TOP-25-003:
Pre-approval comments browse

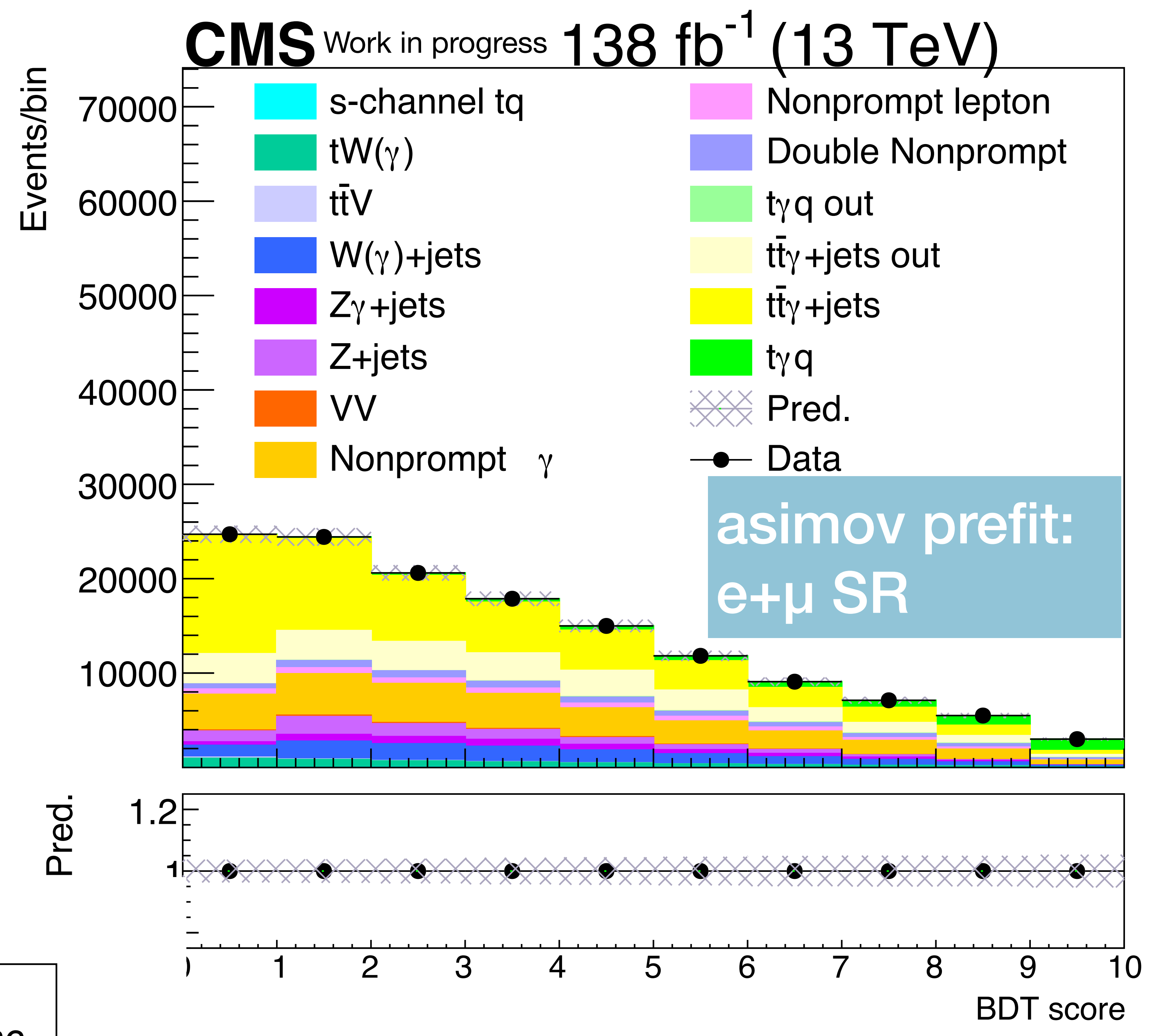
Ying AN, Maria Aldaya, Hugo Becerril, Abideh Jafari, Andreas Meyer

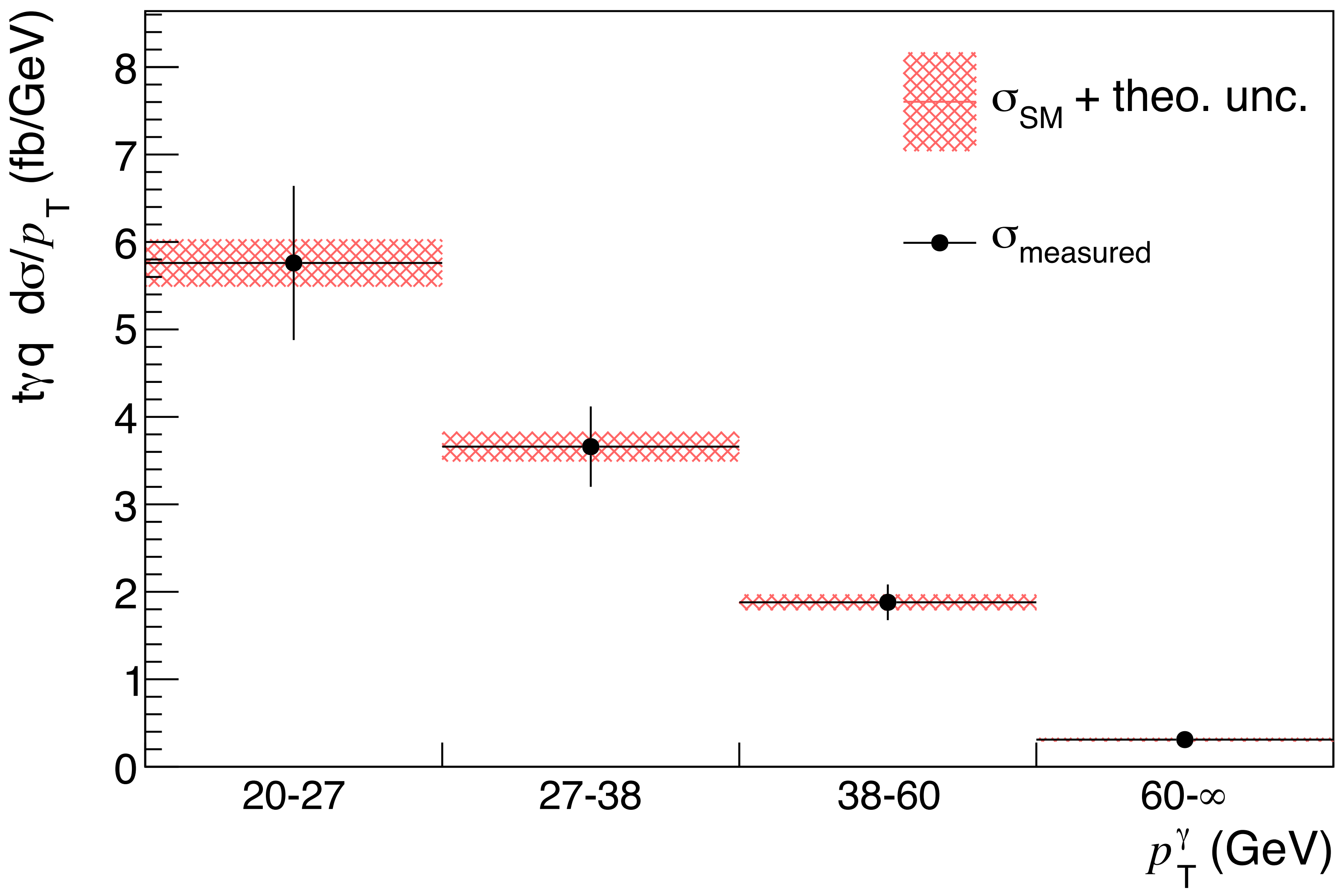
CMS TOP tX meeting

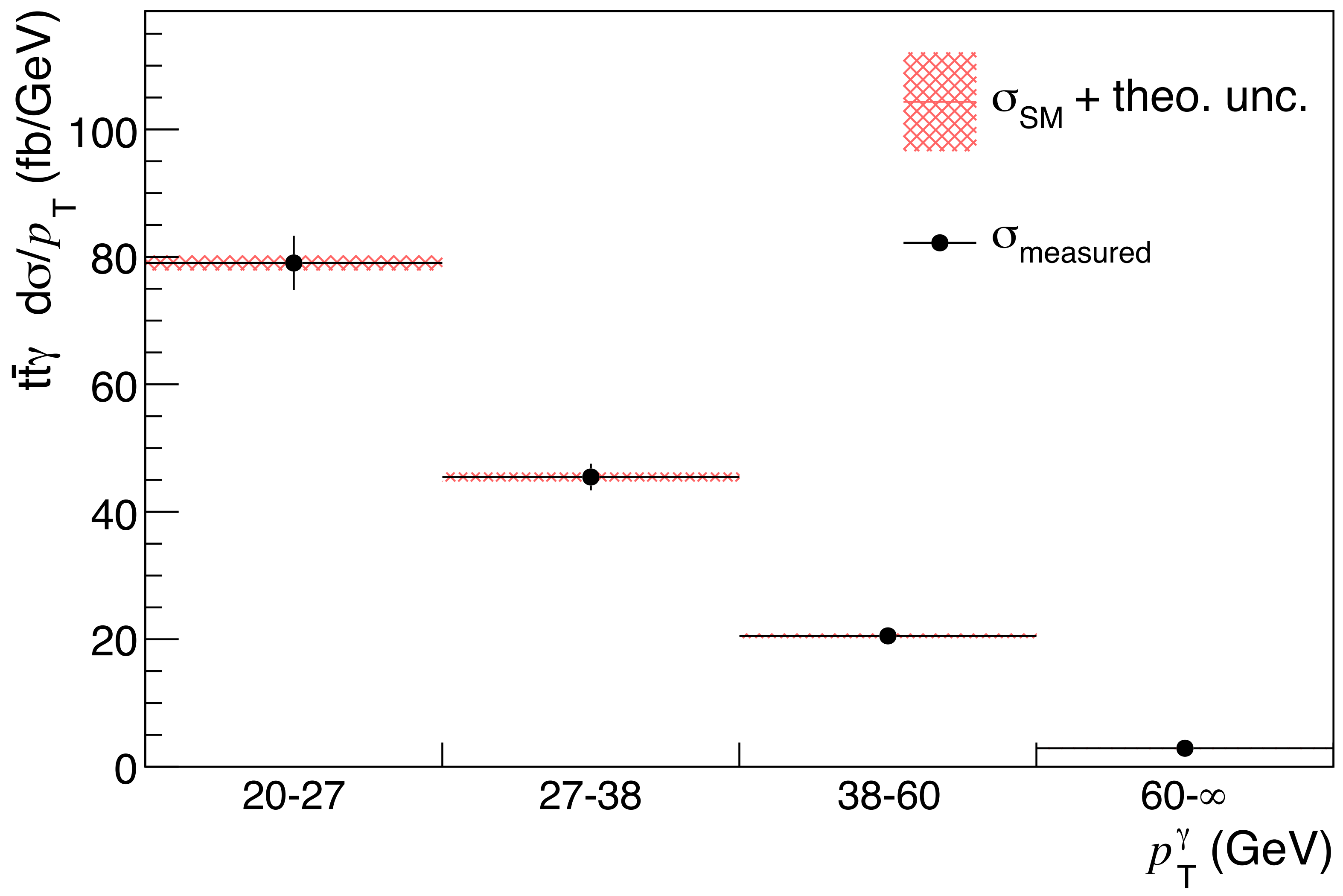
28 May 2025

Recap

- The pre-approval talk happened on April 2nd
- We showed some answers for the main pre-approval comments at the DESY top meeting on April 14th
  - The main change on selection is the photon from decay definition from $t\bar{t}$ shown in next slide and more uncertainties added







