K-Factor Updates

Mohit Srivastav

LHCHWG Meeting

February 20, 2026

ggH k-factor

- Gluon-fusion k-factor is done currently using HNNLO
  - LO → NLO
  - LO → NNLO
  - NLO → NNLO comes for "free" from this since we calculate NLO/NNLO anyways
- K-factor is applied to background as well, using theory assumptions that the bkg, signal, and interference k-factors differ by < 10%
- Need to make a slew of changes to HNNLO code to make it work!
  - Change string length maximum for a bunch of variables internally to allow for NNPDF names as well as designators for systematic variations
  - Edit Higgs width for all resonant points to be 4.14 MeV in br.sm2 file
- Nominal values are calculated, systematic variations require a **lot** of resources.

PDF Set

- The old PDF set used was something different than what we use today for the NLO Powheg samples
- We also have decided to use a slightly different PDF set at NNLO that is hessian
- As such, we revise the Run-2 kfactor as well as the Run-3 value

- We use NNPDF30_lo_as_0130 for LO (same)

- NNPDF31_NNLO_hessian_pdfas for NLO samples

- NNPDF31_NNLO_as_0118_mc_hessian_pdfas for final NNLO calculation

Mass Points

- Mass points are sampled individually
- Points below 100 GeV in HNNLO with very unphysical cross-sections at higher orders
  - i.e., 70 GeV
    - LO → 3.3973338433323967E-003
    - NLO →8.2368421491823014
    - NNLO →1.6262855329656050
- Go from 100 GeV to 6000 GeV in spaced points

```
hmasses = (
    # be very precise in the H125 on-shell region
    list(np.arange(100, 140, 2, dtype=int)) +
    # between on-shell and off-shell don't sample too much
    list(np.arange(140, 180, 10, dtype=int)) +
    # starting at 2mZ start sampling a lot relative to the size
    list(np.arange(180, 1000, 5, dtype=int)) +
    # above a TeV things start to break down anyways
    list(np.arange(1000, 6000, 50, dtype=int))
)
```

Figure 6: The hmasses as defined in Python code with GeV sampling. There are 228 of them.

Systematic Variations

- 9-point variation for the QCD-scales jointly
  - New guidance on k-factors likes an enveloped profile instead of separate uncertainties
- 100 PDF uncertainties
- 2 α_S variations
- **~50k jobs required per run**

$$N_{\text{jobs}} = N_{\text{hmasses}} + 2 \times N_{\text{hmasses}} \times \left(N_{(\mu_F,\mu_R)} + N_{\text{PDF}} + N_{\alpha_S}\right)$$
$$= 228 + 2 \times 228 \times (9 + 100 + 2)$$
$$= 50{,}844$$

- μ_R, μ_F Nominally set to $\frac{m_{4\ell}}{2}$, this number is varied to $\frac{m_{4\ell}}{4}$ and $m_{4\ell}$. The 2 values are changed side-by-side to profile the whole space of Eq. 3.

, where $\{\frac{1}{2}, \frac{1}{2}\}$ denotes the nominal case, in order to derive an uncertainty using Eq. 4, where the maximum and the minimum cross-sectional variations are chosen to define the uncertainty band around the nominal point.

$$(\mu_R, \mu_F) = \left\{ \left\{\frac{1}{4}, \frac{1}{4}\right\}, \left\{\frac{1}{4}, \frac{1}{2}\right\}, \left\{\frac{1}{4}, 1\right\}, \left\{\frac{1}{2}, \frac{1}{4}\right\}, \left\{\frac{1}{2}, \frac{1}{2}\right\}, \left\{\frac{1}{2}, 1\right\}, \left\{1, \frac{1}{4}\right\}, \left\{1, \frac{1}{2}\right\}, \{1, 1\}, \right\} \quad (3)$$

$$(\delta\sigma)_{\text{QCD-scale}} = \frac{\sigma_{\max}(\mu_R, \mu_F) - \sigma_{\min}(\mu_R, \mu_F)}{2} \quad (4)$$

- PDF set Nominally PDF set 0 is used. However, in the hessian PDF sets chosen in Table. 2, there are 100 extra sets to choose from. The systematic uncertainty is then defined by Eq. 5, where $\sigma^{(i)}$ denotes PDF set i, with 0 being the nominal case.

$$(\delta\sigma)_{\text{PDF-set}} = \sqrt{\sum_{i=1}^{100} \left(\sigma^{(i)} - \sigma^{(0)}\right)^2} \quad (5)$$

- α_S value The final 2 PDF sets (101 and 102) for the hessian keep the PDF the same but change α_S from its nominal value of 0.118. There is a variation up (0.116) and a variation down (0.120). These variations are simply taken at face value, with the systematic variations up and down being the calculations from HNNLO.

