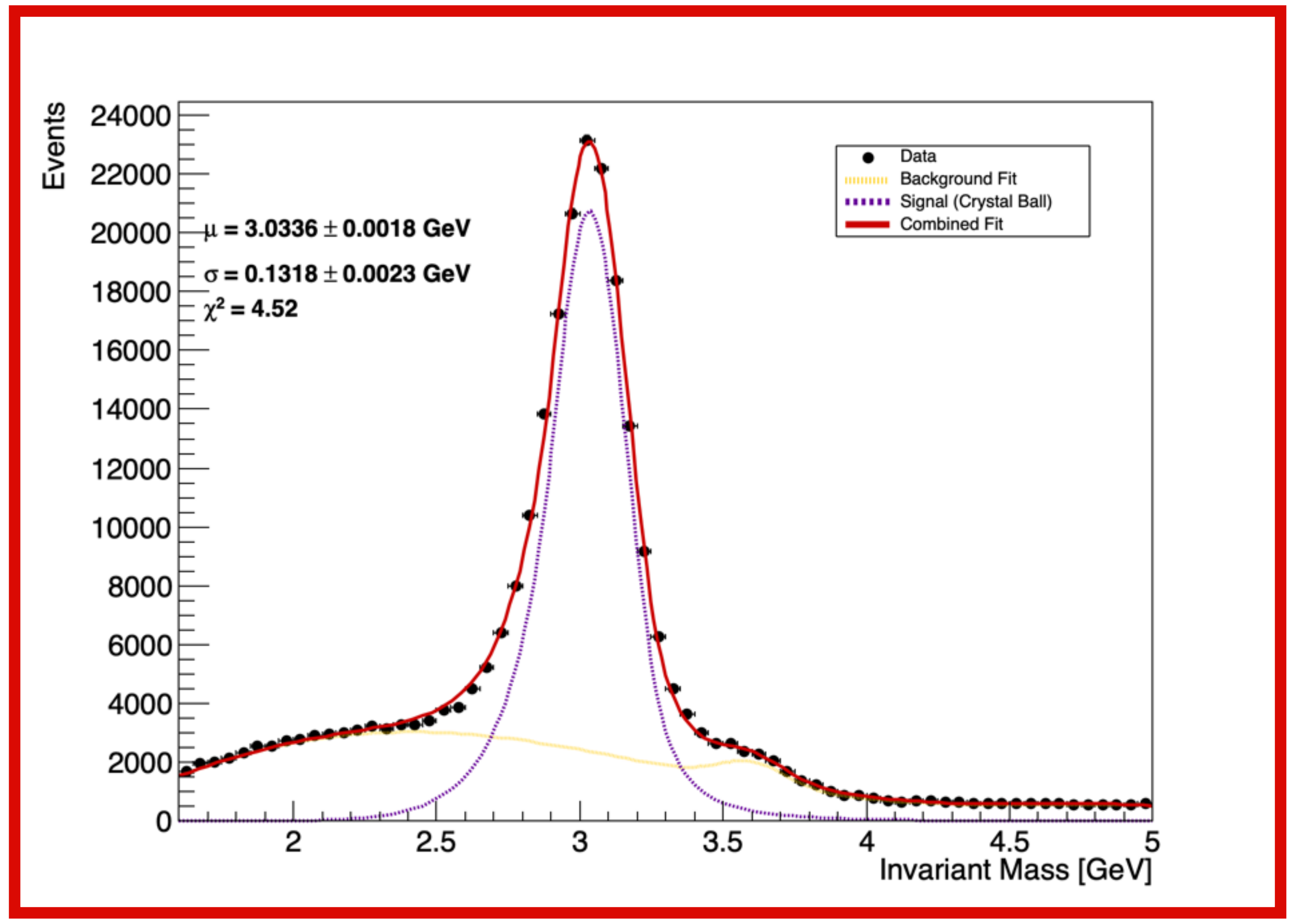
- Background modeled with a 4-th degree polynomial (+ gaussian for $\psi(2s)$)
- Background fit on sidebands, then constrain parameters and fit the signal (with tails parameters constrained to MC values)
- The ID trained on low energies allows uniform behaviour in different p_T intervals (unlike cut-based IDs)