The expected inclusive fit results are:

$$r_{t\gamma q} = 1.0^{+0.081}_{-0.078} = 1.0^{+0.073}_{-0.070}(\mathrm{syst})\,^{+0.036}_{-0.036}(\mathrm{stat})$$

$$r_{t\bar{t}\gamma} = 1.0^{+0.043}_{-0.042} = 1.0^{+0.042}_{-0.041}(\mathrm{syst})\,^{+0.01}_{-0.01}(\mathrm{stat})$$

Comments posted during pre-approval talk

https://cms-pub-talk.web.cern.ch/t/follow-ups-from-pre-approval/36939

1.) Please quantify the expected contribution of non-Zjets fake photons from electrons, both in the signal regions and sidebands used to estimate the fake photon background. This background was quite significant in TOP-18-010 (sometimes they were of the same order of magnitude as the signal, see table 6 of the paper). Could you clarify why this analysis is so different that it is negligible in your case?

I wouldn't say that the e mis. γ is negligible in my case, it just gives us different performance in electron and muon channels.

In the muon channel, it is mainly from ttbar→2ℓ. Post 2018 MC as an example, the prompt and ele mis. γ yields from ttbar→2ℓ are around 4332 and 3508. And the ele mis. γ part doesn't show a dependency on any variable, moreover, we don't observe discrepancies out of the uncertainties either in VR or CR/SR low BDT regions, so we think these ele mis. γ from ttbar→2ℓ can be estimated well in simulation

However, In the electron channel, this part is pretty large with showing shape dependency and mainly from Z to ee. The discrepancies can be observed in all regions.

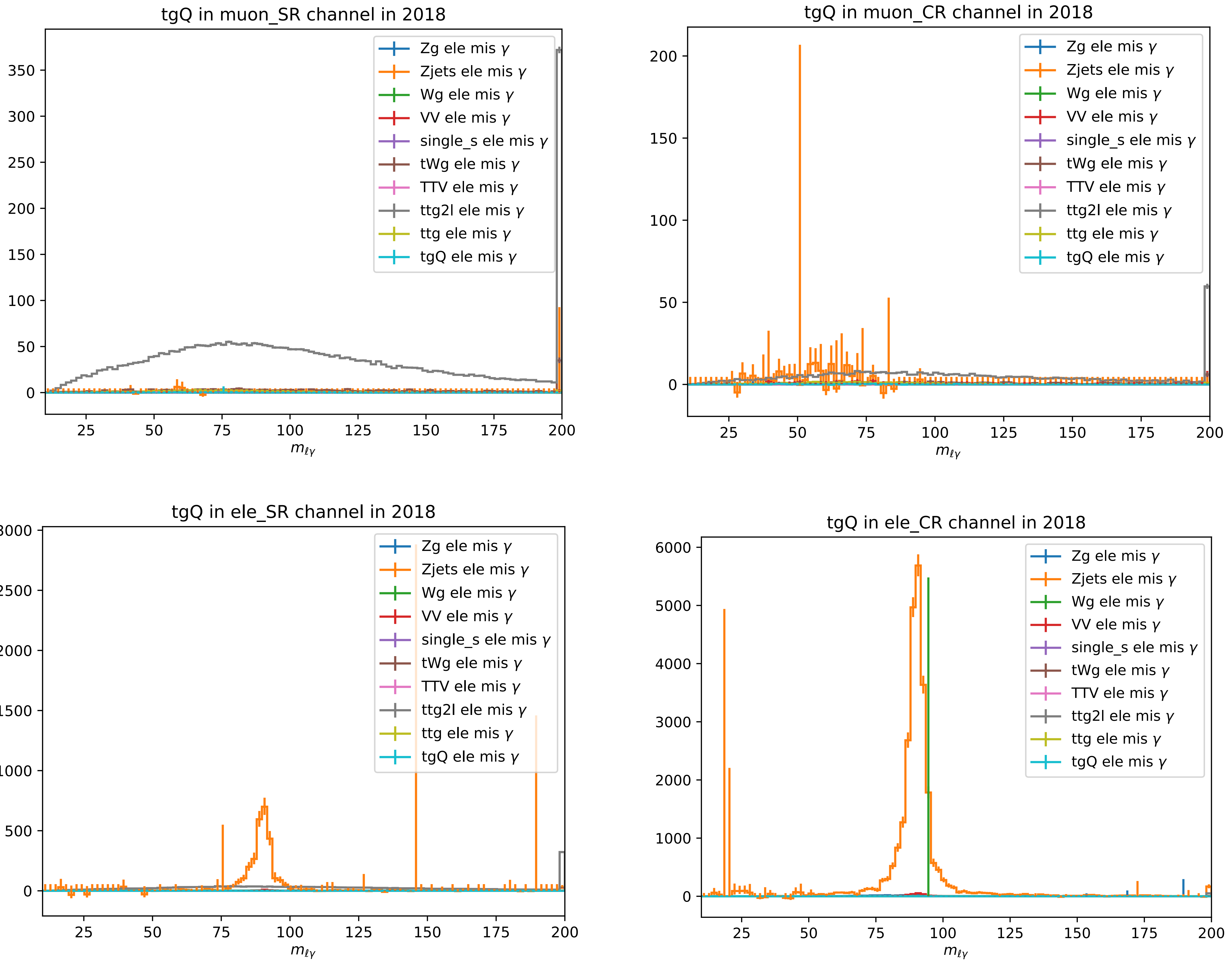
So we only consider this for the electron channel and estimate it by floating Z to ee normalization for the electron channel only and choose the mℓγ as the fit variable in the final fit.

Table 6: The observed number of events for the SR3 and SR4p signal regions in the e and μ channels, and the predicted yields and total post-fit uncertainties in each background component.

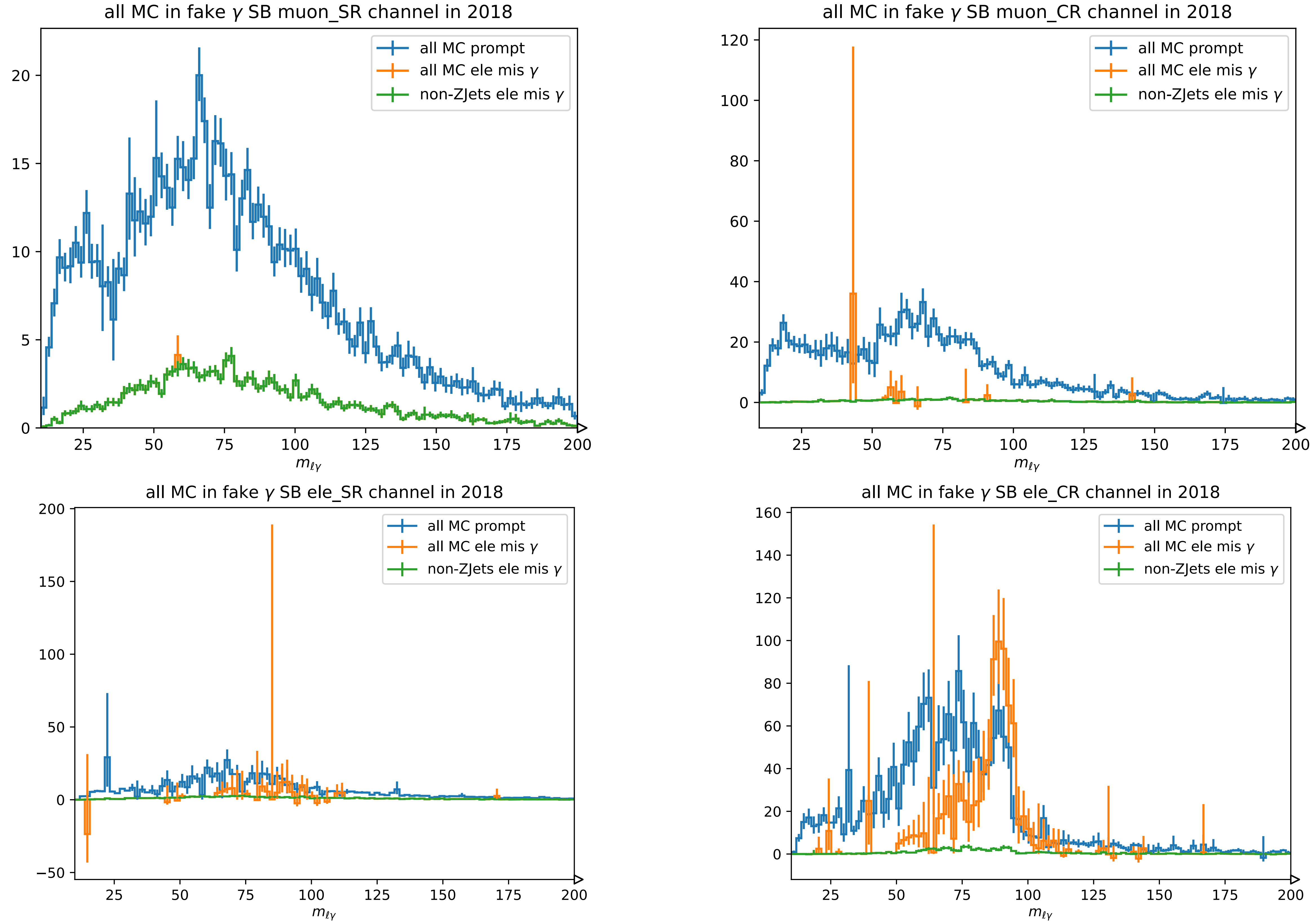
Process	SR3		SR4p	
	e	μ	e	μ
$t\bar{t}\gamma$	4995 ± 168	7821 ± 251	6174 ± 192	9495 ± 280
Misid. e	3710 ± 200	3322 ± 220	1904 ± 134	2015 ± 153
Nonprompt γ	2621 ± 107	4077 ± 161	2315 ± 124	3580 ± 149
Other	1136 ± 102	1866 ± 159	857 ± 110	1360 ± 166
Wγ	1082 ± 77	1486 ± 108	585 ± 48	864 ± 74
Multijet	560 ± 104	762 ± 140	302 ± 65	472 ± 102
Zγ	356 ± 38	640 ± 68	189 ± 25	306 ± 40
Total	14459 ± 178	19976 ± 196	12326 ± 150	18093 ± 173
Observed	14479	19885	12305	18184