Smoothing done via Gaussian Regression

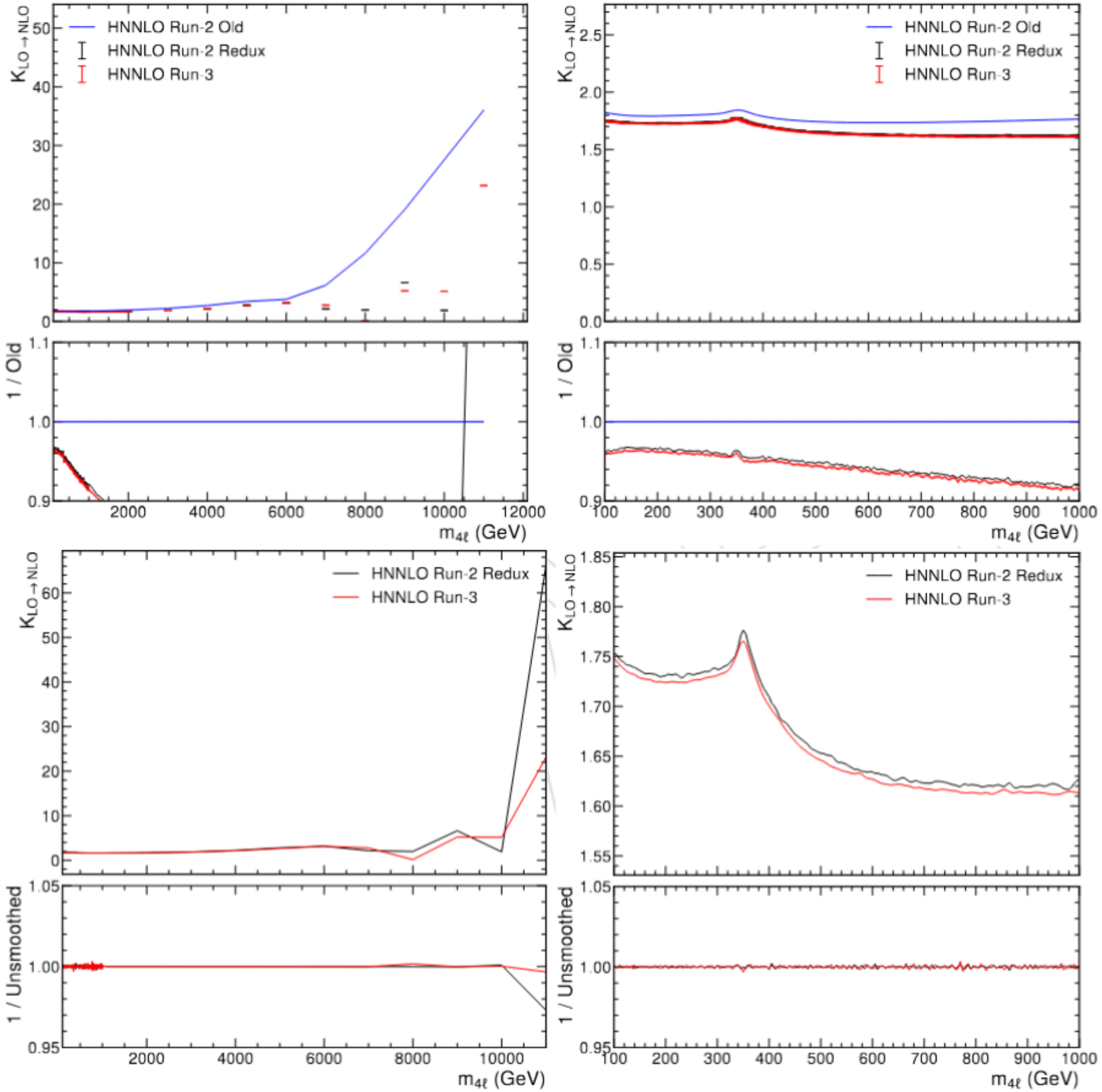


Figure 7: The gluon-fusion NLO k-factor as a function of $m_{4\ell}$, for the whole range (left), and zoomed in to the range [100, 1000] GeV (right), for the unsmoothed case comparing to the old k-factor (top) and the smoothed case comparing to the unsmoothed case (bottom).

Smoothing done via Gaussian Regression

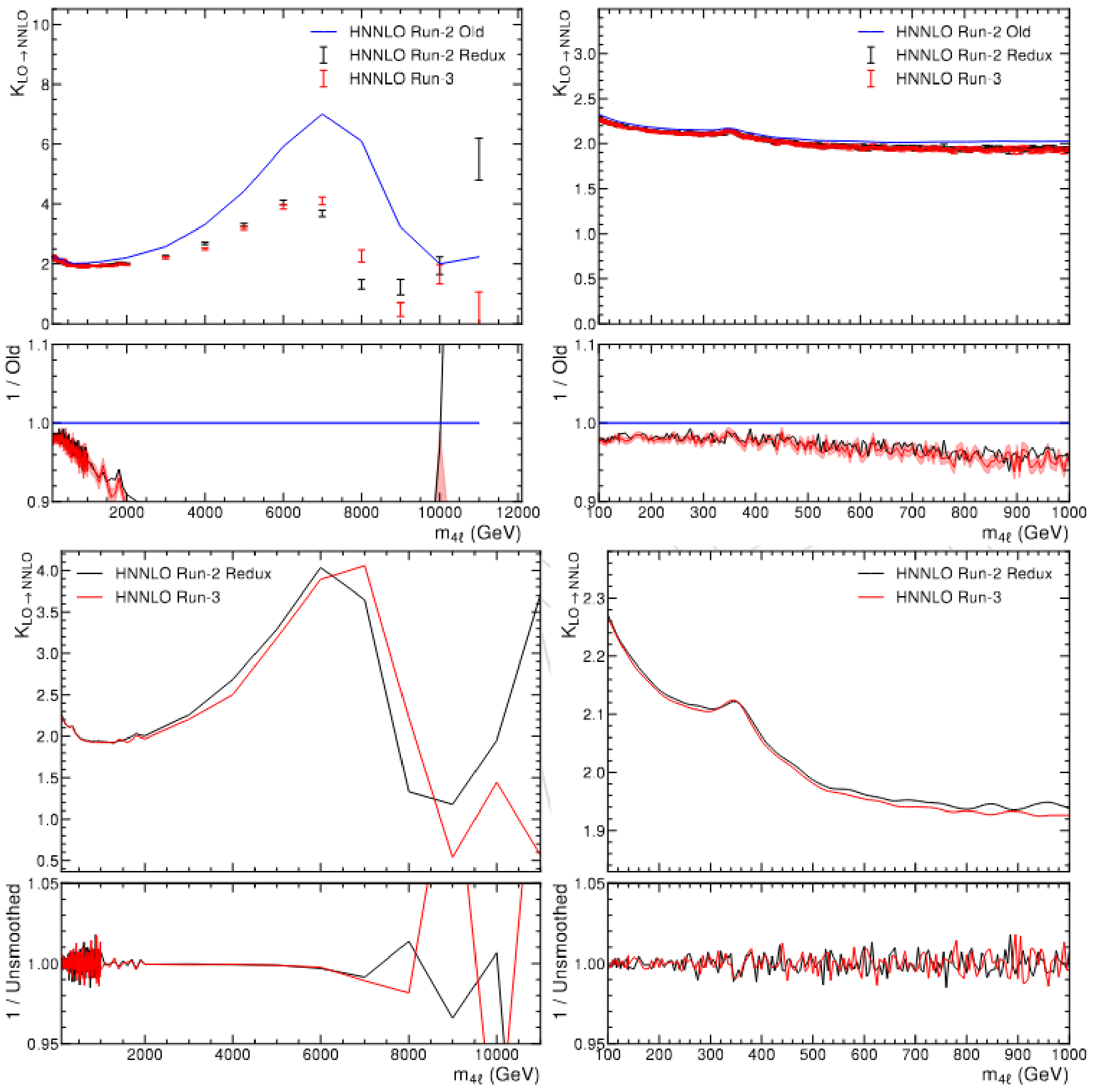


Figure 8: The gluon-fusion NNLO k-factor as a function of $m_{4\ell}$, for the whole range (left), and zoomed in to the range [100, 1000] GeV (right), for the unsmoothed case comparing to the old k-factor (top) and the smoothed case comparing to the unsmoothed case (bottom).