↓

Better fit performance







4 GeV $< p_T <$ 7 GeV

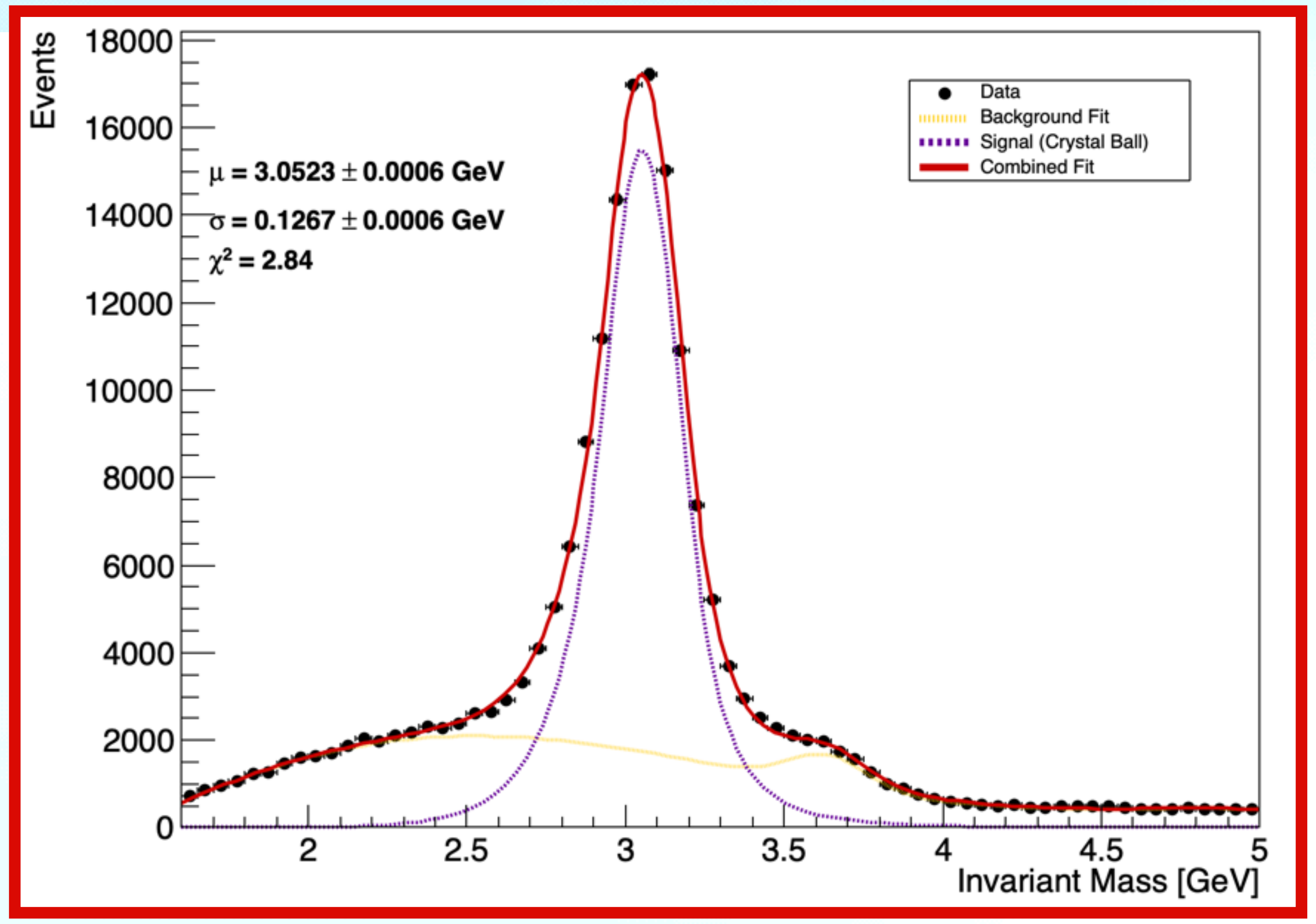
7 GeV $< p_T <$ 9 GeV

9 GeV $< p_T <$ 11 GeV

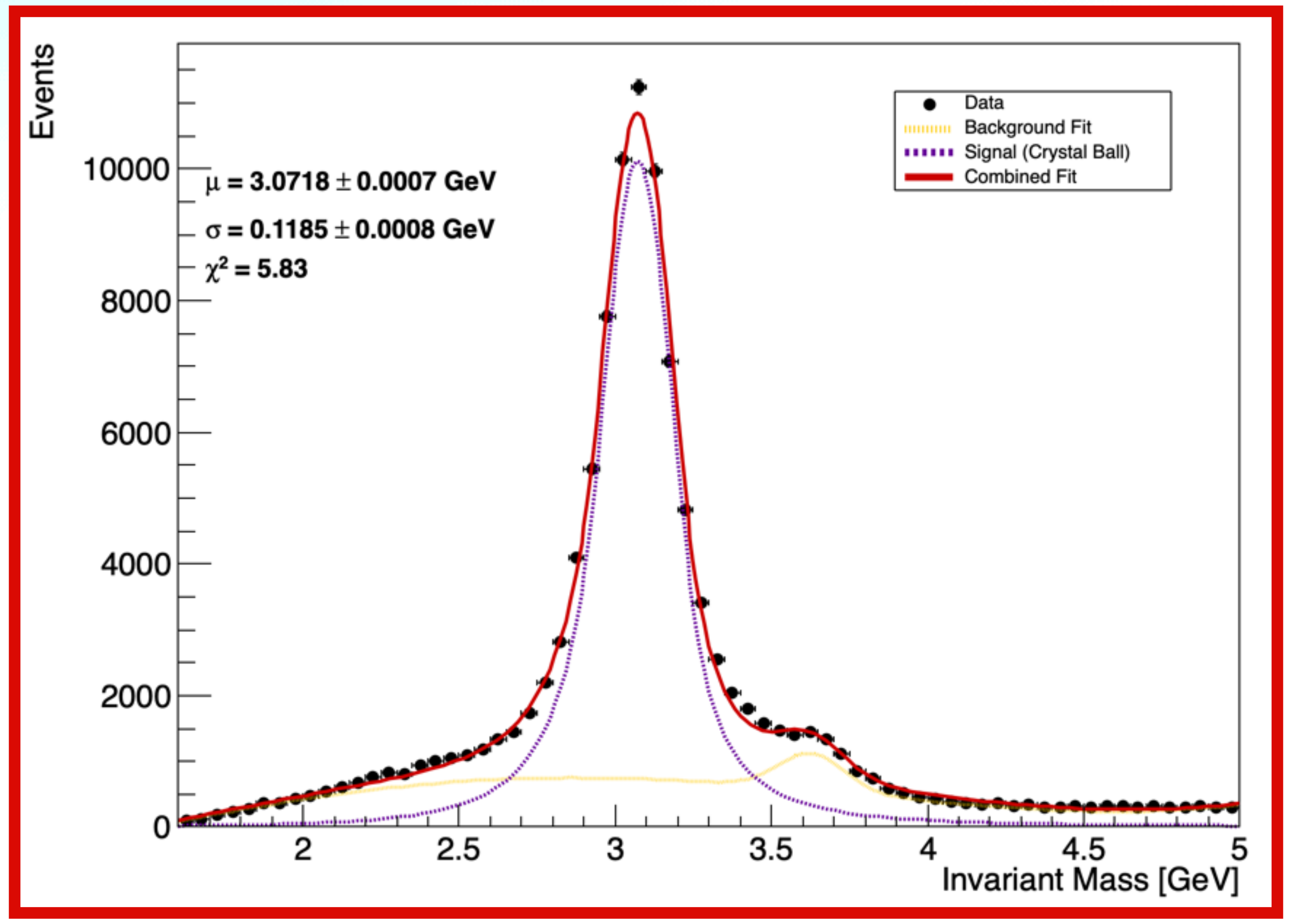
Fit on Data

- Background modeled with a 4-th degree polynomial (+ gaussian for $\psi(2s)$)
- Background fit on sidebands, then constrain parameters and fit the signal (with tails parameters constrained to MC values)
- The ID trained on low energies allows uniform behaviour in different p_T intervals (unlike cut-based IDs)