1.) Please quantify the expected contribution of non-Zjets fake photons from electrons, both in the signal regions and sidebands used to estimate the fake photon background. This background was quite significant in TOP-18-010 (sometimes they were of the same order of magnitude as the signal, see table 6 of the paper). Could you clarify why this analysis is so different that it is negligible in your case?

ele mis. γ in CR and SR



Prompt and ele mis. γ in fake photon SB region

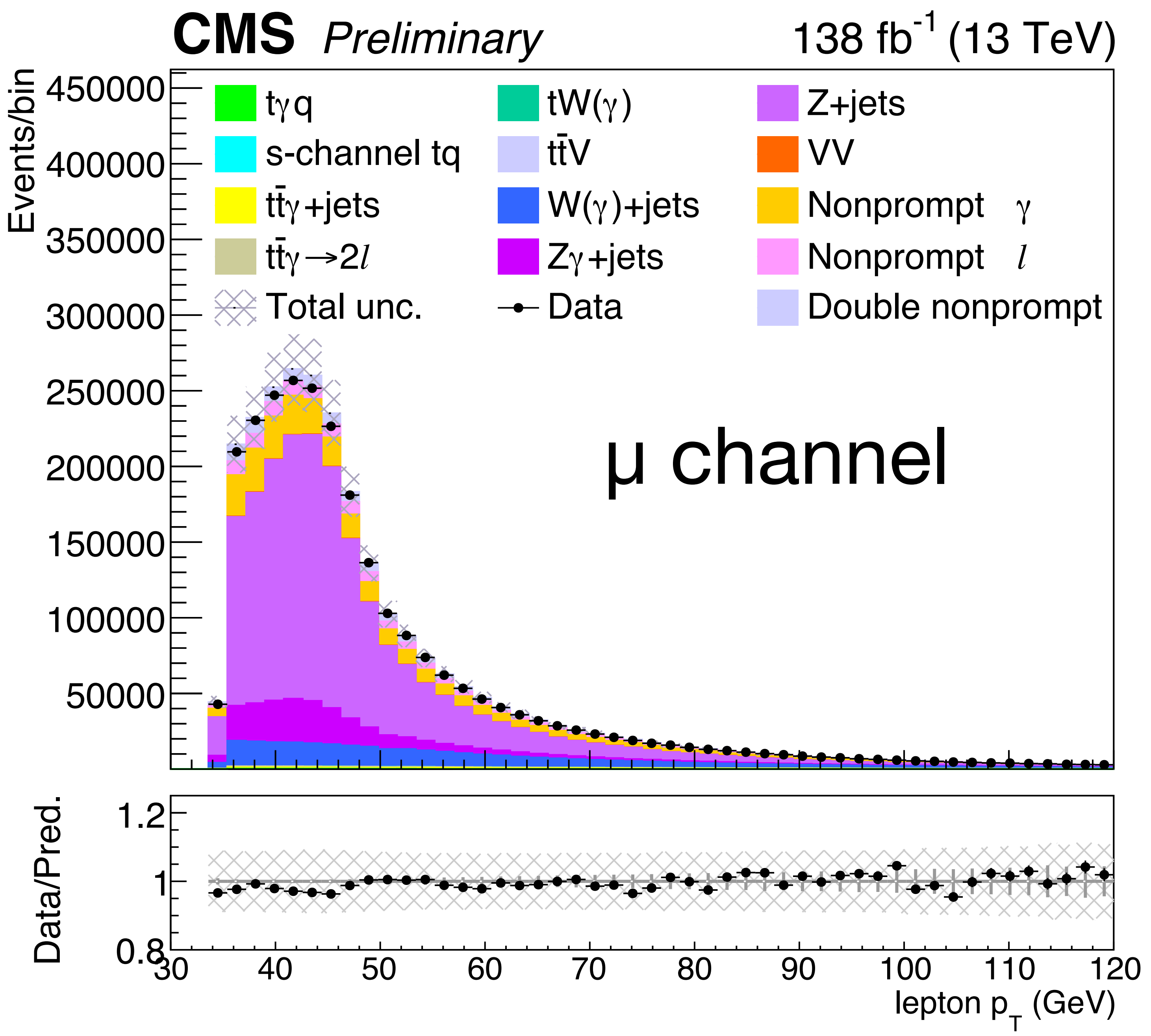


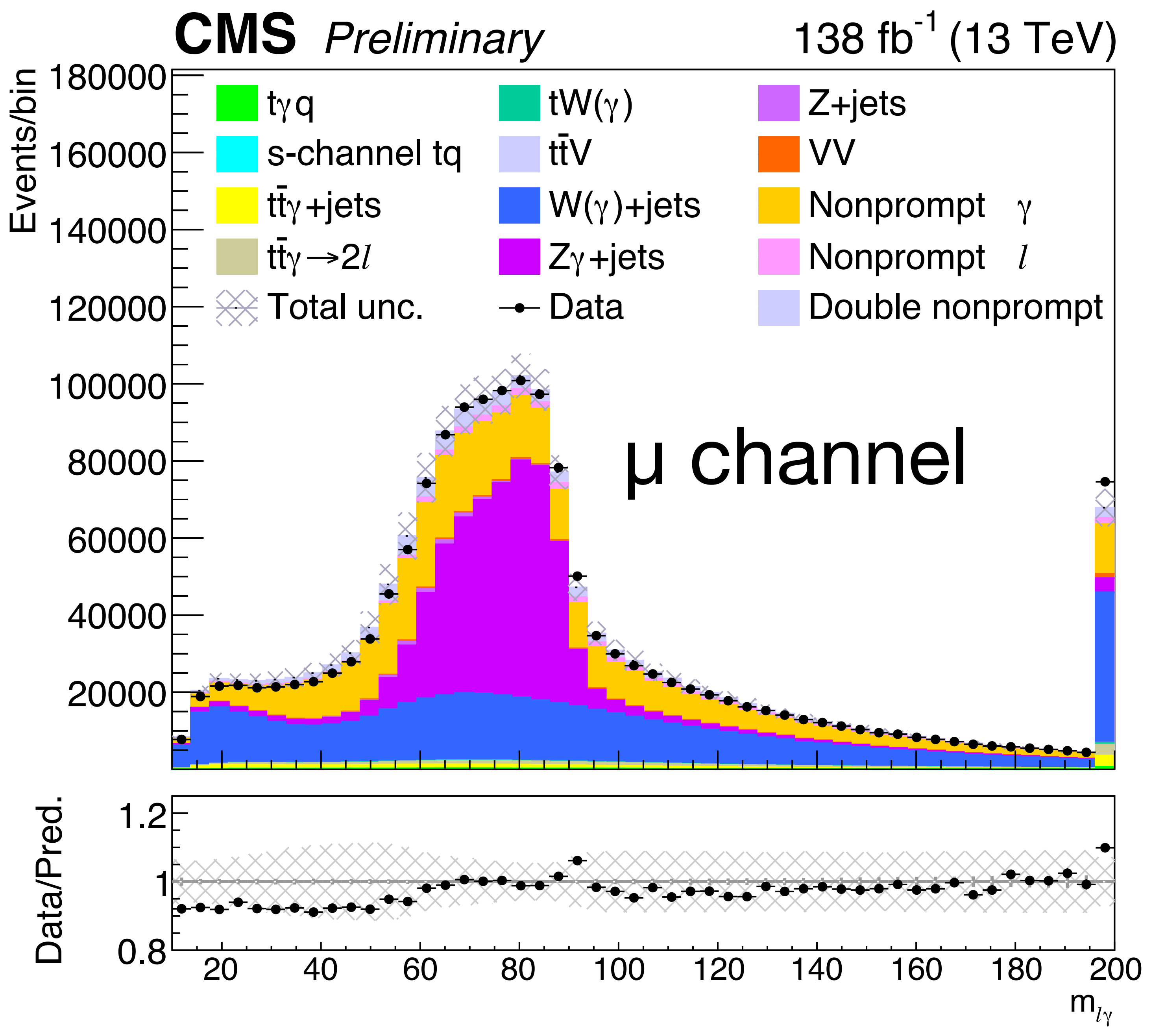
Small contribution in SB regions

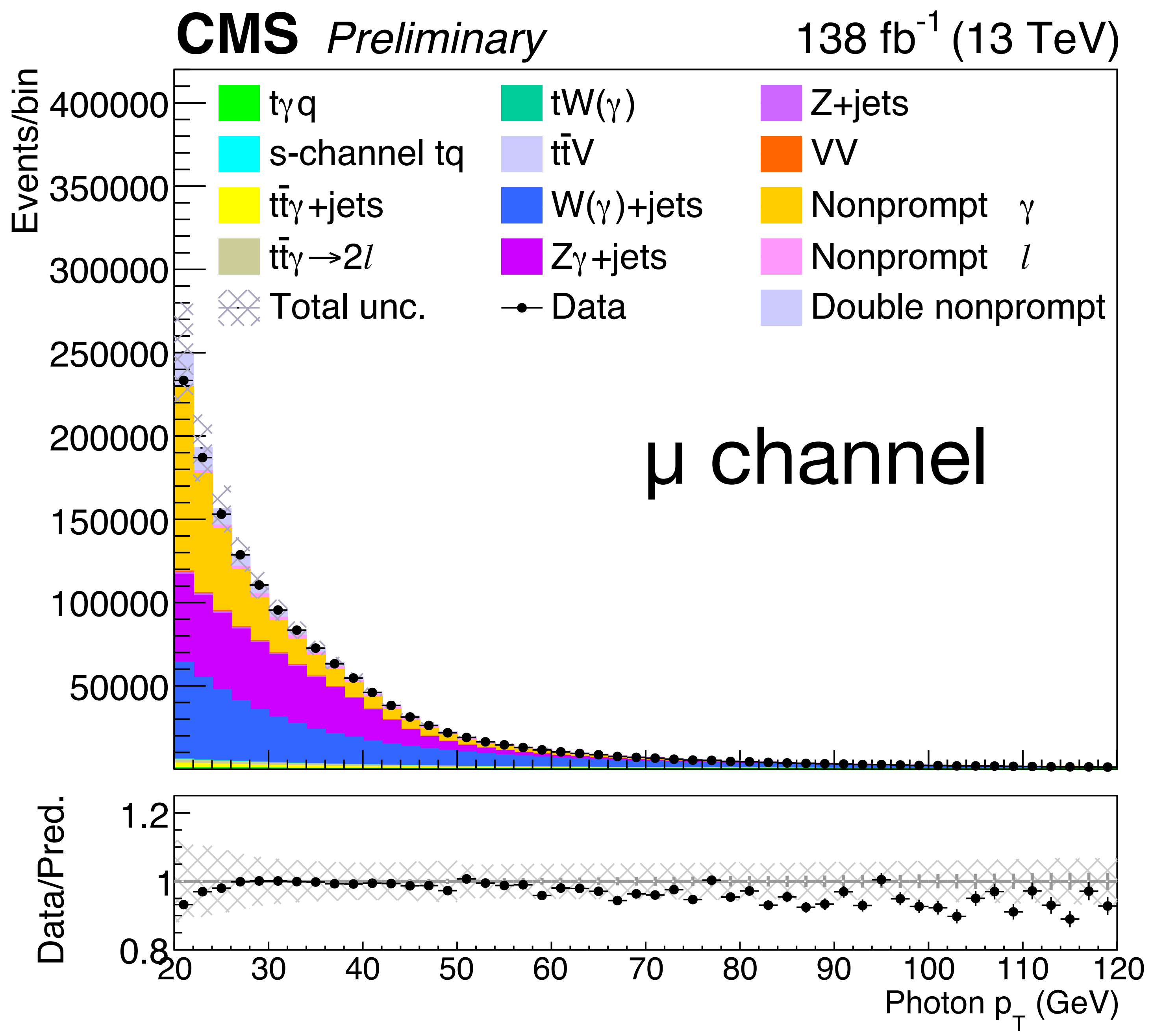
2.) Both electron and muon channels in VR show disagreement below the Z-peak, but in the muon channel, the uncertainty appears much smaller and the agreement is well outside of the uncertainty. Please check this. The point of the VR is to show that agreement is within uncertainty. Suggest looking into both whether the uncertainties on the plots are correct, and whether an additional uncertainty on photon fake rate is needed to cover the low part of the m_lgamma distribution.

I found the potential bug from the Rebin function in the ROOT. Ideally, we think uncertainties among bins are correlated, so when rebinning the histogram, the new uncertainty is the sum of uncertainties from nearby bins. However, the Rebin function of ROOT returns the square root of the quadratic sum, causing the underestimation. After fix it, the VR muon plots are as follows. **Within uncertainties, agreement is acceptable.**