Smoothing done via Gaussian Regression

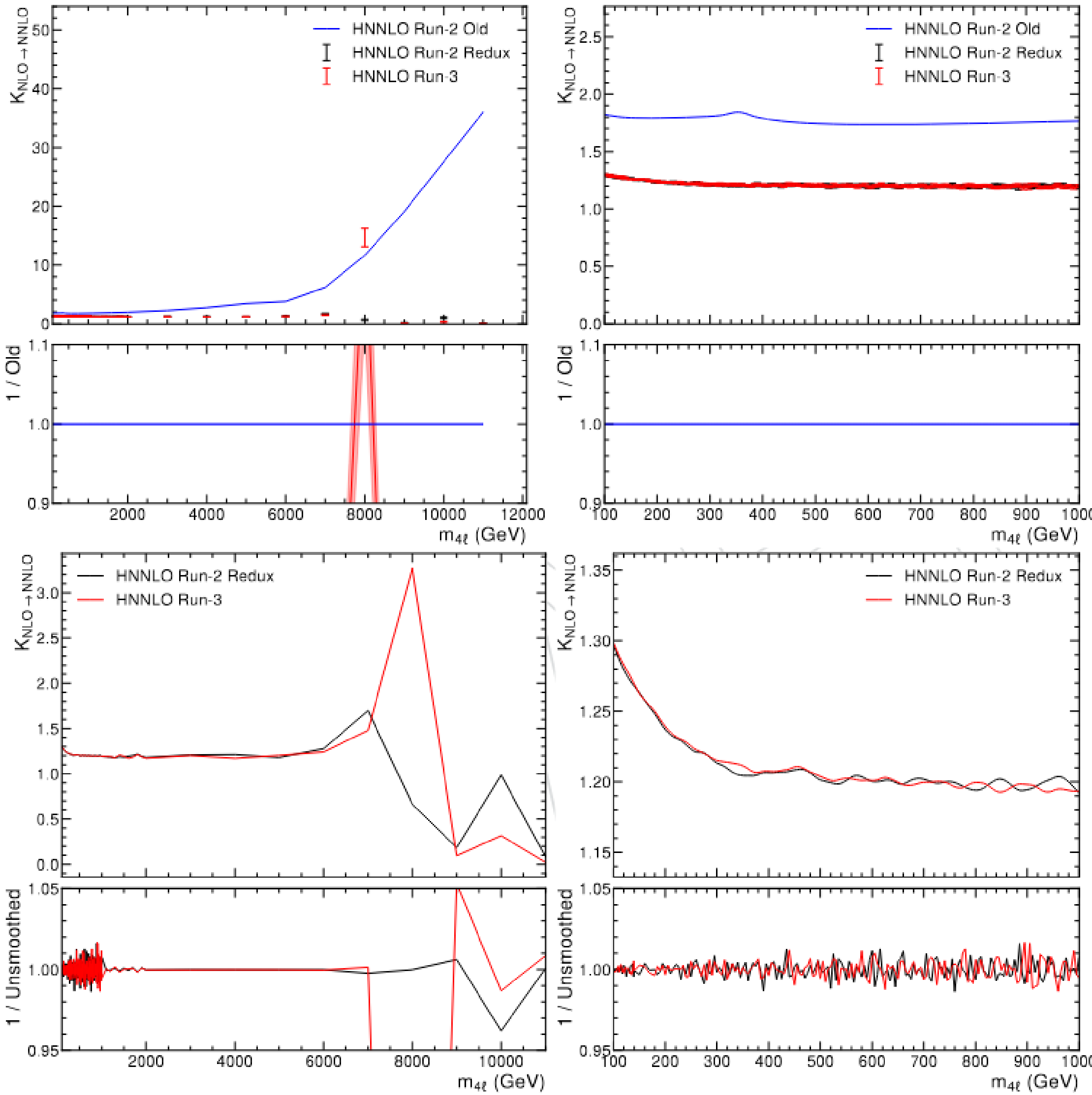


Figure 9: The gluon-fusion NLO → NNLO k-factor as a function of $m_{4\ell}$, for the whole range (left), and zoomed in to the range [100, 1000] GeV (right), for the unsmoothed case comparing to the old k-factor (top) and the smoothed case comparing to the unsmoothed case (bottom).

qqZZ k-factor

- Here we use MATRIX

- Lots of work done here since November to try and understand the EW correction
- Back and forth with MATRIX authors

- Running MATRIX on afs requires some small changes to the code internally that had to be iterated when at higher precision
- This background is a very important quantity used for a slew of analyses, and we need to run them again for Run-3

Setup

The exact values for cuts are shown in Table 1 amongst other quantities, which are also taken from the cuts utilized in analyses shown in Ref. [10]. The photon ΔR becomes relevant once EW effects are included.

variable	cut
muon p_T	≥ 5 GeV
muon $\lvert\eta\rvert$	≤ 2.4
electron p_T	≥ 7 GeV
electron $\lvert\eta\rvert$	≤ 2.4
photon ΔR	≤ 0.5

Table 1: Cuts and other particle quantities for the K-factor setup

PDF: NNPDF31_nnlo_as_0118_mc_hessian_pdfas

$$\mu_{F,R} = 1$$

Dealing with EW Photon Emission

- **This was by far our hardest problem to deal with**

- MATRIX emits an extra photon as either FSR or ISR, but does not distinguish the 2
- Photons are merged back into the leptons around a ΔR cone, but are otherwise lost to the user

- These FSR/ISR effects warp the distribution!

Dealing with EW Photon Emission

- Large deviations seen on the left side of the Z-peak and $2m_Z$ threshold due to EW photon emission.

- Peak gets shifted to the right due to missing energy from unmerged FSR for the EW correction, and things look weird!