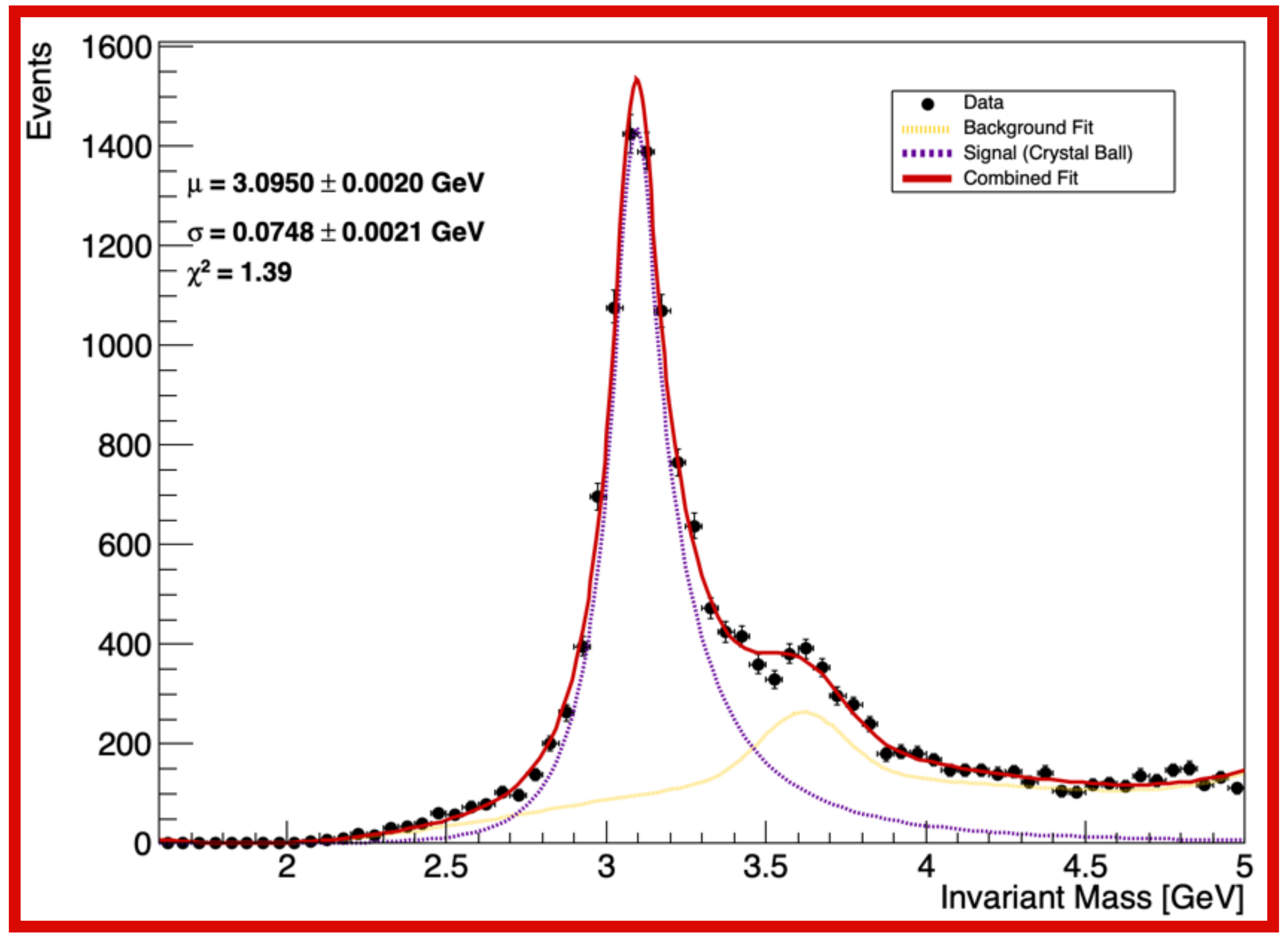
Better fit performance



11 GeV $< p_T <$ 14 GeV



14 GeV $< p_T <$ 20 GeV

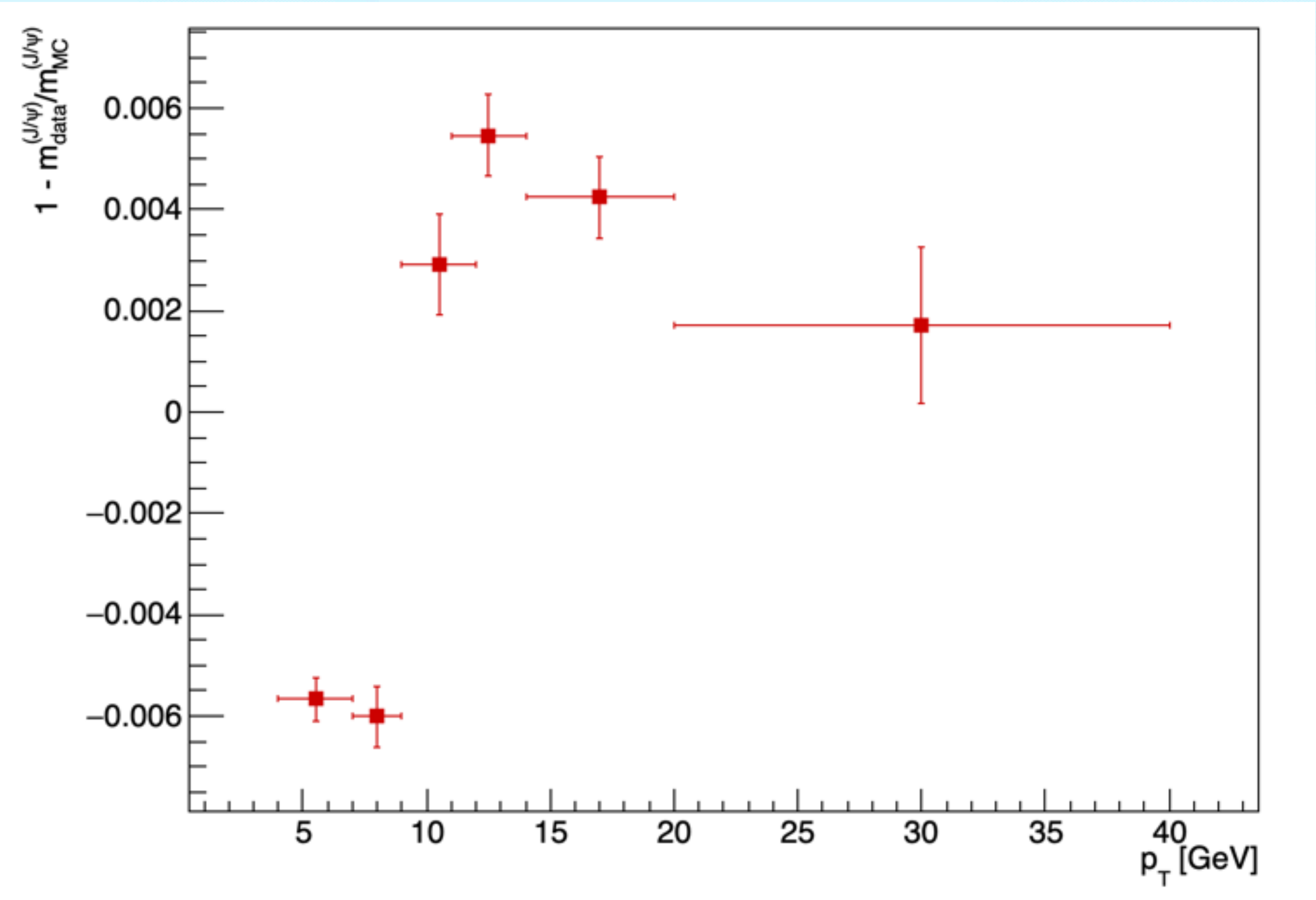


20 GeV $< p_T <$ 40 GeV

Scale and Smearing vs p_T

Scale

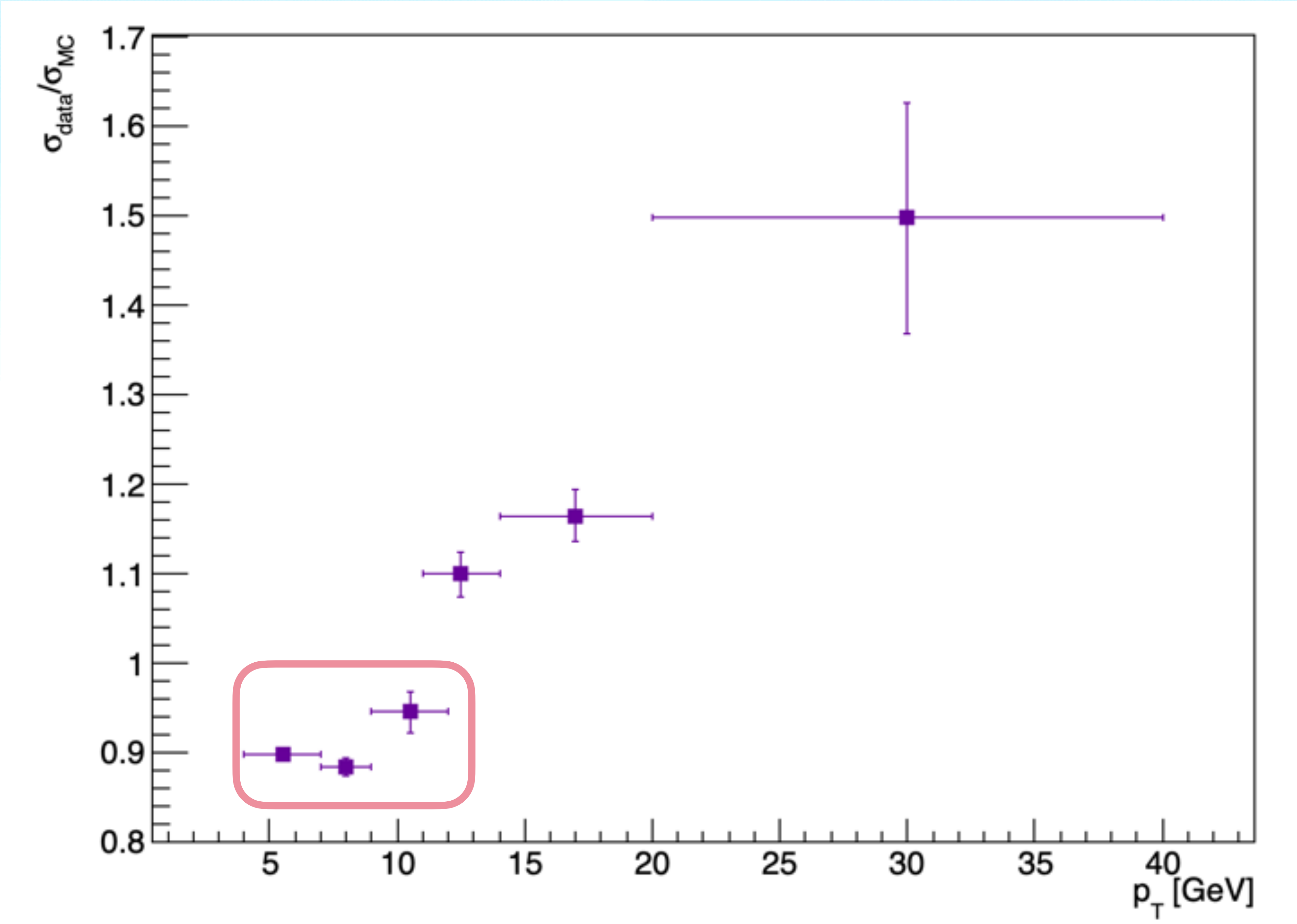
$$1 - \frac{m_{data}}{m_{MC}}$$



Scale is within 1% in all p_T bins

Smearing

$$\frac{\sigma_{data}}{\sigma_{MC}}$$



Apparently $\sigma_{data} < \sigma_{MC}$ at low p_T : this is unusual

→ **Guess**: the regression algorithm, which is not optimized against p_T might be responsible

Scale corrections for single electron

Our goal is to correct single electron energy in data

$$Scale = 1 - \frac{m_{data}^{(J/\psi)}}{m_{MC}^{(J/\psi)}} = 1 - \frac{\sqrt{2E_{data}^{e1}E_{data}^{e2}(1-\cos\theta)}}{\sqrt{2E_{MC}^{e1}E_{MC}^{e2}(1-\cos\theta)}} = 1 - \sqrt{\frac{E_{data}^{e1}E_{data}^{e2}}{E_{MC}^{e1}E_{MC}^{e2}}}$$

→ To obtain ***Scale=0*** need to multiply E_{data} by a factor

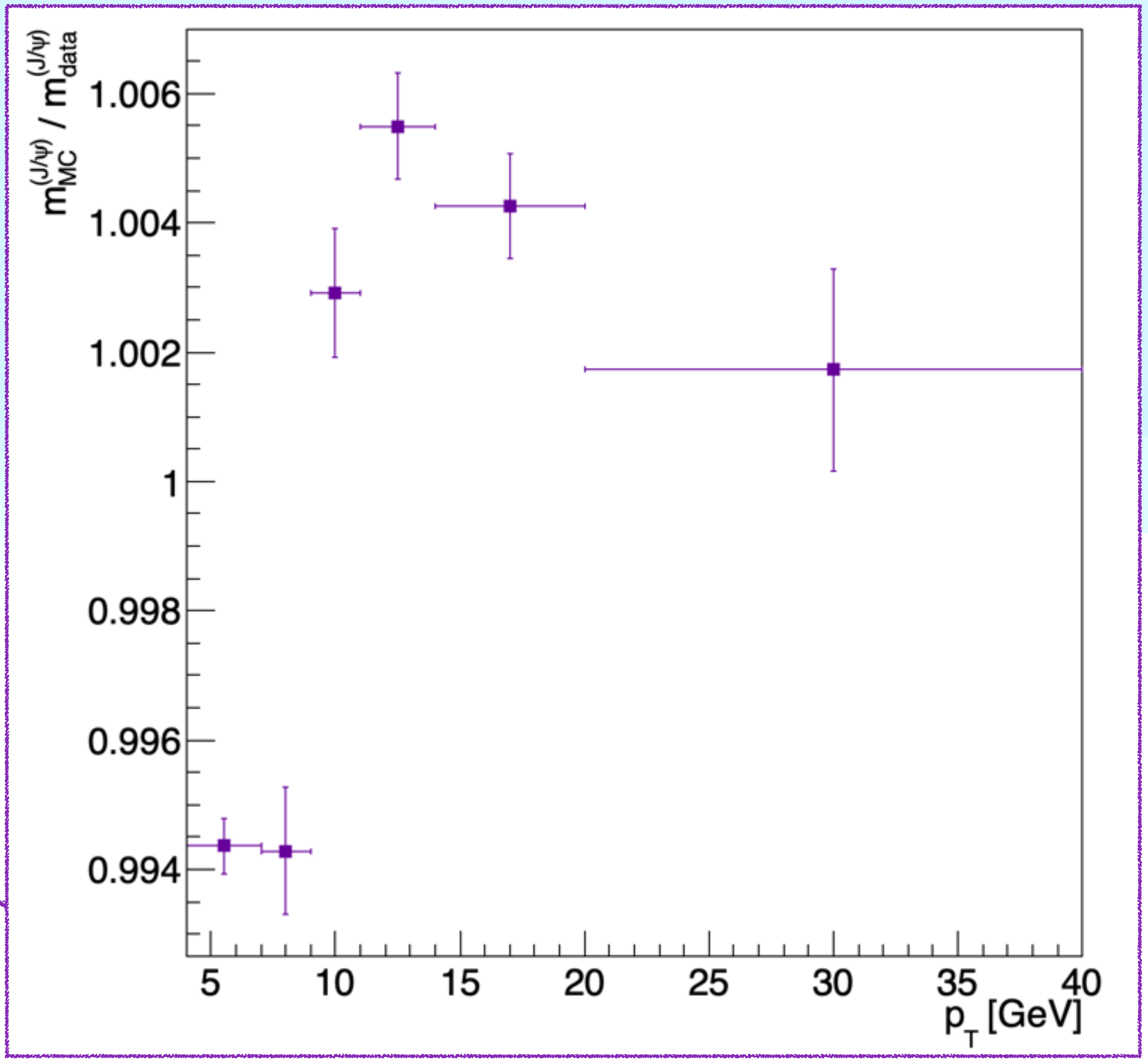
$$\delta = \frac{E_{MC}}{E_{data}}$$

select **diagonal categories** → same correction for both electrons

$$1 - \frac{m_{data}^{(J/\psi)}}{m_{MC}^{(J/\psi)}} = 1 - \sqrt{(\delta^{p_T bin})^{-2}\frac{E_{MC}^{e1}E_{MC}^{e2}}{E_{MC}^{e1}E_{MC}^{e2}}}$$

↓

$$\boxed{\delta^{p_T bin} = \frac{m_{MC}^{(J/\psi)}}{m_{data}^{(J/\psi)}}}$$



Scale correction in bins of the single electron p_T

OBS: from Z peak corrections in the range 1.003 - 1.007 are found for 20 GeV < p_T < 40 GeV