On the other hand, additional handling is considered including remove electron and photons for 2016preVFP data and MC in an affected region, extra uncertainty in 2018 in BPIX affected region.







3.) Also regarding the VR, please explicitly state the Zgamma normalization factors used for these control plots, and compare to values from a CR-only partially unblinded fit.

Z+jets normalization factor in Vγ e channel

Year	e_minus	e_plus	ele
2016pre	1.56	1.55	1.55
2016post	1.67	1.67	1.67
2017	1.34	1.35	1.34
2018	1.40	1.40	1.40

- In the fit, the uncertainties applied to ZJets normalisation are:
  Zjets_norm_lightCR16, Zjets_norm_lightCR1718
  Zjets_norm_bSR16, Zjets_norm_bSR1718

- For a CR-only partially fit, we have
  Zjets_norm_lightCR16 **1.7 ± 0.2**
  Zjets_norm_lightCR1718 **1.2 ± 0.1**

4.) Please before unblinding, produce separate fits of the electron and muon channels.

The expected inclusive fit results:

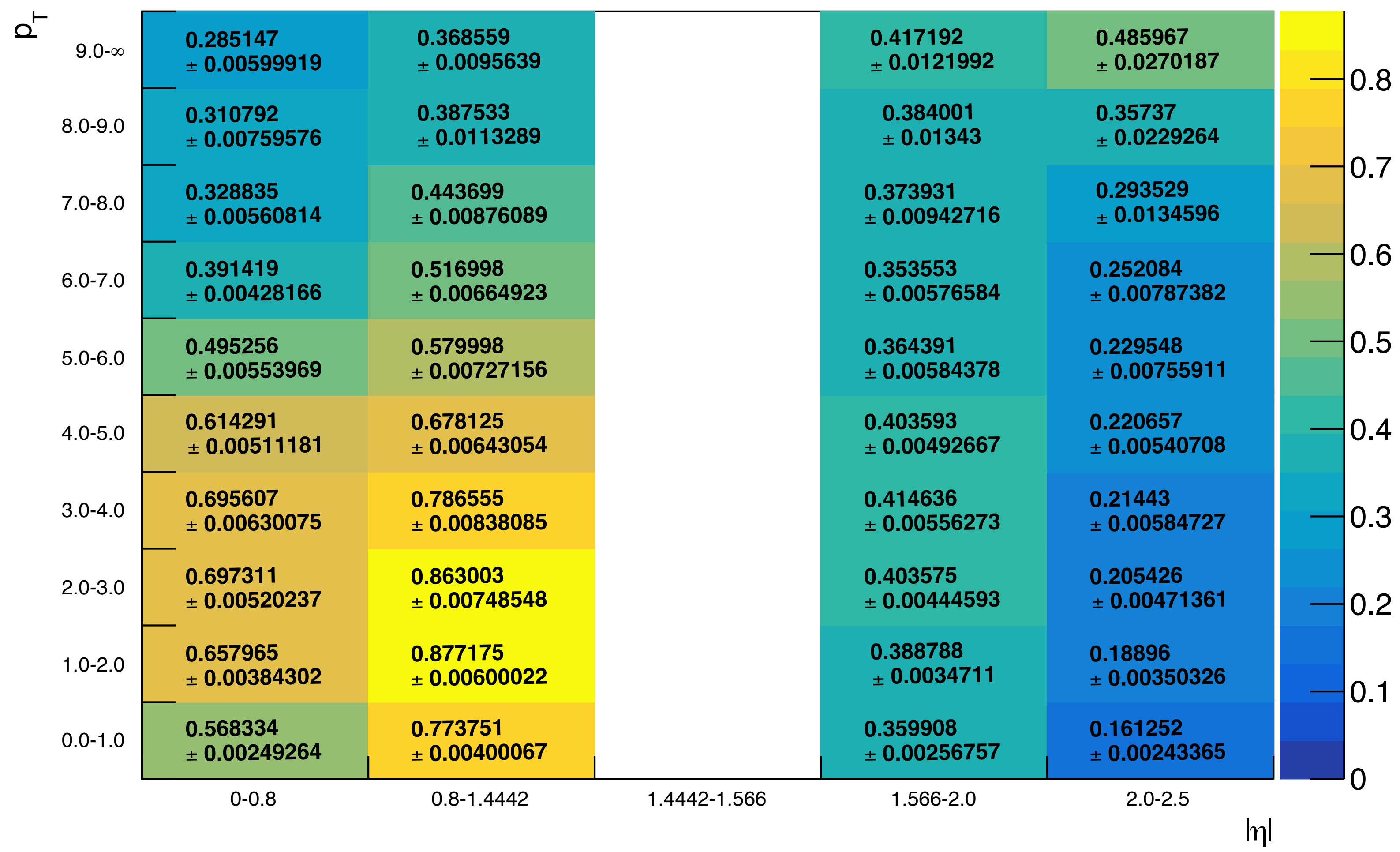
- Muon channel:
  - r_tgQ : +1.0 -0.109/+0.120 (68%)
  - r_ttg : +1.0 -0.054/+0.057 (68%)
- Electron channel:
  - r_tgQ : +1.0 -0.138/+0.153 (68%)
  - r_ttg : +1.0 -0.101/+0.094 (68%)

- Combined:
  - r_tgQ : +1.0 -0.100/+0.110 (68%)
  - r_ttg : +1.0 -0.050/+0.052 (68%)