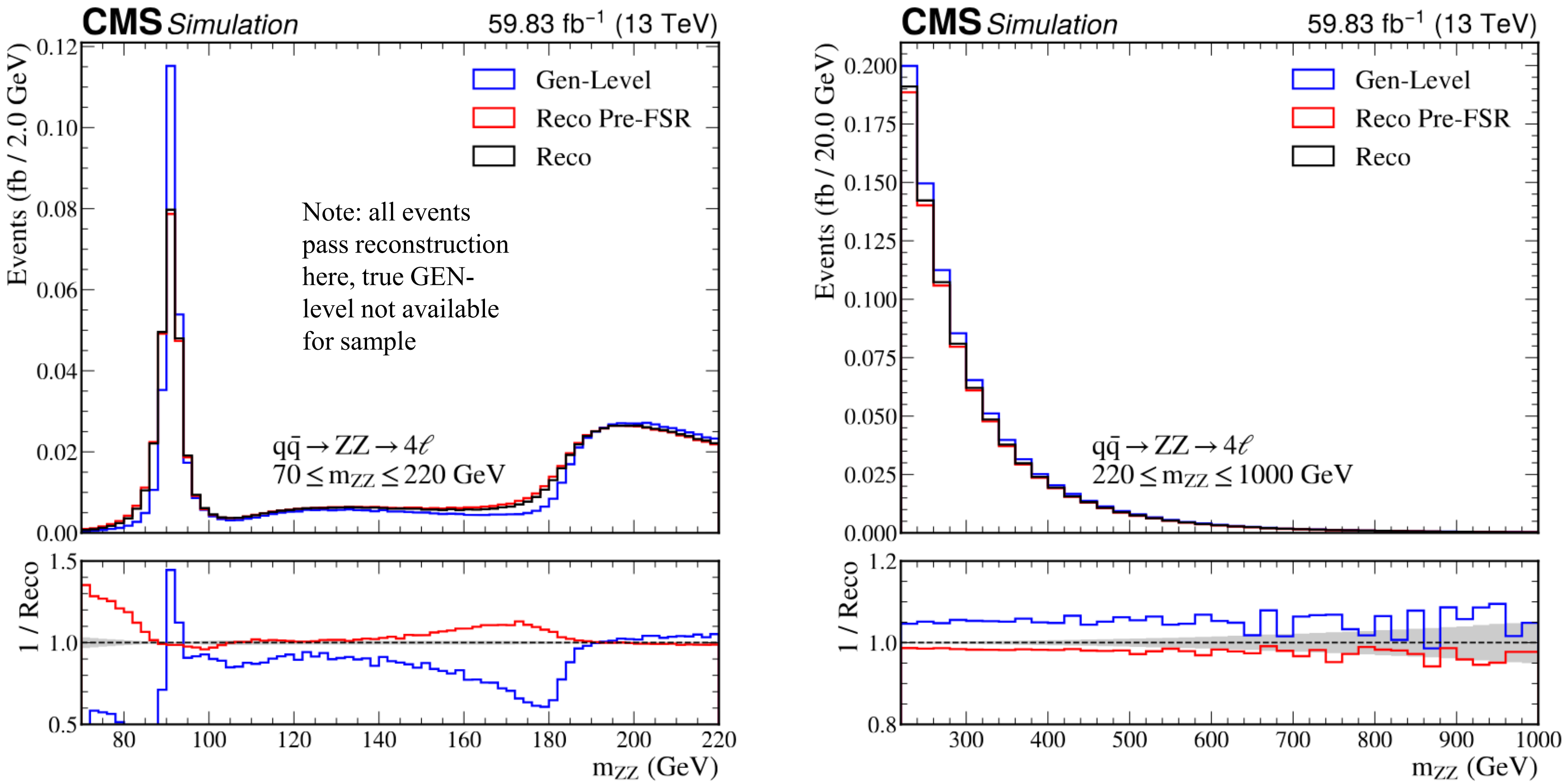
- We decided to check what we do in experiment…

Dealing with EW Photon Emission

- You can see similar behavior in reconstructed Monte-carlo for qq4l!

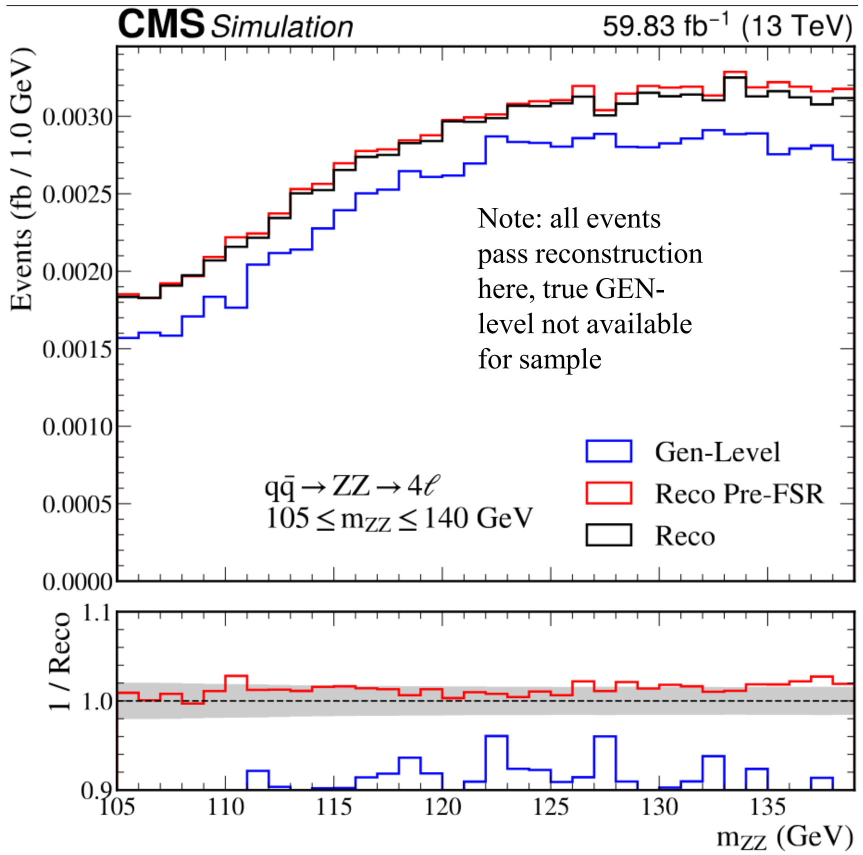
- Nominally flat in off-shell region, but makes a large difference at the Z and $2m_Z$ thresholds