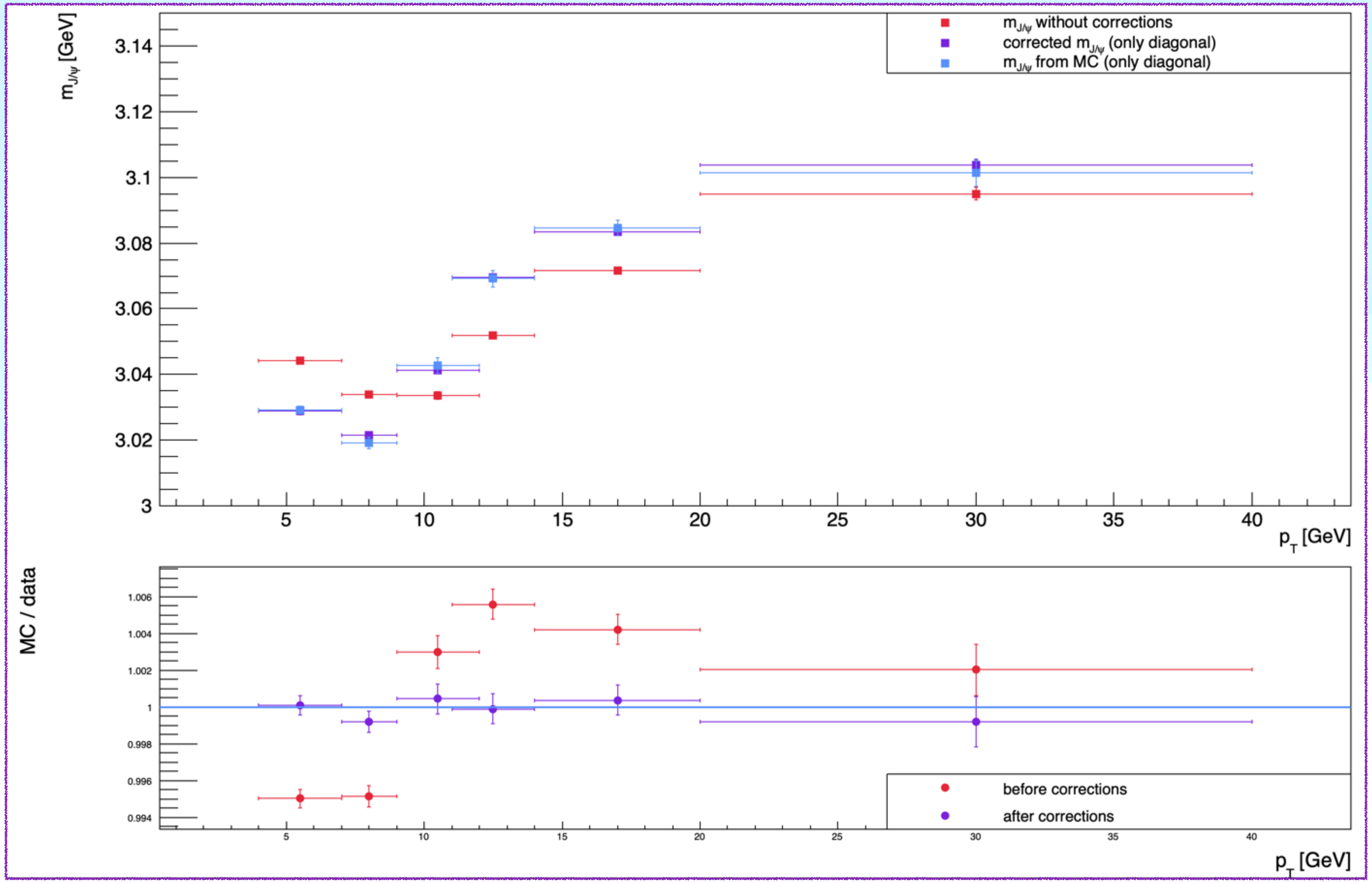
https://twiki.cern.ch/twiki/bin/view/CMS/EgammSFandSSRun3#Scale_And_Smearings_Correctionli

Corrections Validation - diagonal categories

- Single electron corrections are applied to energies in data

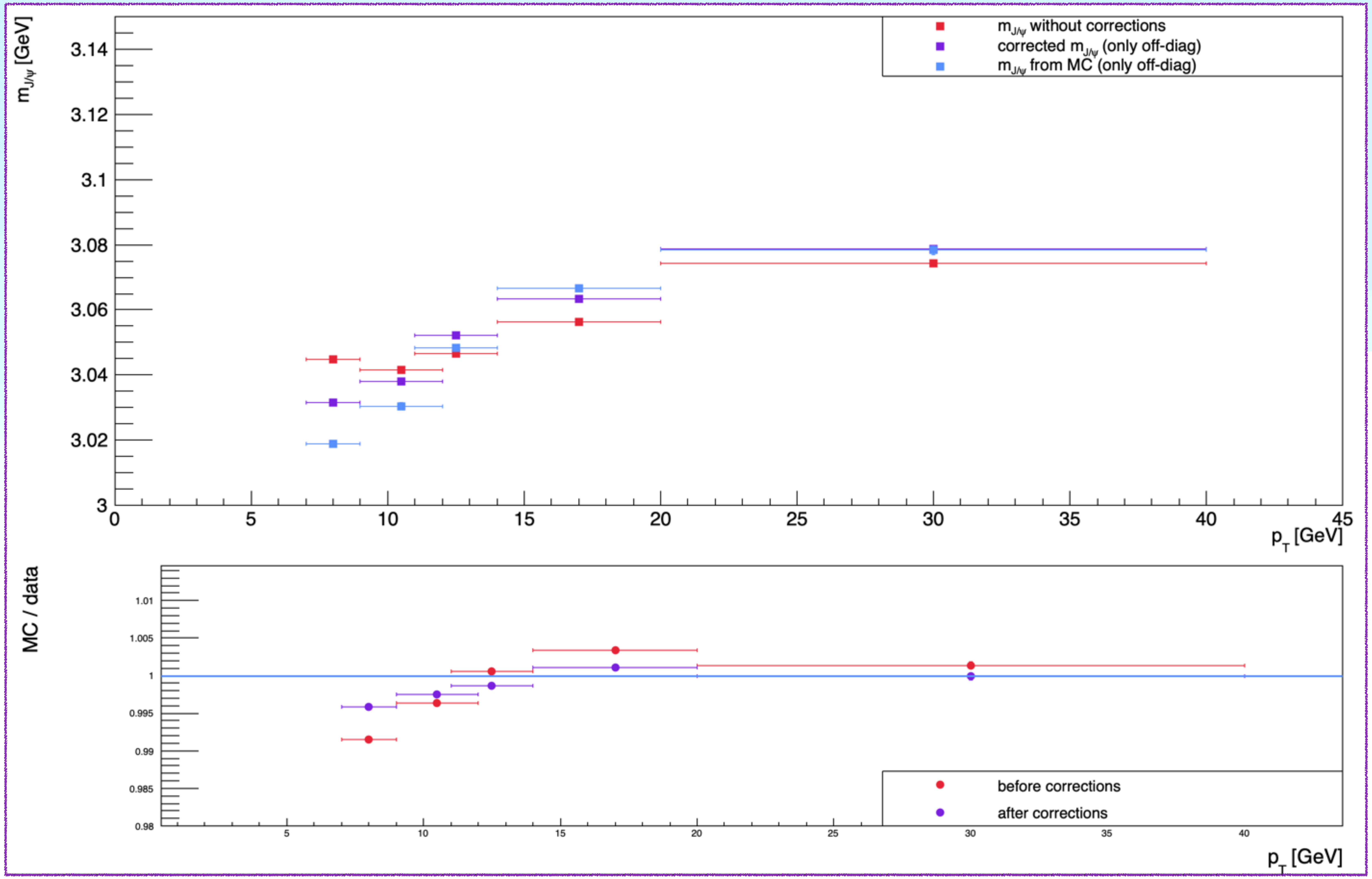
- Invariant mass from corrected electron energies fitted again in p_T bins