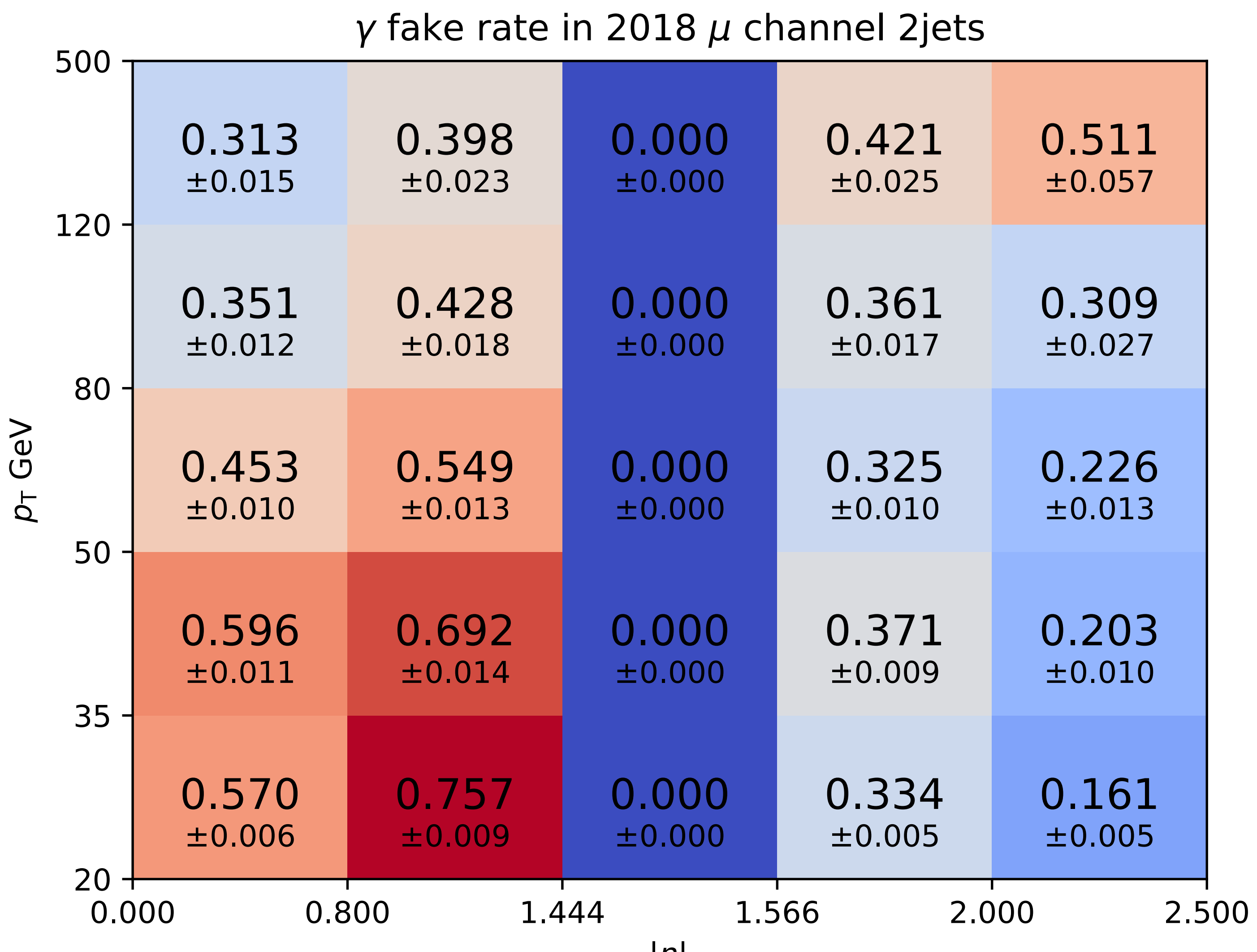
5.) Please check the effect of binning the photon fake rate more finely in photon pT, eliminating the Njet dependence if needed. This could be done as a cross-check, to see if it affects the background estimate or result.

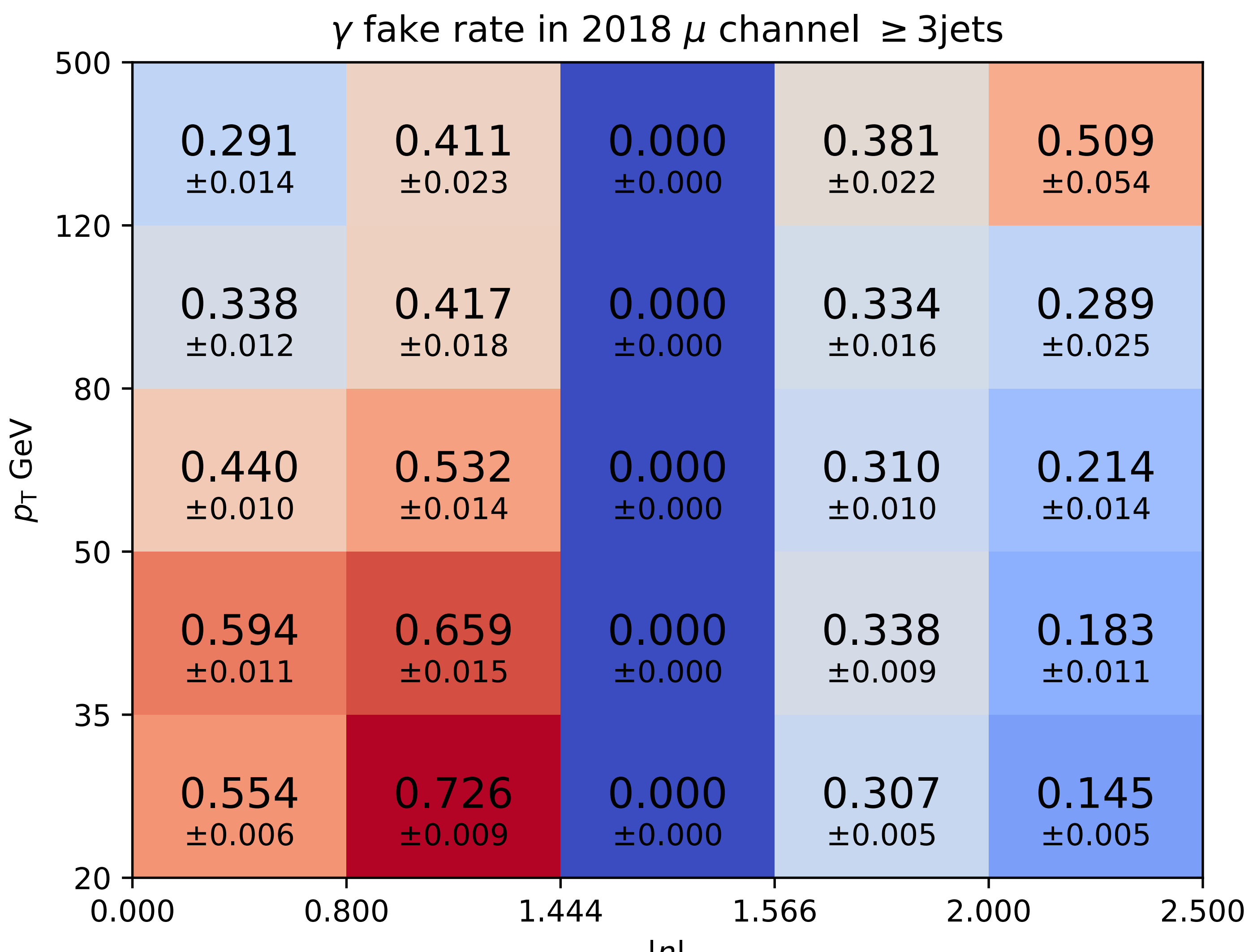
- Photon p_T: [20,25,30,35,40,50,60,80,100,120,∞]
- Photon |η|: [0,0.8,1.4442,1.566,2,2.5

2018 fake photon weight μ channel



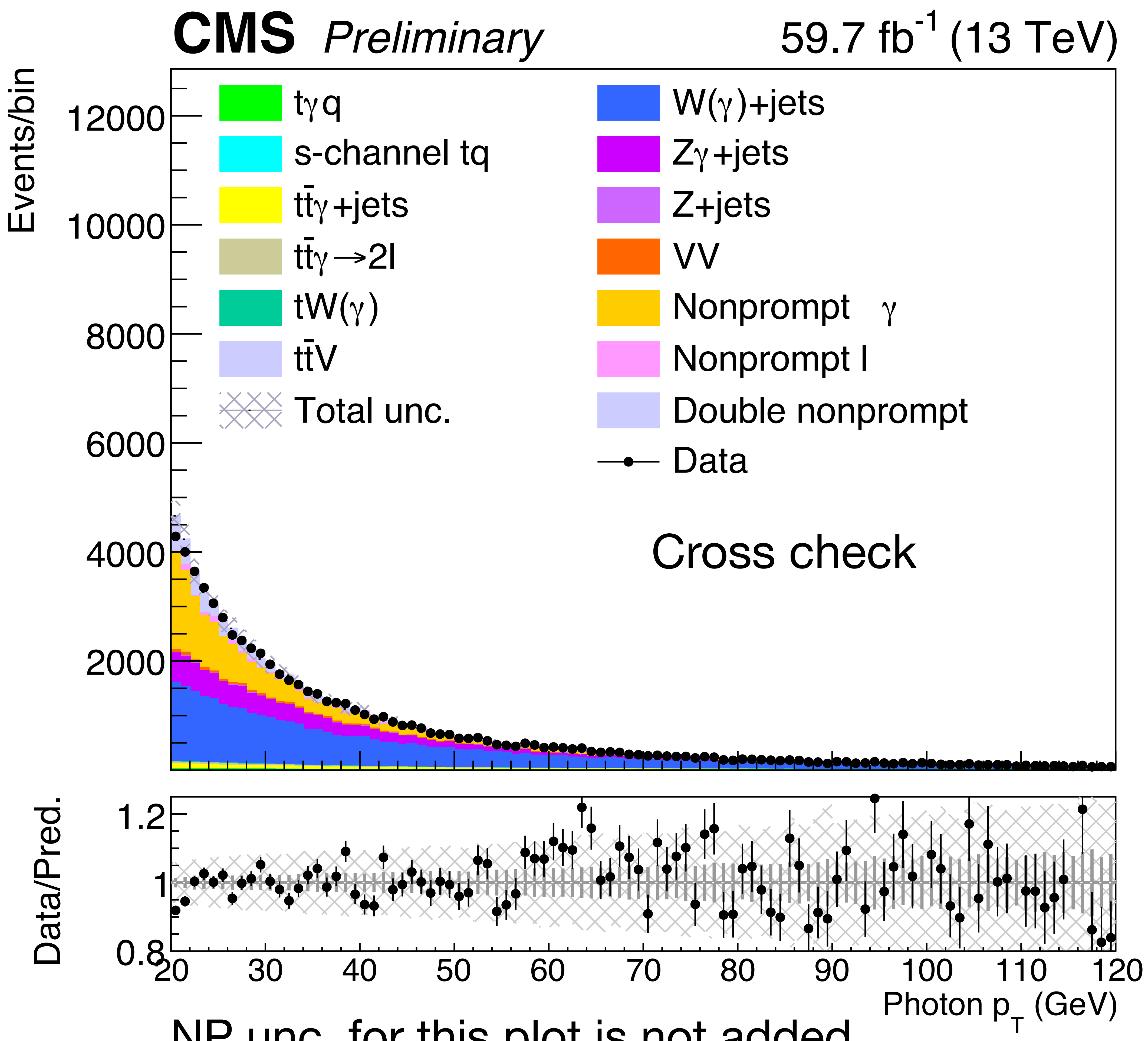
- Photon p_T: [20,35,50,80,120,∞]
- Photon |η|: [0,0.8,1.4442,1.566,2,2.5
- Nj: [0,1,2,3,∞]