- MATRIX not *exactly* the same, since GEN-level distribution looks different



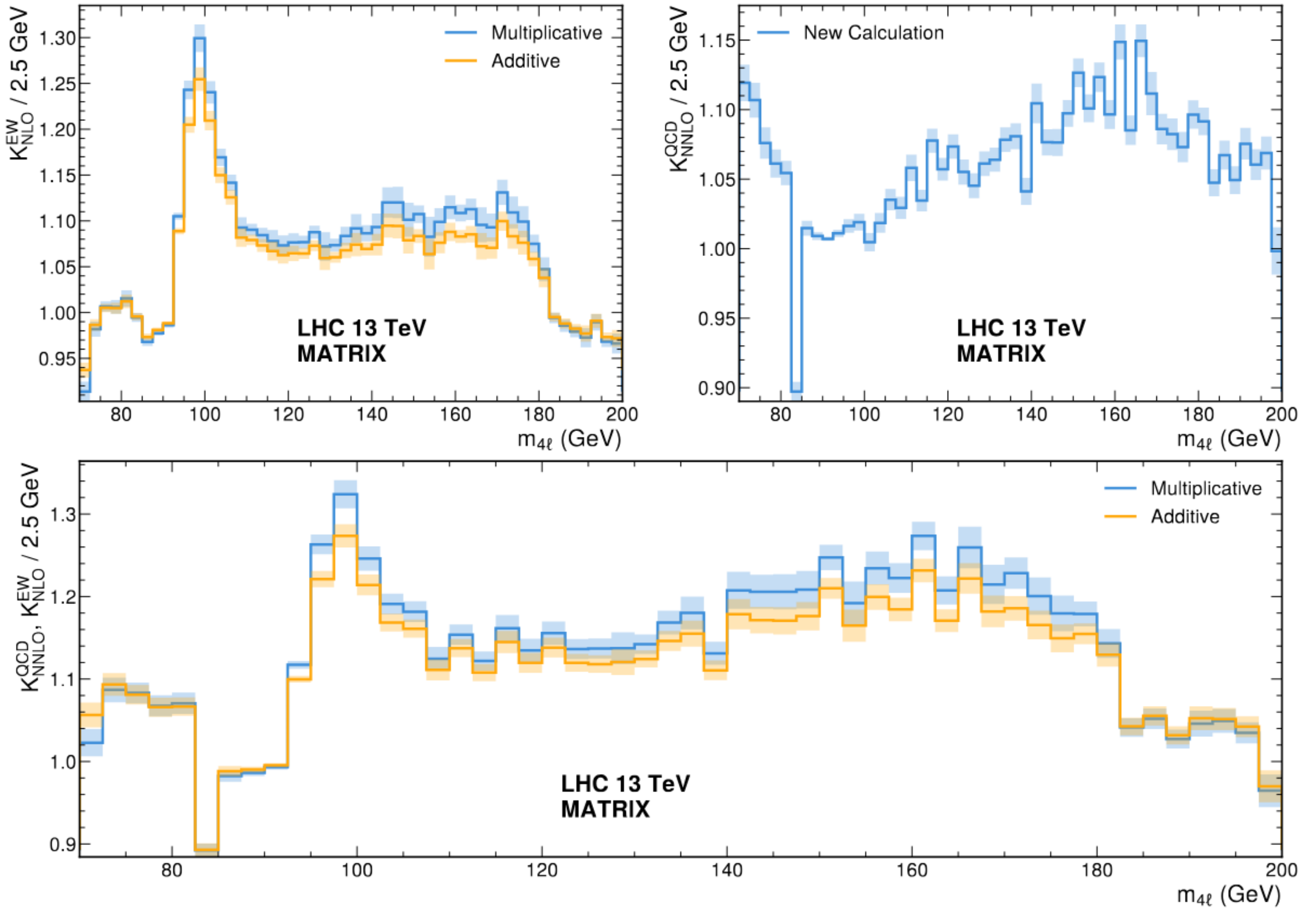
Dealing with EW Photon Emission

- The normal on-shell analysis region is also relatively flat in reconstructed MC
- NOT the case in GEN (necessarily)