- Results show agreement within 1 σ with MC for all diagonal categories



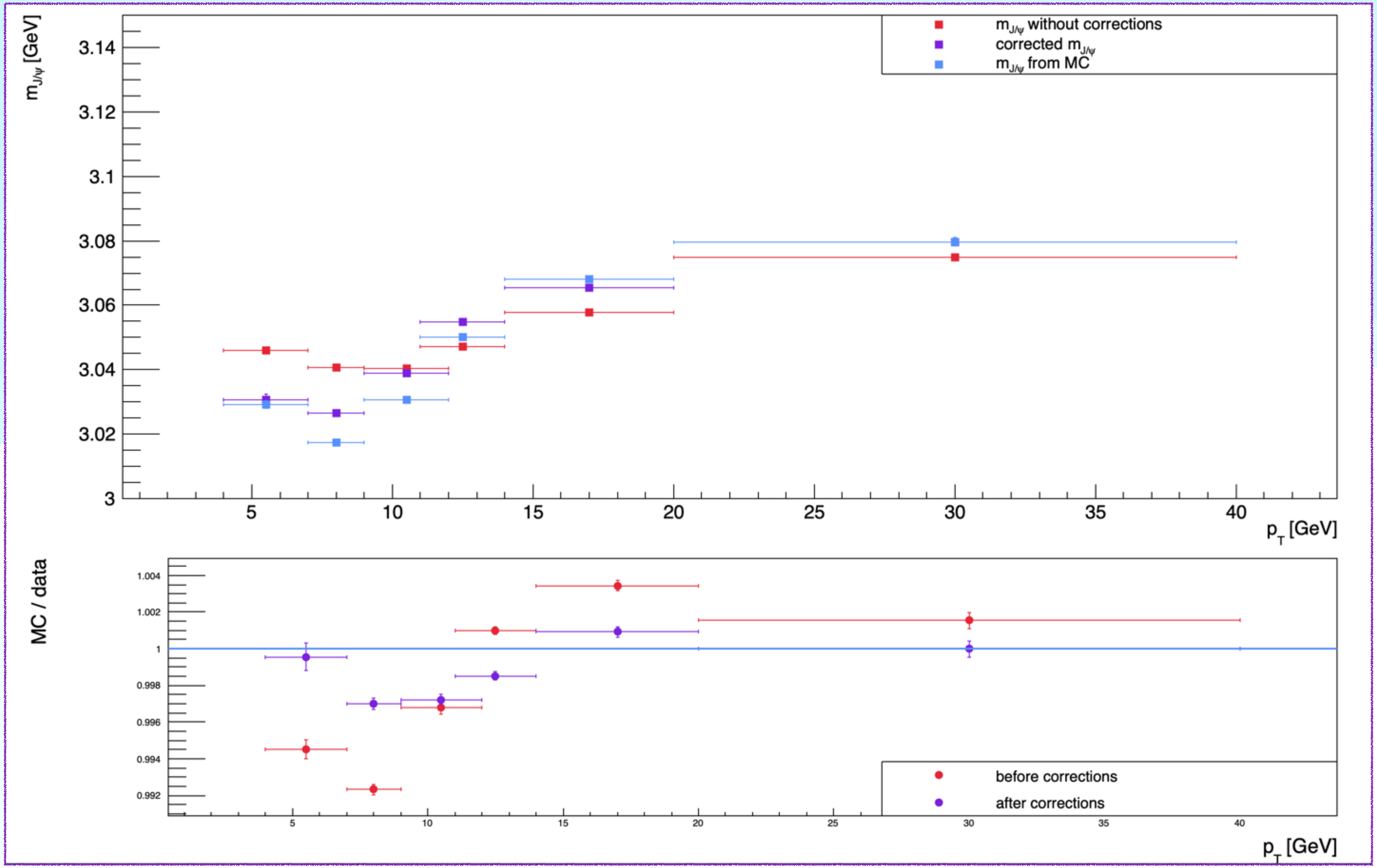
Corrections Validation - off diagonal categories

- Worse agreement, but still some improvement in most of the bins
- No off-diagonal data in the 1st bin, as, by construction the 1st electron has always higher p_T than the 2nd electron



Overall performance of scale corrections

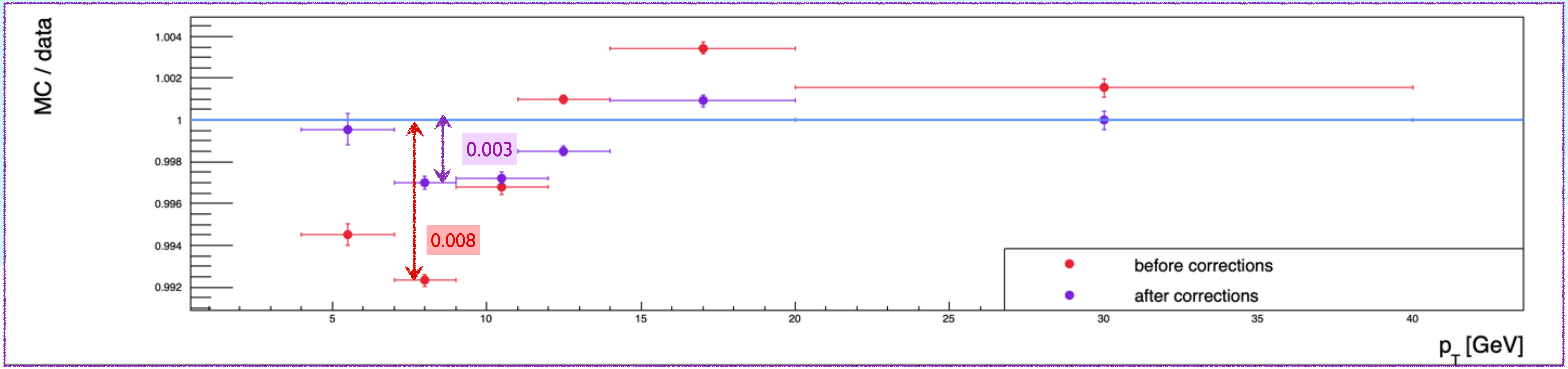
Overall improvement in the agreement between data and MC



Maximum spread before corrections

Maximum spread after corrections

Overall performance of scale corrections



Maximum distance from 1 in the MC/data ratio decreased by 63% after corrections

Conclusions…

- Scale and smearing dependence on p_T checked for $p_T > 4$ GeV
- Pointed possible issue in the regression at low p_T
- Scale corrections for single electron energy **extracted** from J/ψ for the first time
- Corrections have been **validated**, with **overall improvement** (63%) in data-MC agreement
- Same workflow can be applied to extract scale & smearing for photons

…and next steps

- Study of systematic errors
- Time-dependent corrections binned on Run number
- Smearing corrections
- Study of corrections impact on full reconstructed electron

References

[1] Shervin Nourbakhsh. "Studio degli eventi J/ in due elettroni con i primi dati di CMS". MA thesis. Università La Sapienza, 2009/2010.

[2] Albert M Sirunyan et al. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: JINST 16.05(2021), P05014. doi: 10.1088/1748- 0221/16/05/P05014. arXiv: 2012.06888 [hep-ex].

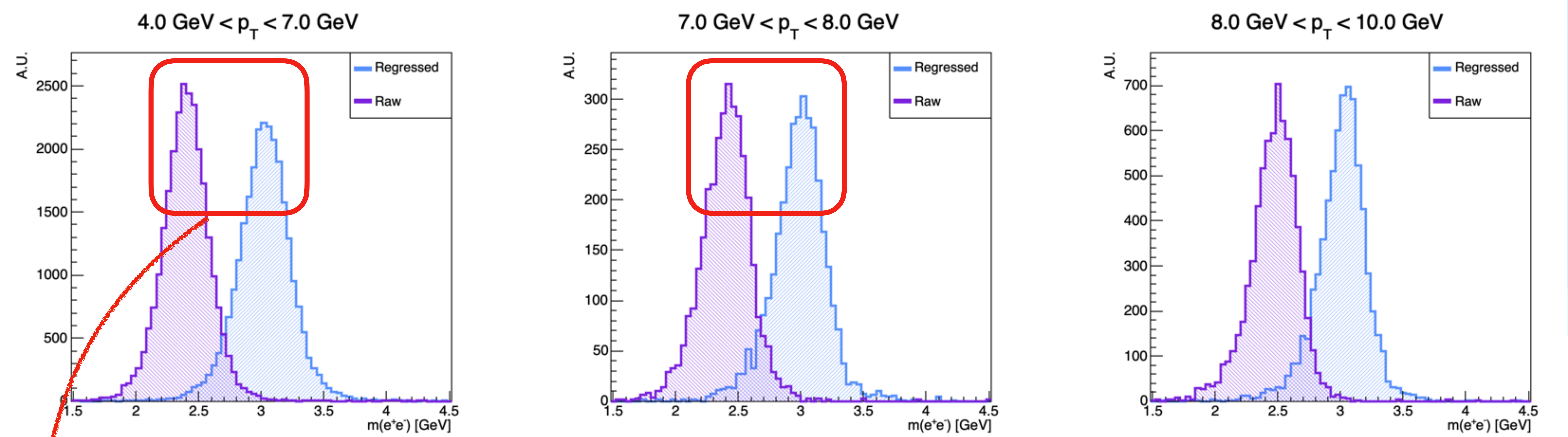
[3] Raffaella Tramontano. "Searches for Heavy Neutral Leptons in B Meson Decays with the CMS Experiment". PhD thesis. Scuola di Dottorato in Scienze Astronomiche, Chimiche, Fisiche e Matematiche 'Vito Volterra', 2023.

[4] Ntuplizer: https://github.com/Raffaella07/ScaleAndSmearingTools/tree/jpsi

[5] Scale from Z: https://twiki.cern.ch/twiki/bin/view/CMS/EgammSFandSSRun3#Scale_And_Smearings_Correctionli

Backup…

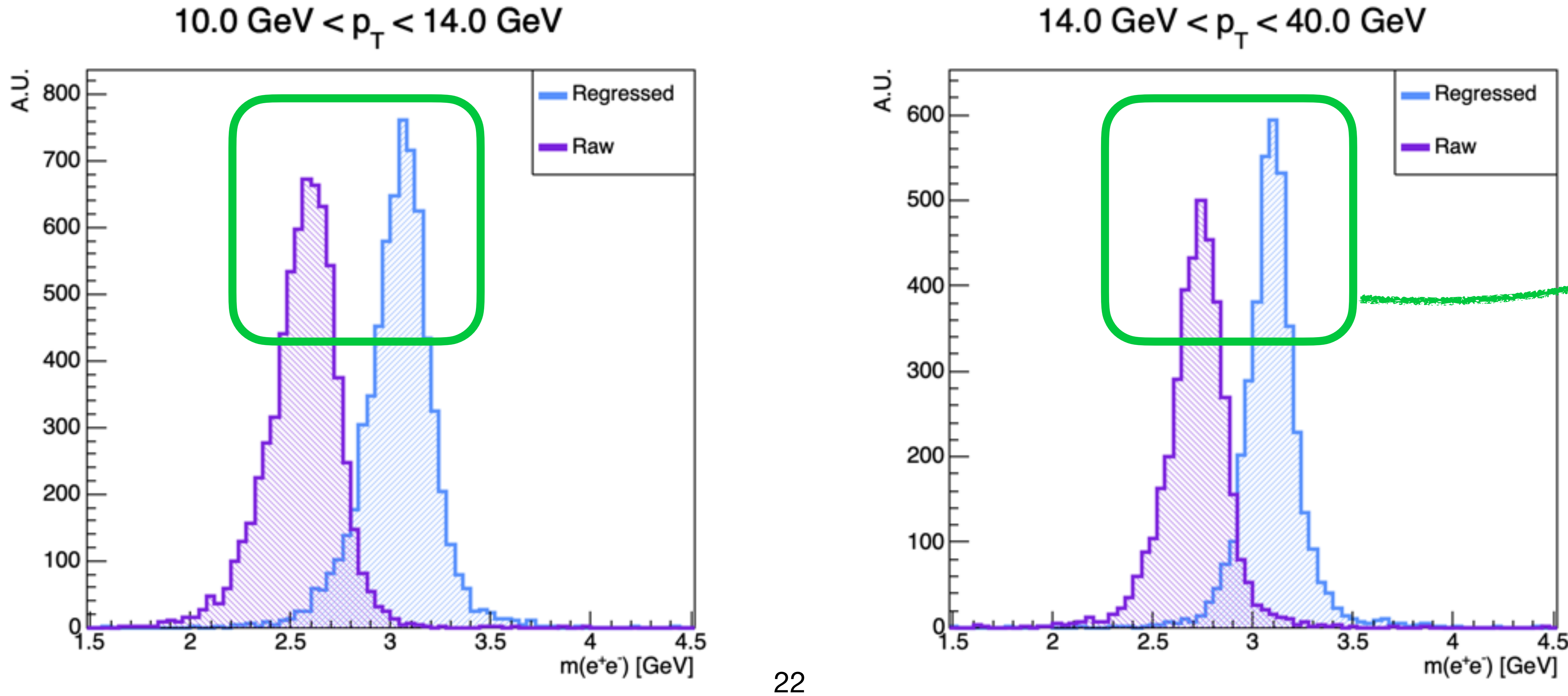
Regression behavior at different p_T - MC