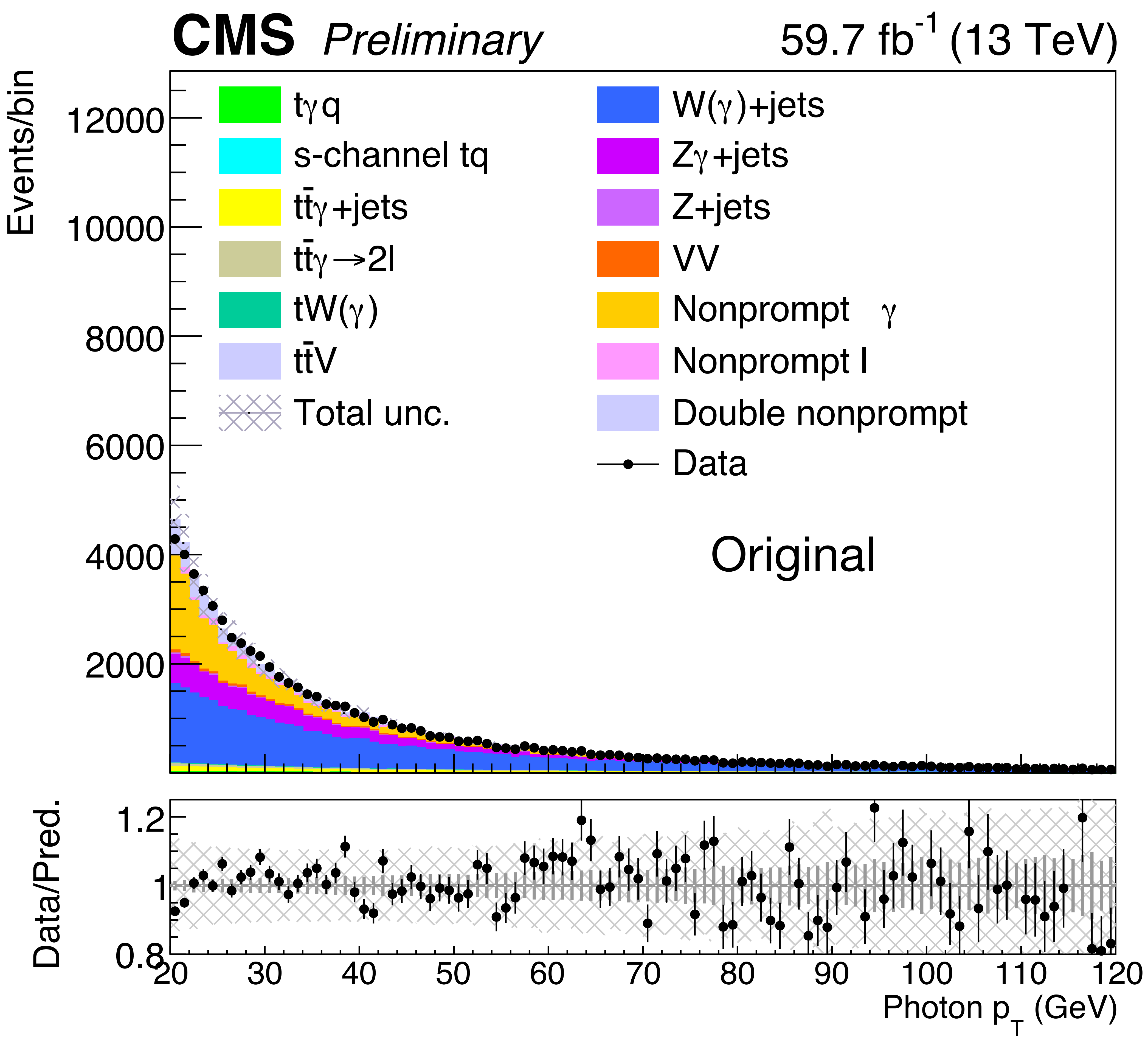


5.) Please check the effect of binning the photon fake rate more finely in photon pT, eliminating the Njet dependence if needed. This could be done as a cross-check, to see if it affects the background estimate or result.

Using 2018 data and MC to do this cross check, the agreement doesn't change obviously in our CR even in the photon pT distribution. It's OK to keep our current binning.



NP unc. for this plot is not added



6.) Please check the significance for t-gamma-q and report in the AN. A statement should be made in the paper at least indicating that it is well above 5 sigma, since CMS has not yet observed the process.

Small modifications are applied to fit the significance data-cards

- Require tγq events to pass reco-level selection only
- tγq is as the only signal, ttγ is as background
- Use full theoretical uncertainties not only shape effect

```
combine -M Significance work_space.root --expectSignal=1 -t -1
```

The expected inclusive fit results:

- Muon channel: 14.1 σ
- Electron channel: 9.68 σ
- Combined: 17.0 σ

7.) Wgamma/Zgamma: clarifying the last round of comments, one way to derive an Njet shape uncertainty is to create a shape template with a sloping trend in Njet large enough to cover the observed disagreement in the CR. This is what was done for Zgamma in TOP-23-002.

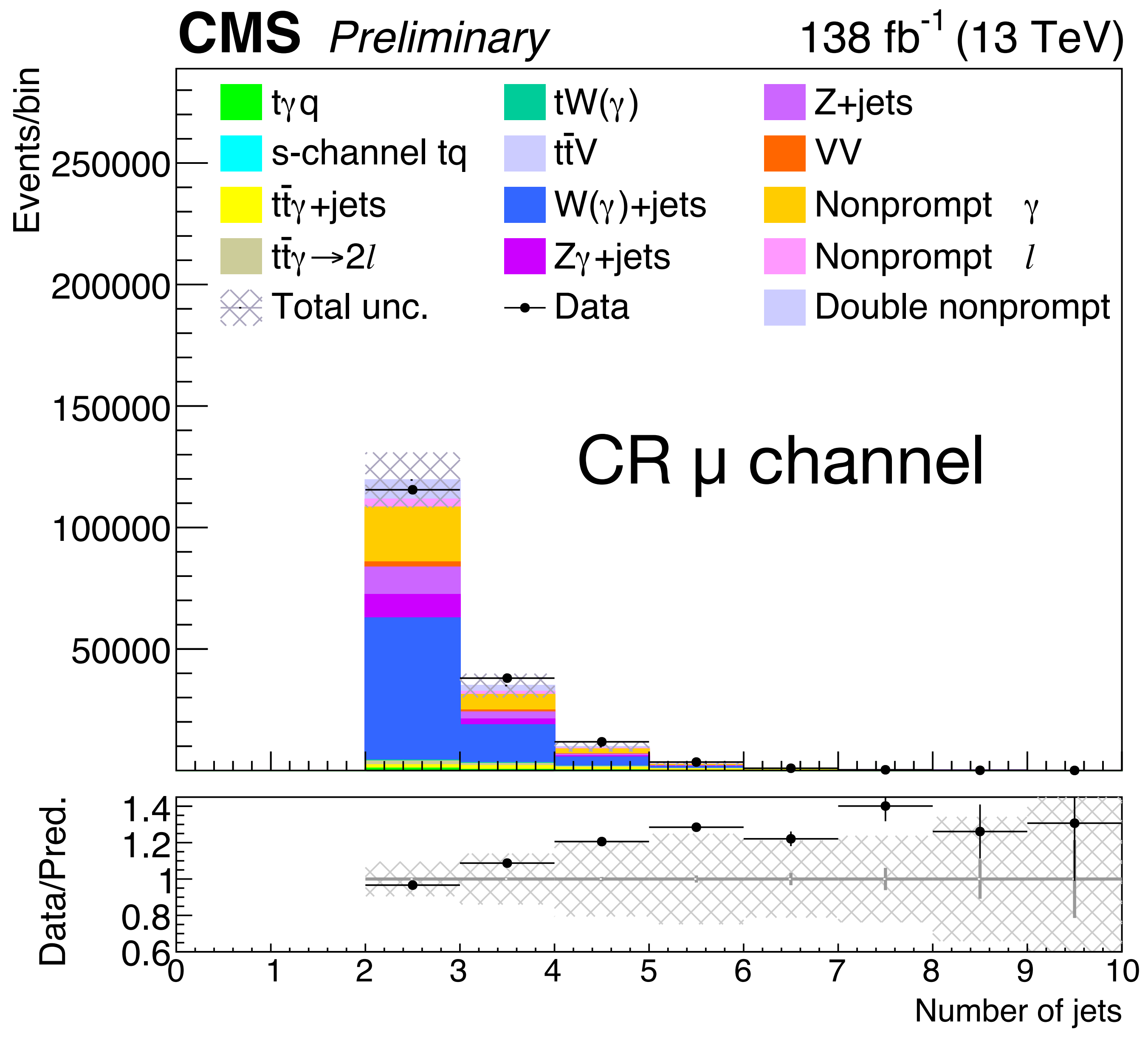
Uncertainties are calculated as function of number of jets are applied to