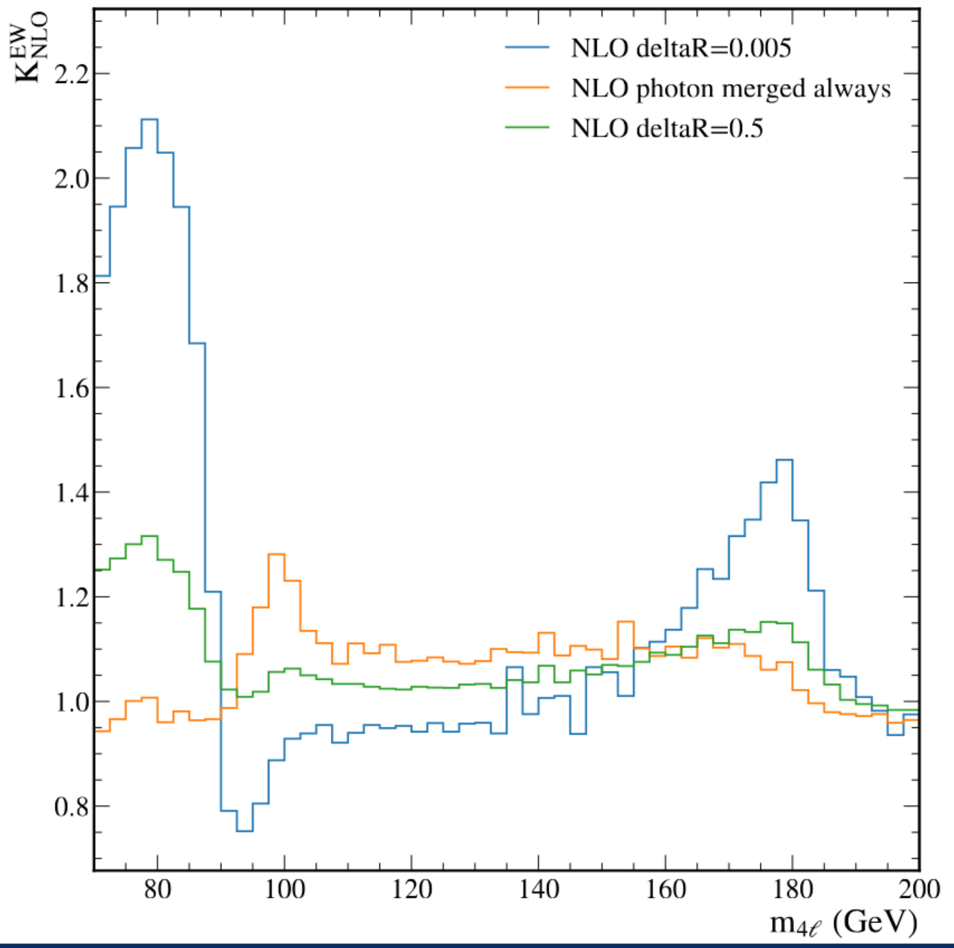


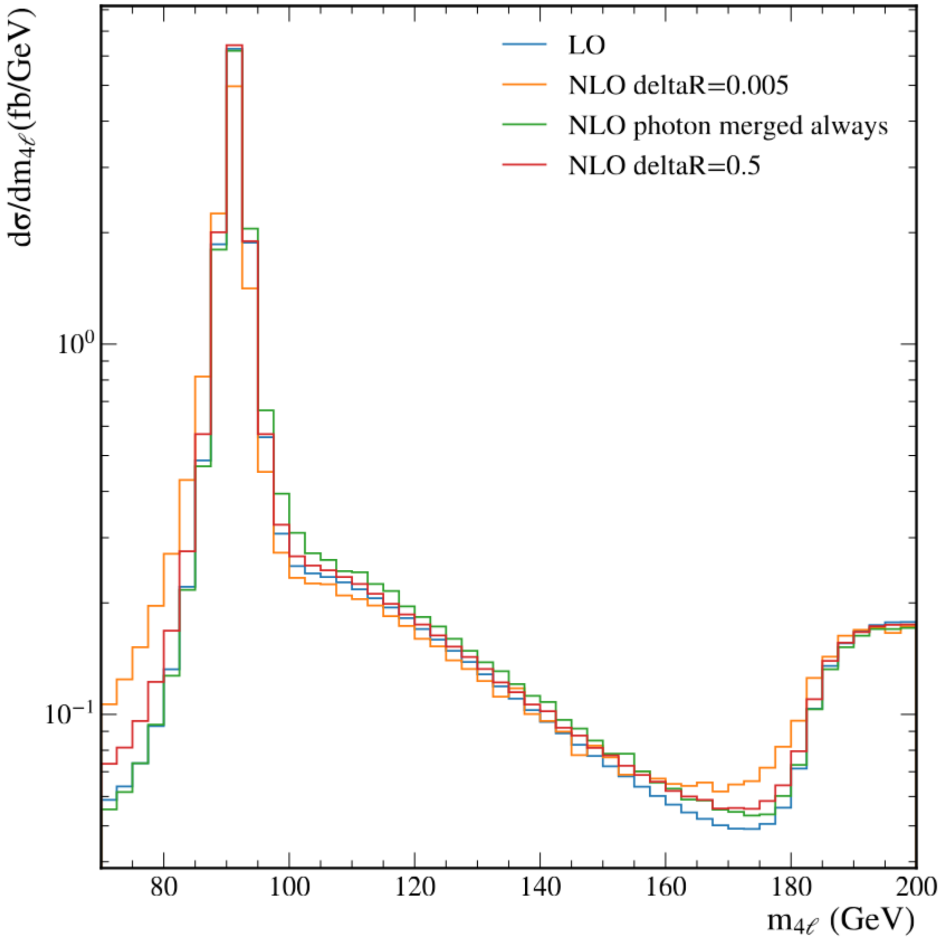
Dealing with EW Photon Emission

- We have the original distributions for $\Delta R = 0.1$
- Talked to MATRIX authors and changed the code such that the photon is always added into the mass distribution
- **Distribution gets shifted the other way!**

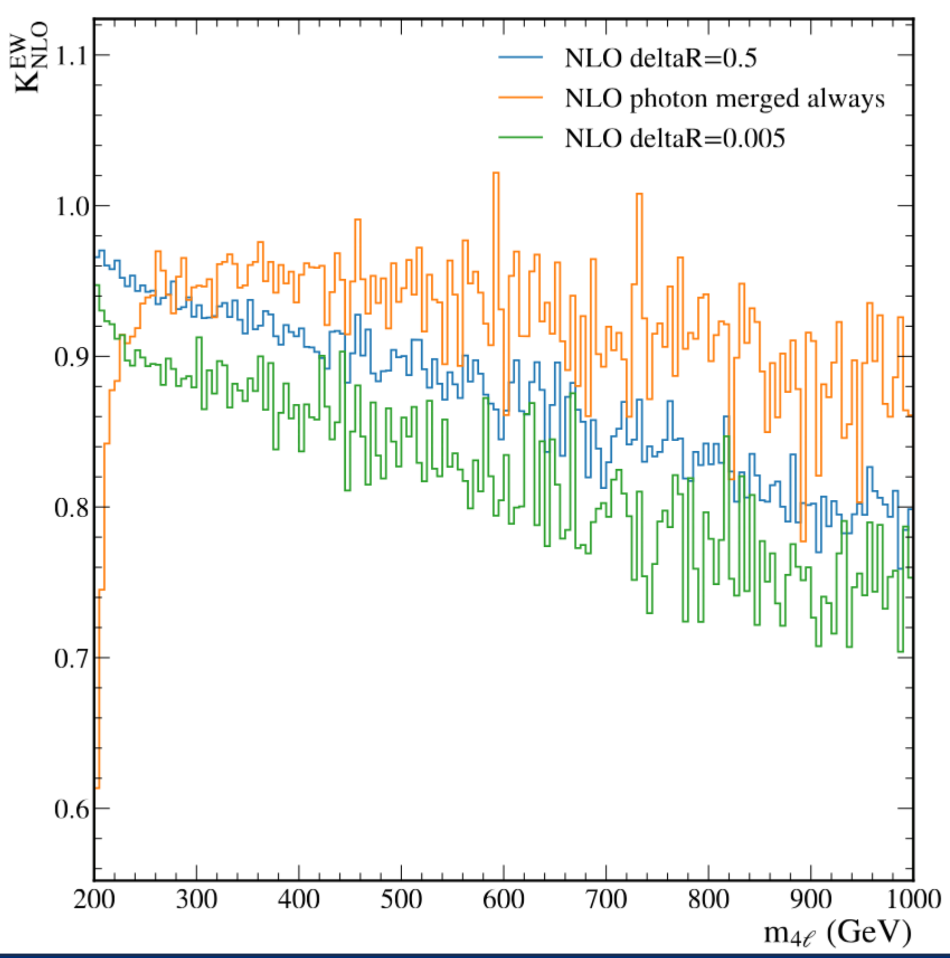


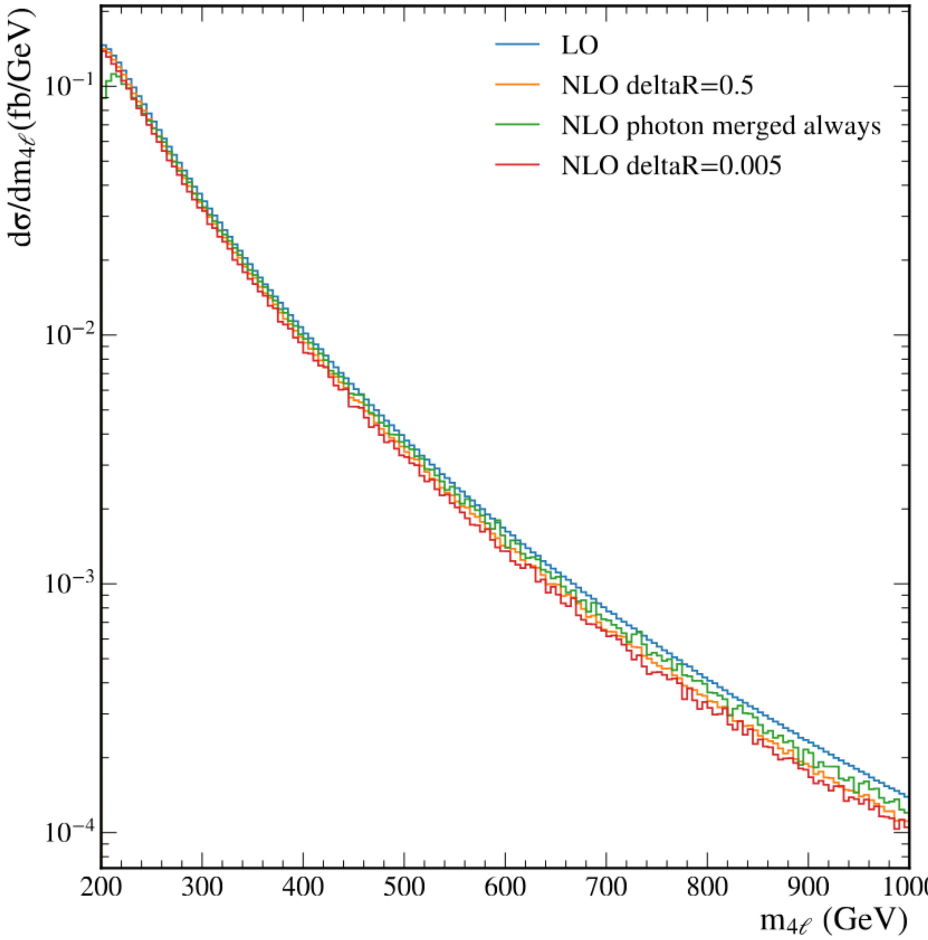
Dealing with EW Photon Emission





Dealing with EW Photon Emission





Dealing with EW Photon Emission

- Including every photon from ISR/FSR is probably the wrong thing to do

- $\Delta R = 0.005$ is too small, probably too much missing energy and/or unable to pass cuts without photon