Same number of events!

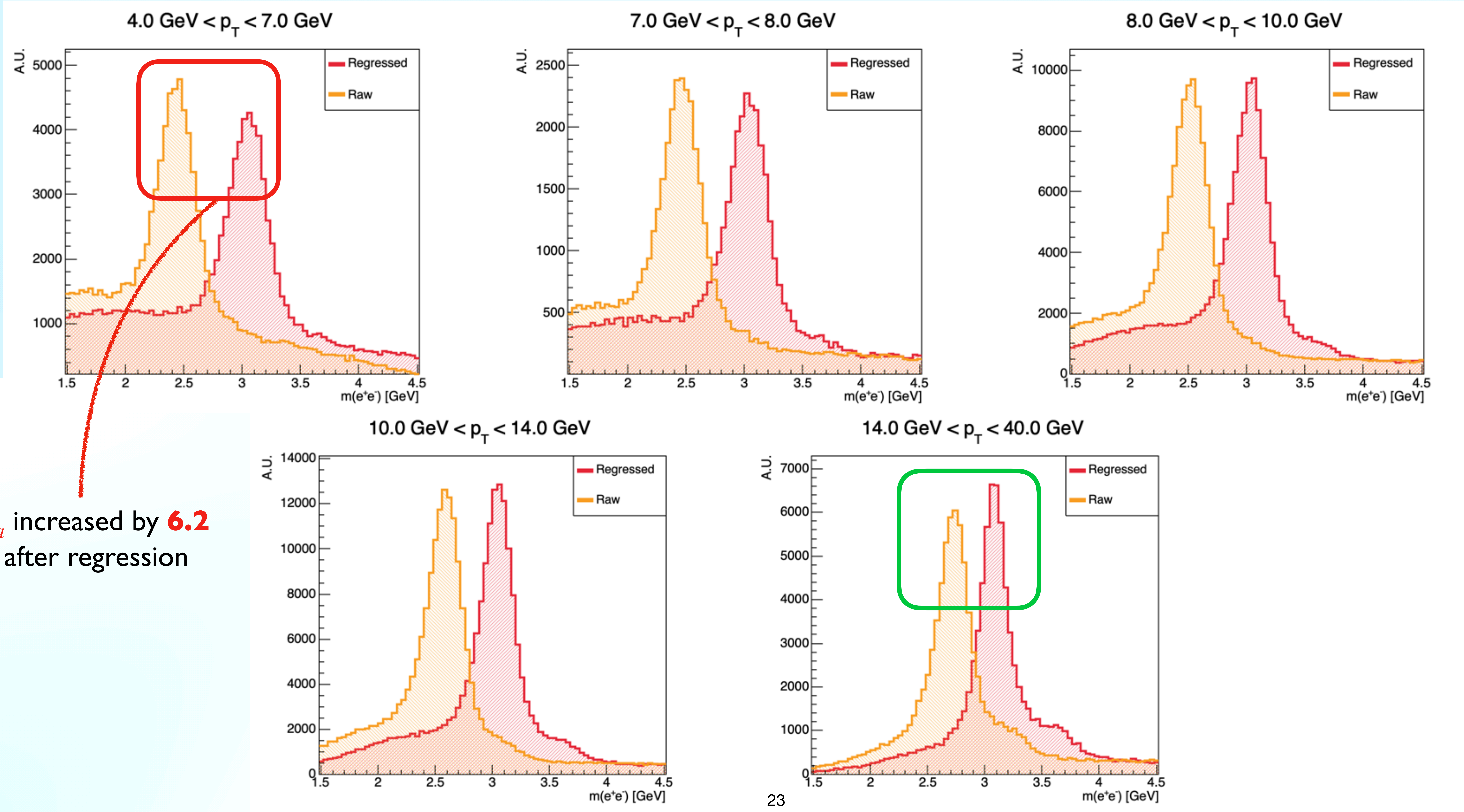
Wider distribution after regression algorithm:

σ_{MC} increased by **9.2 %** after regression



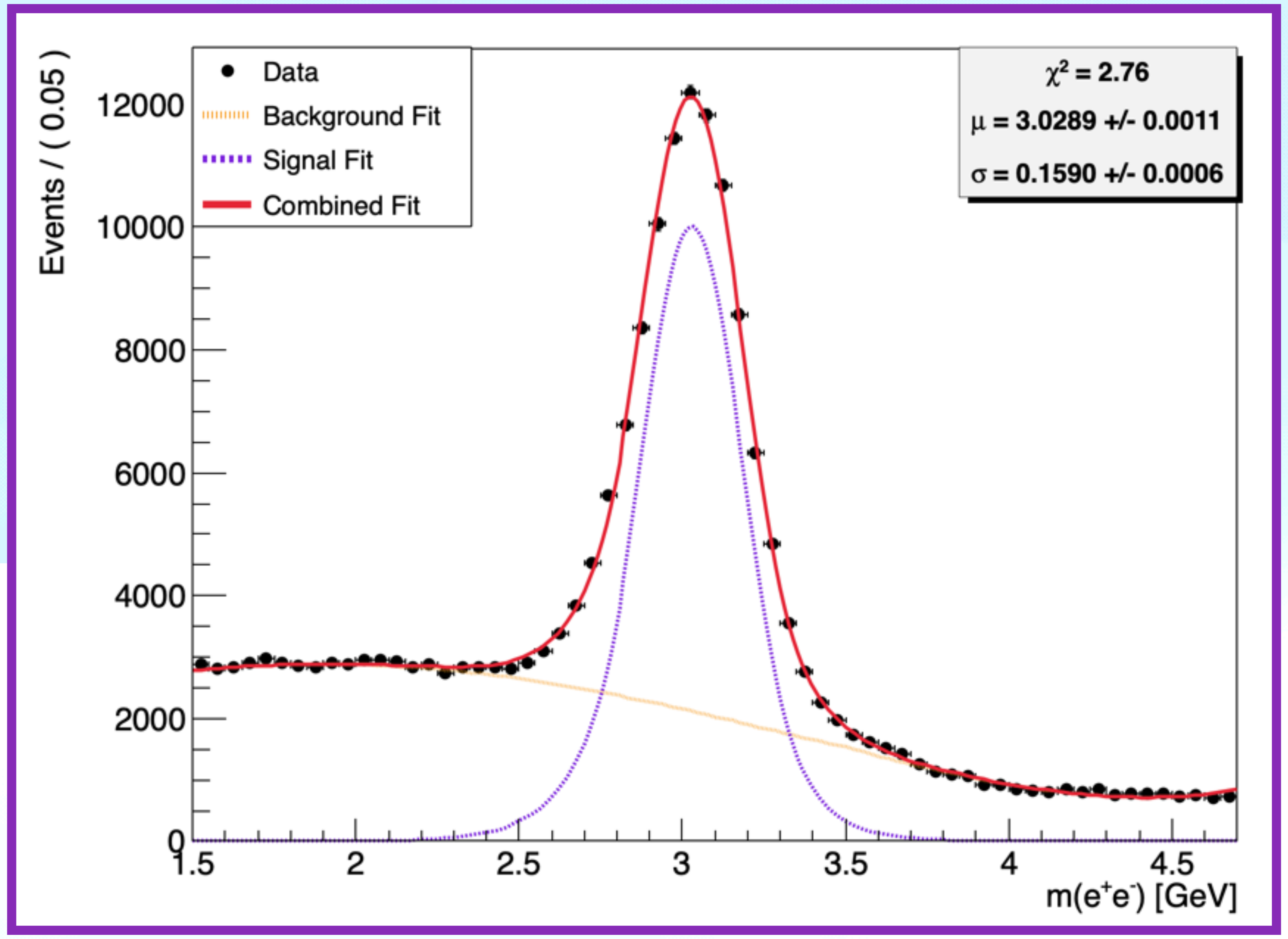
Sharper distribution after regression algorithm