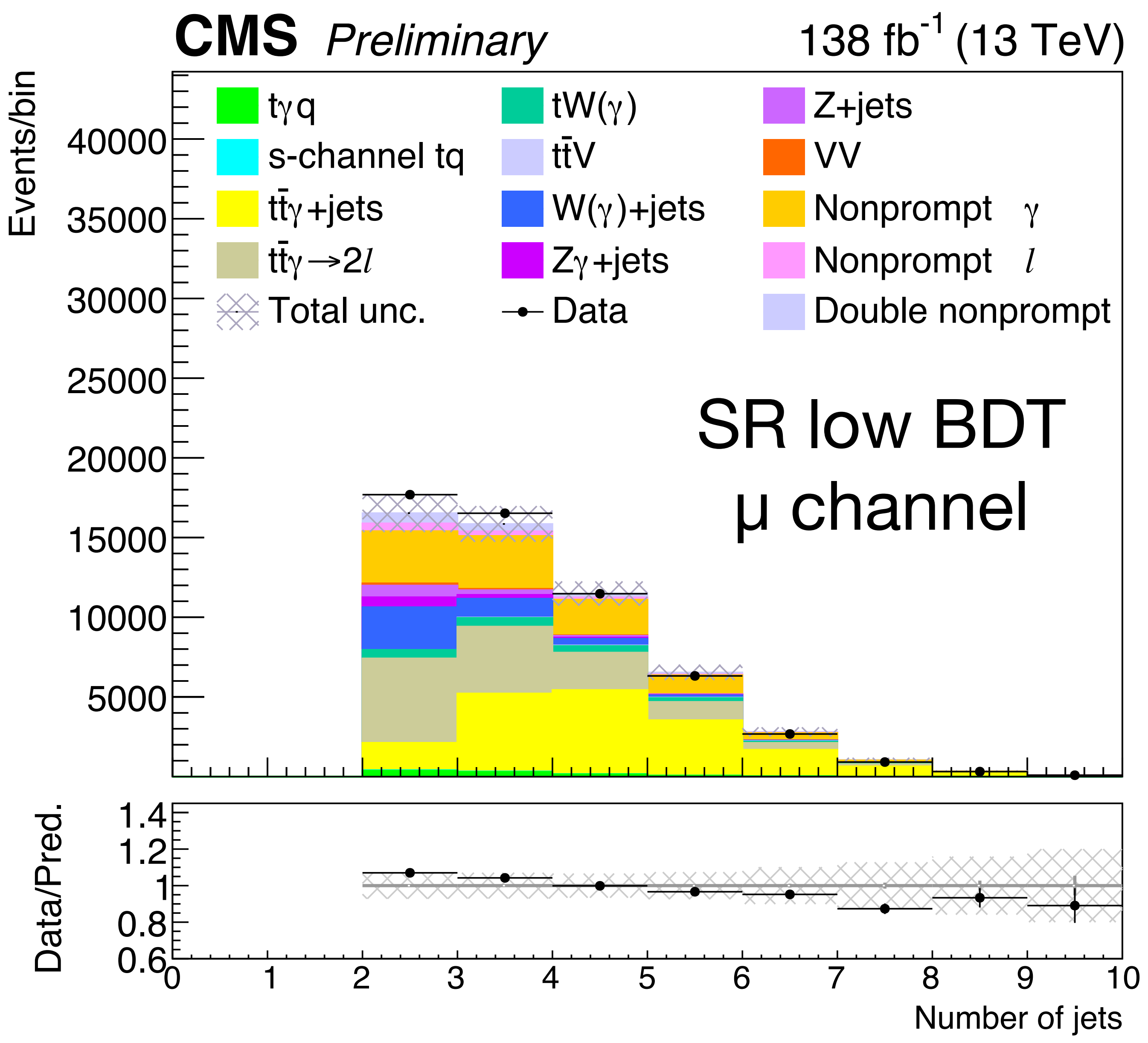
- Wγ and Zγ in muon channel
- Wγ, Zγ and ZJets in electron channel

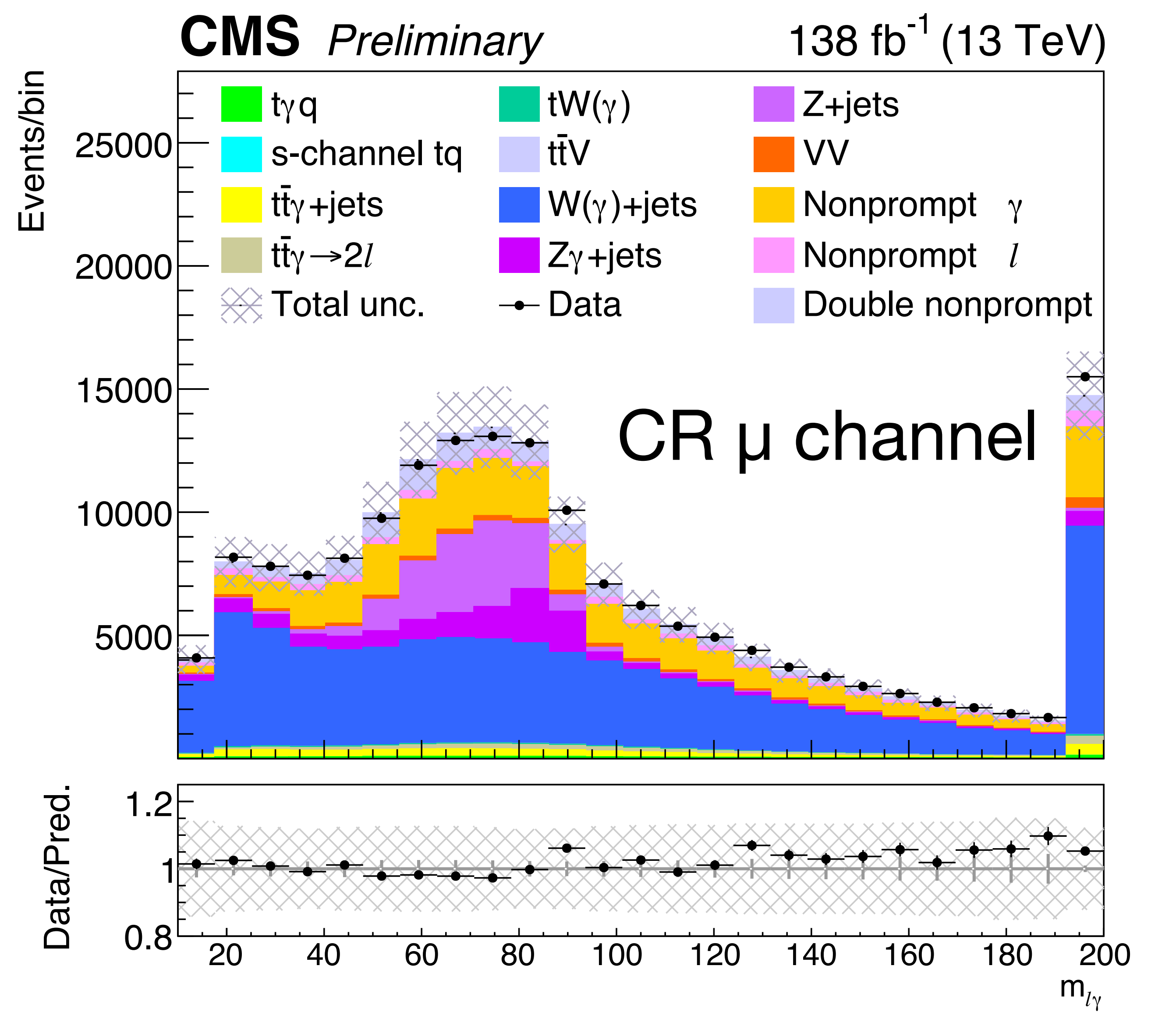
Uncertainties applied in the fit are:

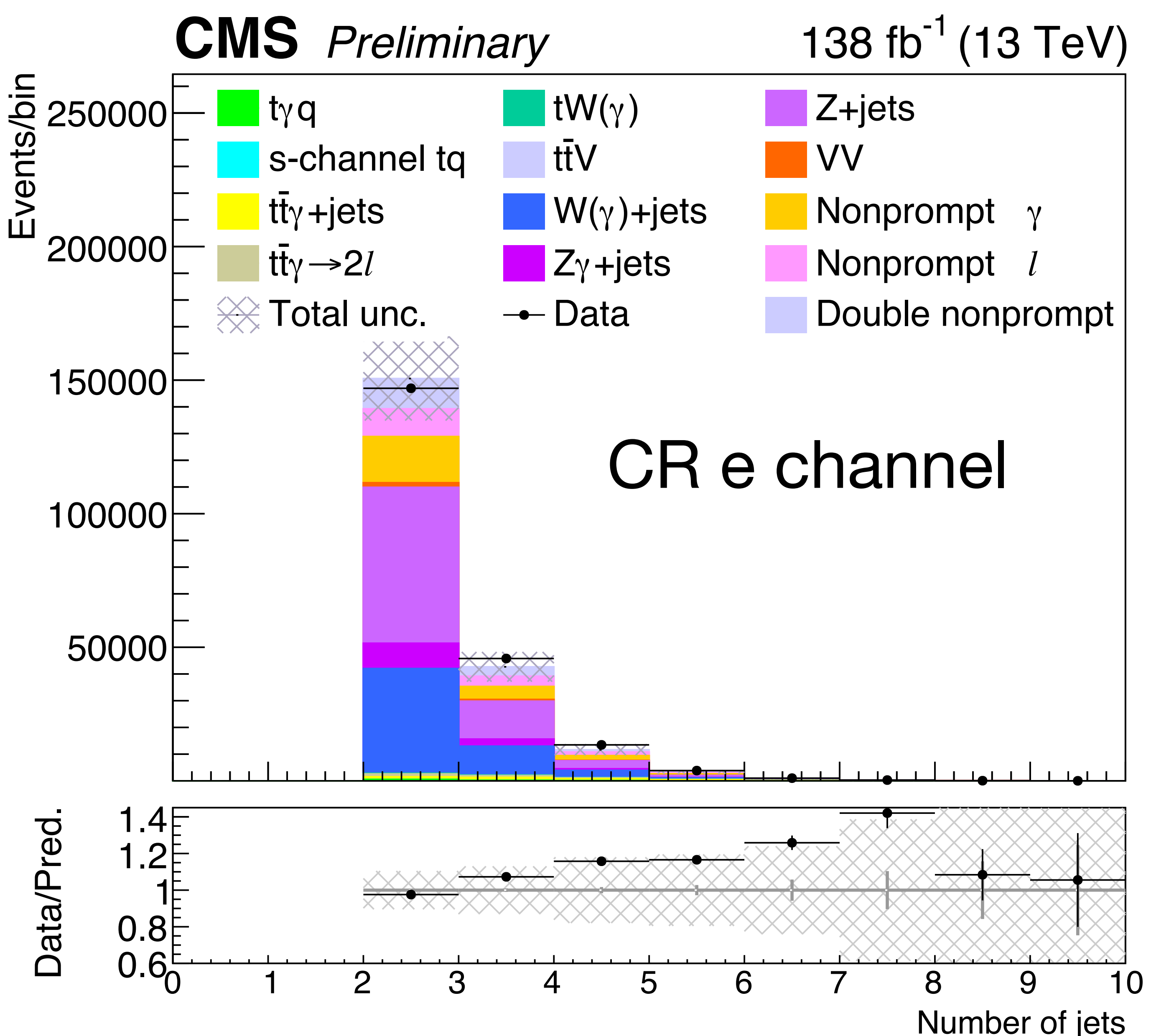
- Wg_njets_SR, Wg_njets_CR for both electron and muon channels
- Zg_njets_SR, Zg_njets_CR for both electron and muon channels
- Zjets_njets_SR, Zjets_njets_CR for only electron channel