- In our analysis, the value for $\Delta R = 0.5$, so that is what we will do moving forward

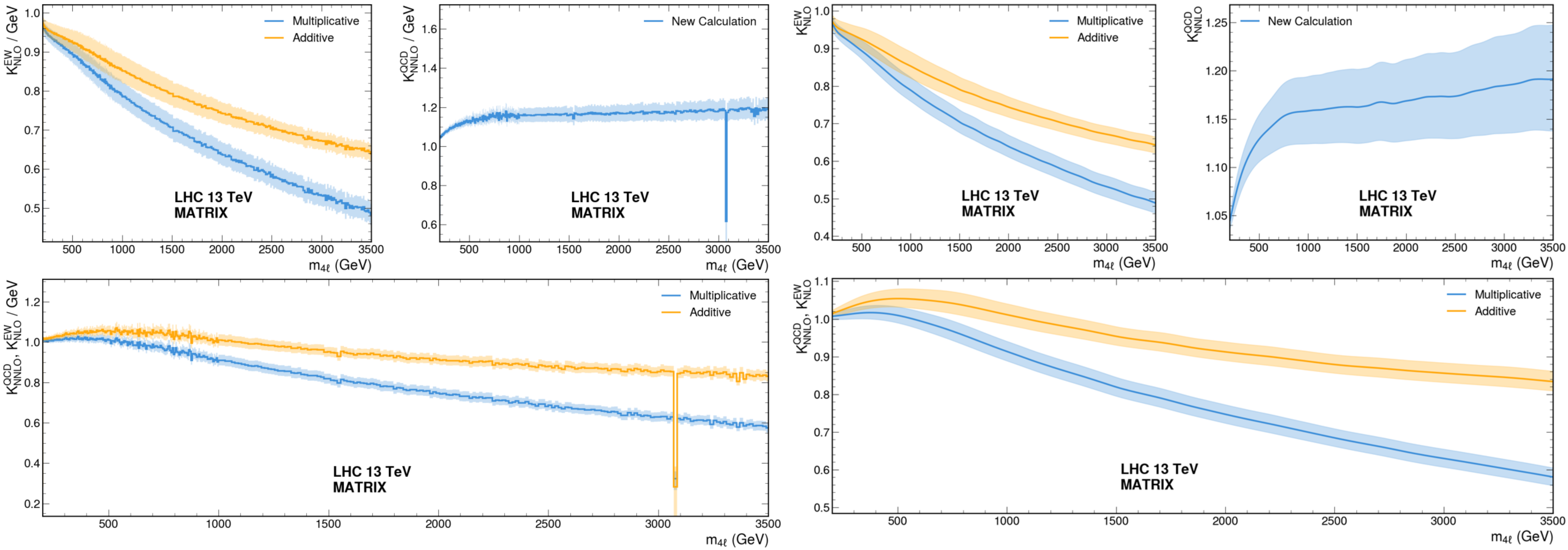
Total Calculation

- Currently, the only variable that we use if $m_{4\ell}$. We are investigating whether more a k-factor in more variables would help considerably.
- The results that we show are for the pure $m_{4\ell}$ k-factor
- Will also run an instance of NLO EW corrections for the 4e final-state to see if there are any large differences
- 2 options for systematics

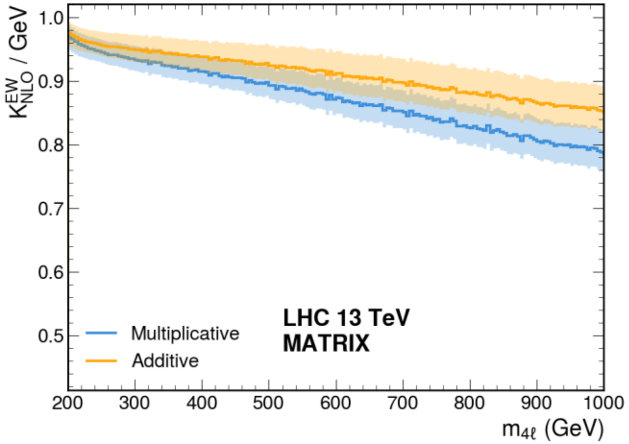
1. Have separate QCD and EW k-factors with their own systematic variations for $\mu_{F,R}$
2. Use MATRIX inbuilt functionality to calculate $K_{QCD} \times K_{EW}$ and $K_{QCD} + K_{EW}$ with a combined systematic variation for $\mu_{F,R}$