Regression behavior at different p_T - Data

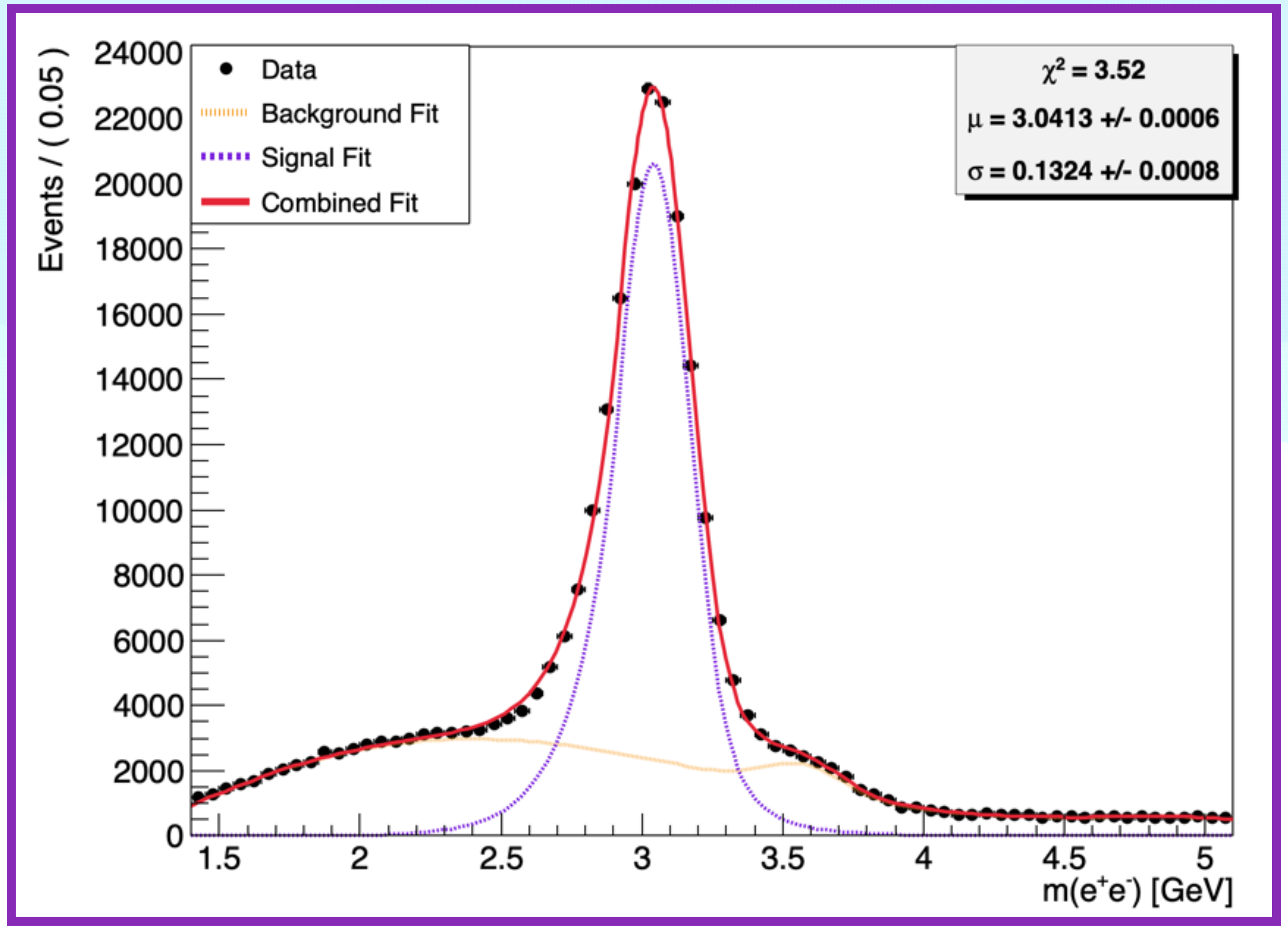


σ_{data} increased by **6.2 %** after regression

Fits for validation: diagonal

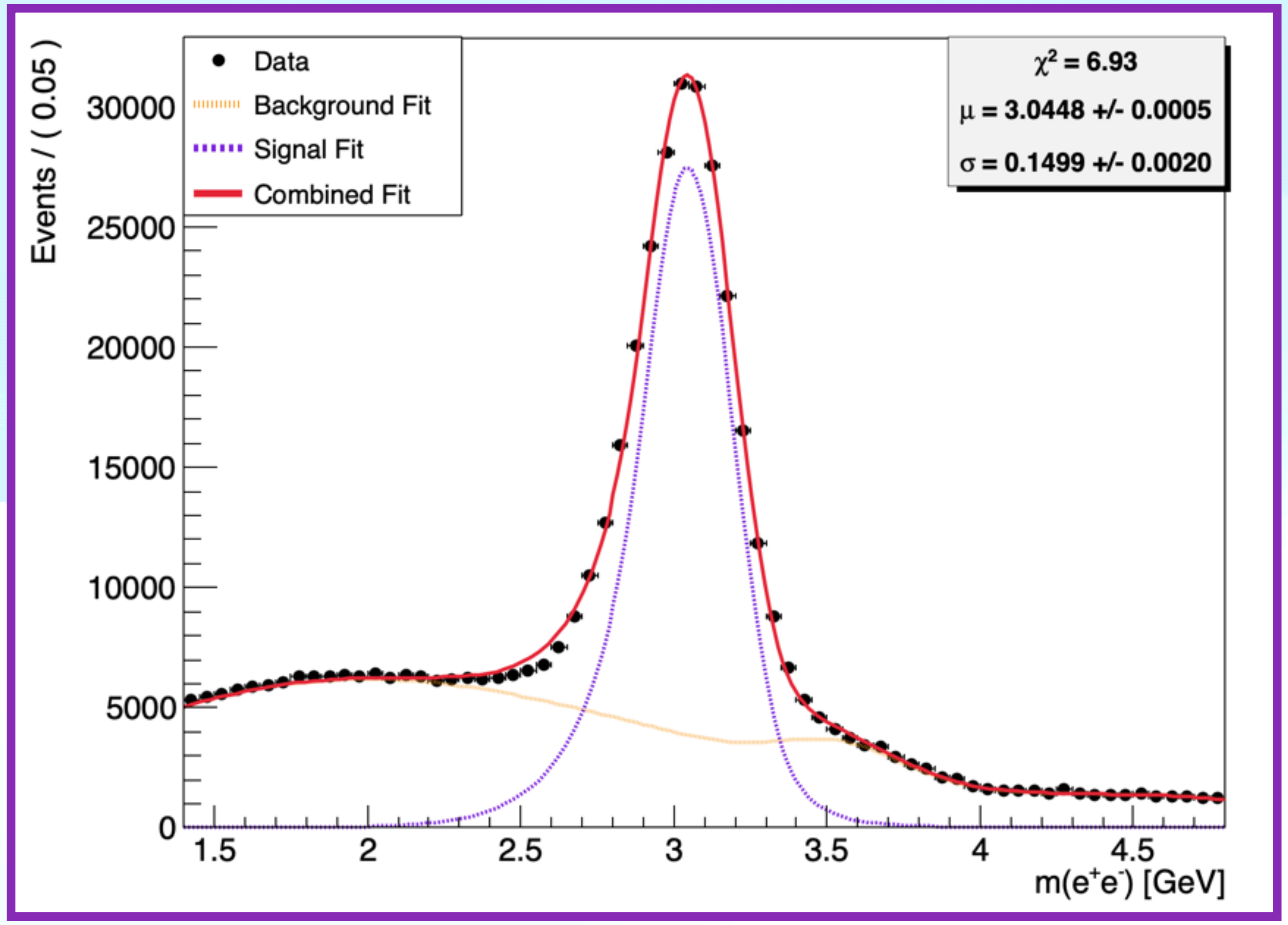


4 GeV < p_T < 7 GeV

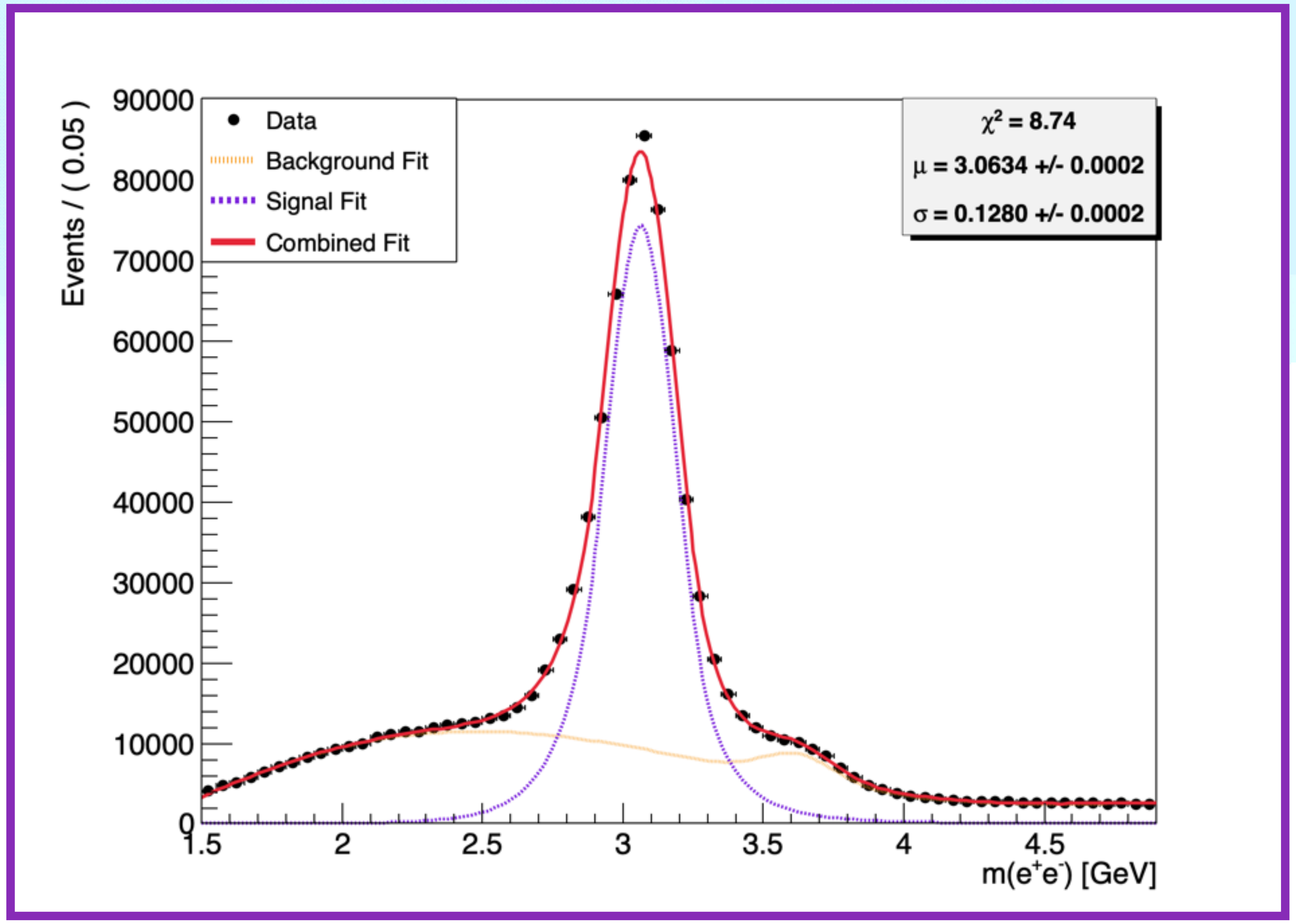


9 GeV < p_T < 11 GeV

Fits for validation: off-diagonal

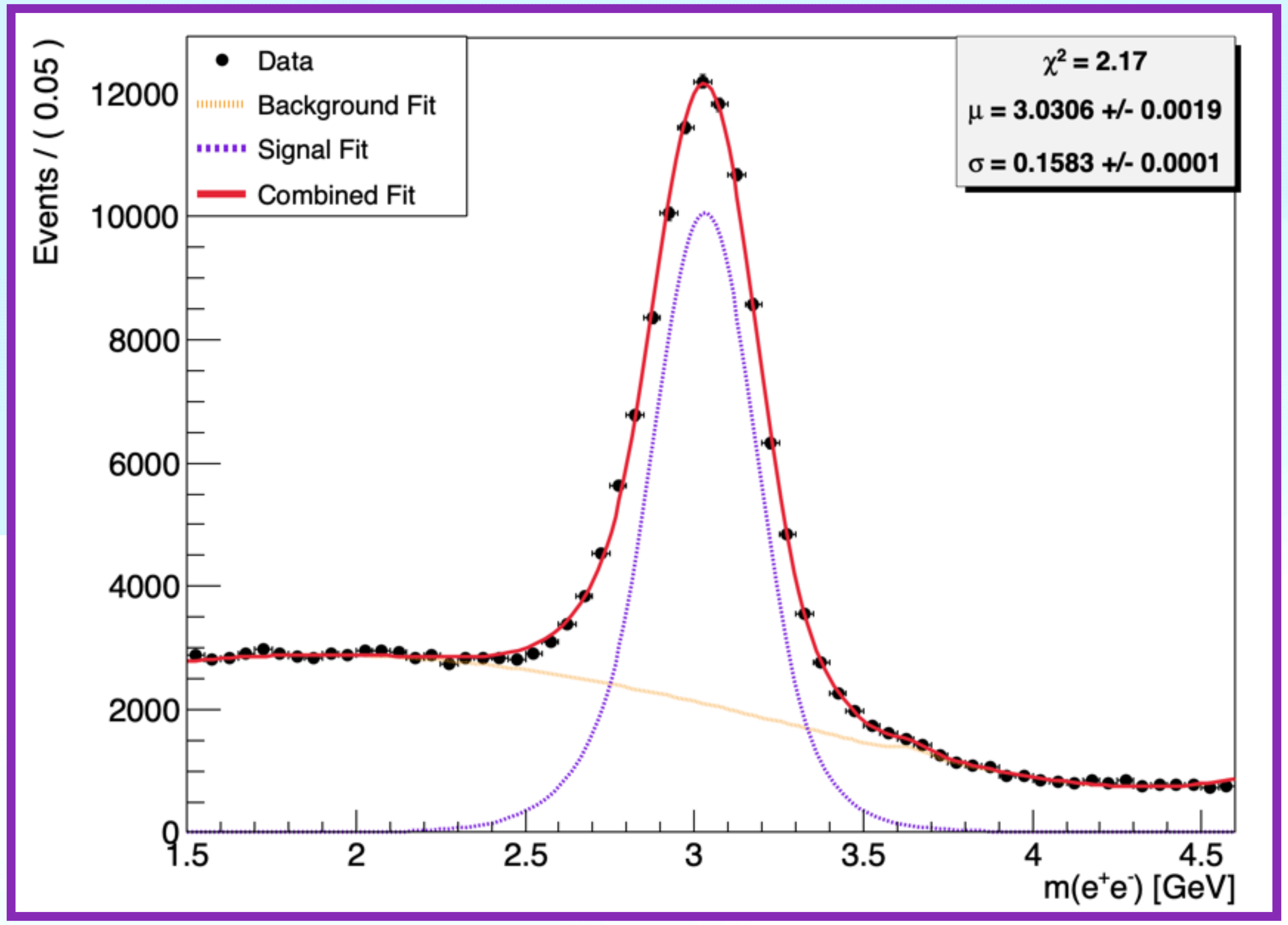


7 GeV < p_T < 9 GeV

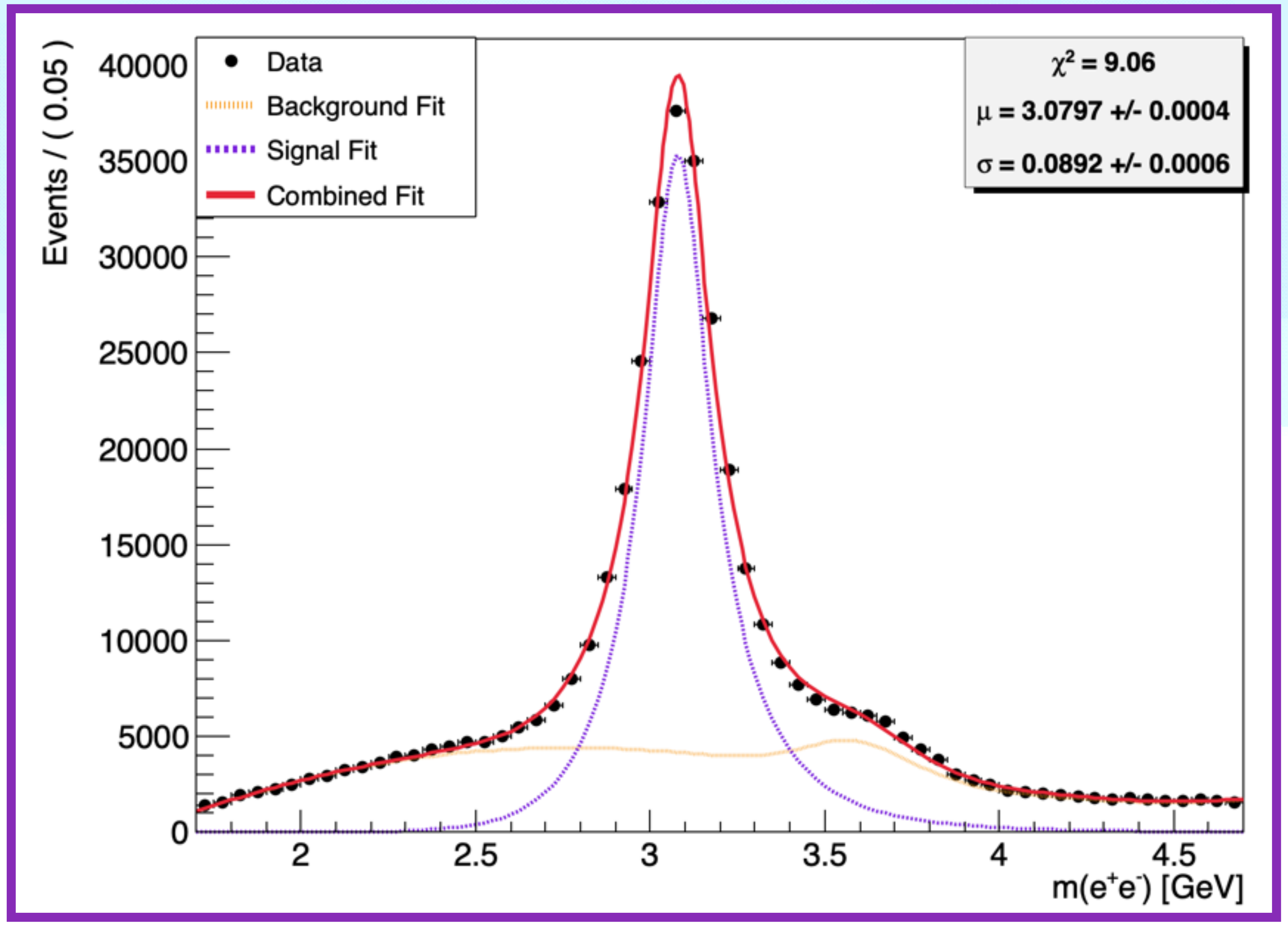


14 GeV < p_T < 20 GeV

Fits for validation: all categories



4 GeV < p_T < 7 GeV



20 GeV < p_T < 40 GeV

More on: https://eldesant.web.cern.ch/Fit_BinOnPt/

Scale consistency between J/ψ and $\psi(2s)$

	$m_{J/\psi}\ [GeV]$	$m_{\psi(2s)}\ [GeV]$	$\Delta m\ [GeV]$
PDG	3.0969	3.6861	0.5892
Bin 5	3.0717 +/- 0.0007	3.627 +/- 0.005	0.556 +/- 0.006
Bin 6	3.095 +/- 0.002	3.61 +/- 0.01	0.51 +/- 0.01