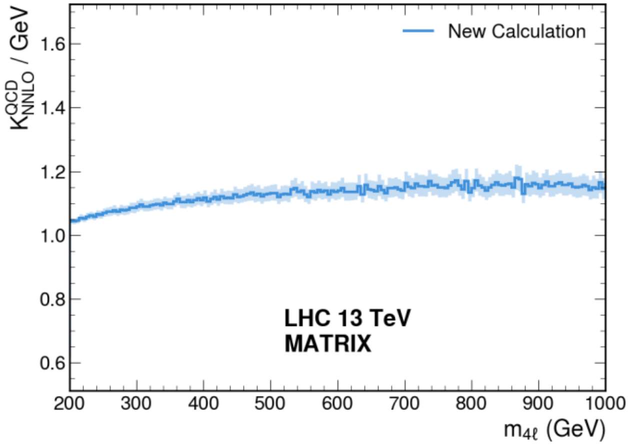
We prefer option 2

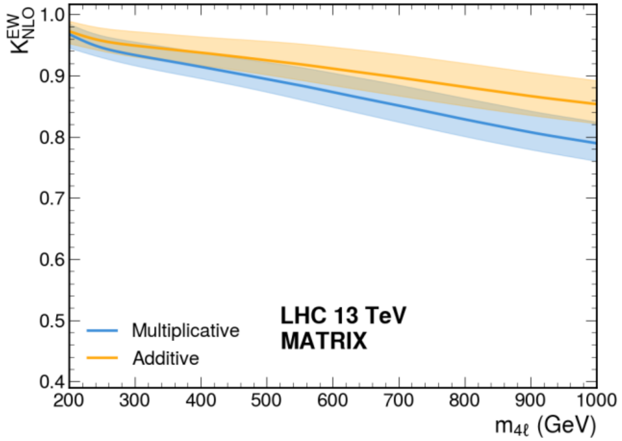
Total Calculation

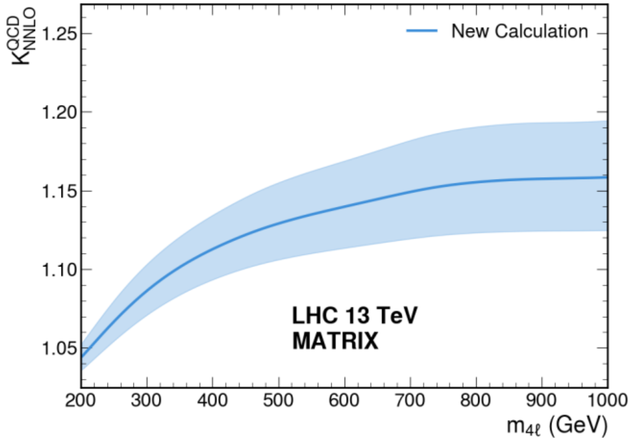


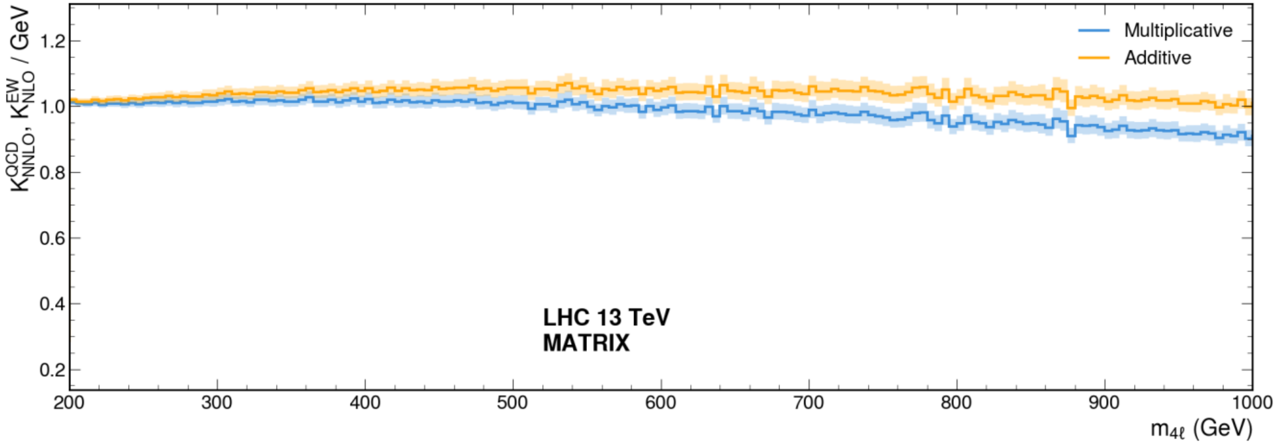
Total Calculation

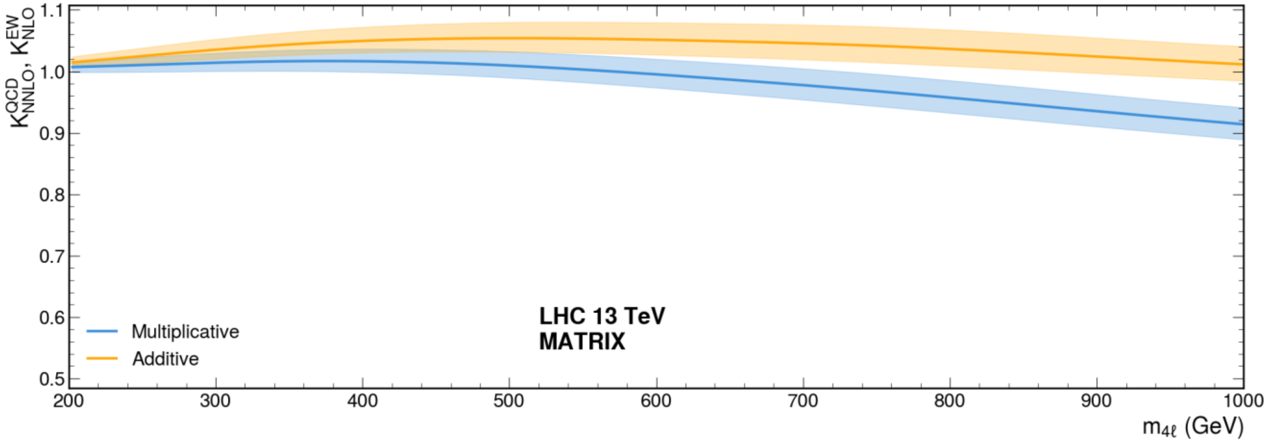












Systematic Variations

- μ_F and μ_R variations are handled internally by MATRIX in a 7-point variation
- K-factor's inherent uncertainty is in the difference between $K_{QCD} \times K_{Ew}$ and $K_{QCD} + K_{EW}$
- Recommend that the native sample's PDF uncertainties be used (unless someone wants to run MATRIX 100 times!)
  - If anyone has a different suggestion this would be useful
- α_S uncertainties **could** be done, but would require 2 more runs of MATRIX for the overall calculation

Further Work

- The assumption that the ggH k-factor can be used for signal processes holds, but adds an uncertainty to the background/interference distributions
  - One of the most impactful uncertainties in off-shell!

- New products for NLO calculation exist and can help to lessen this uncertainty
  - Gluon-fusion continuum POWHEG production is at NLO for H signal, diboson background, and interference
  - MATRIX has NLO gluon-fusion calculations available in QCD

- We have been trying to get MATRIX with NLO ggH to work to little success lately

- Personal suggestion: Submit central production to CMS for POWHEG samples (they are **very** useful for background with jets anyways) and check those cross-sections compared to LO.
  - POWHEG generates LO cross-sections anyways too
  - Go from LO to NLO with that factor, then NLO to NNLO with computed k-factor
  - Could make the case to make uncertainty smaller (maybe 5% instead of 10%)

Conclusion

- The k-factor for ggH and qqbar is now well documented and in the final stages of moving forward
  - We need to do a final investigation on the variables needed for qqbar and push forward with systematic variations for ggH

- It would also be a bonus to calculate the k-factors for signal, background, and interference separately for gluon-fusion to try and get the uncertainties down in analysis

- **These are important quantities that are beneficial to a wide array of analyses on both CMS and ATLAS**