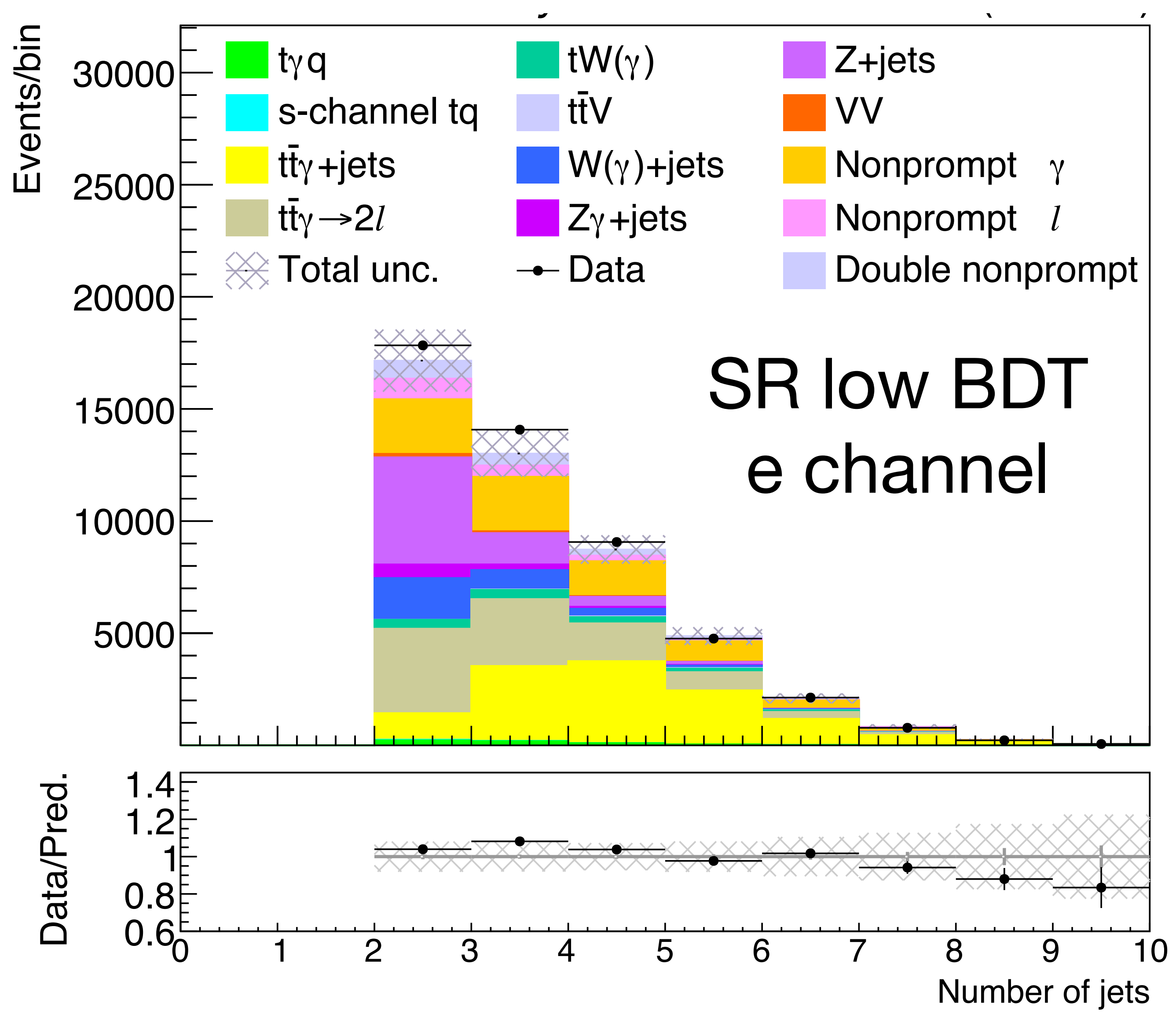
7.) Wgamma/Zgamma: clarifying the last round of comments, one way to derive an Njet shape uncertainty is to create a shape template with a sloping trend in Njet large enough to cover the observed disagreement in the CR. This is what was done for Zgamma in TOP-23-002.

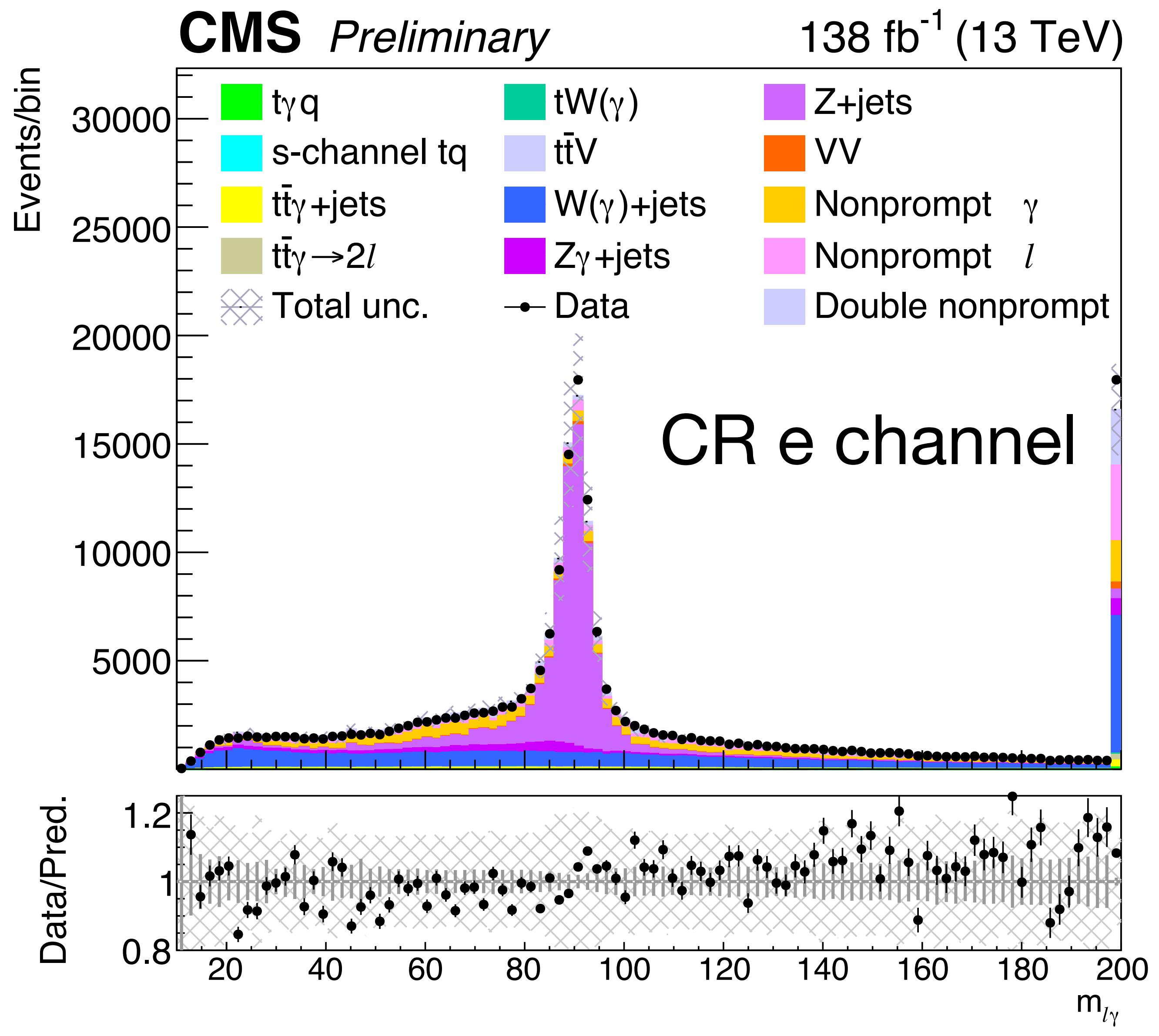












8.) Photon from decay definitions (Table 6 in AN 26/3/2025): We need to follow up jointly with Beatriz to make sure that in cases where these differ from TOP-23-002, the change is well-motivated. Currently, the photon definitions make the two analyses hard to compare. The values chosen for TOP-23-002, for example the deltaR cone cutoff, were validated and we need to have the same level of investigation here if not using the same definitions.

Also Table 3 is different, e.g. please clarify and check the NLO ttgamma cross section from FxFx. Beatriz quotes (corrected for BR) a value of 3.7 pb. Which cross section did you use for tW. Make sure it is the same for top and antitop and ideally the same as Beatriz (35.9 pb). What are the efficiencies for photon emission in the tq sample from Powheg? How do you use the cross section of 136 resp. 80.95 pb you quote in the table?

The good genphoton* definition: **Previous**

- $|pdgId| = 22$
- $status = 1$ → stable particle
- $p_T > 20$ GeV, $|\eta| < 2.5$
- *Not have any hadron or meson ancestor*
- $\Delta R(\gamma, part.) > R0gamma$ for every status=1 genparticle (not γ or ν)
  - A. Decay: 1. ℓ → γ; 2. top → W/b → γ
  - B. Production: not from decay

The good genphoton* definition: **New**

- $|pdgId| = 22$
- $status = 1$ → stable particle
- $p_T > 20$ GeV, $|\eta| < 2.5$
- *Not have any hadron or meson ancestor*
- $\Delta R(\gamma, part.) > 0.1$ for every status=1 genparticle (not γ or ν)
- $\Delta R(\gamma, part.) > 0.4$ for every status=1 gen-lepton
  - A. Decay: 1. ℓ → γ; 2. top → W/b → γ; **3. top → top → γ**
  - B. Production: not from decay

- The removal strategy and cross section value used are changed to be same as the Beatriz

  Attention: photon from ttbar → 2ℓ contains two kinds of photon including photon from decay and photon from ele mis. which needs careful handling

Comments before pre-approval

— Evan and David

https://cms-pub-talk.web.cern.ch/t/move-forward-for-documentation-freezing/35995/4

We are preparing the production of tWgamma samples at NLO QCD in the single lepton channel analogue to TOP-25-004. https://gitlab.cern.ch/cms-top-tmg/tmg-mc-requests/-/issues/50 Once ready we kindly ask the analysis team to include this process in the signal definition. A similar strategy was followed in TOP-23-004 where tWZ and ttZ are combined. This is in particular important since a main motivation for the analysis is EFT interpretation where the EFT operators usually affect tWgamma as well.

Lastly, if the authors still intend to add an EFT interpretation to the analysis down the line, we emphasize that this should at that time be presented in a tX subgroup meeting, meaning in a dedicated talk on the EFT interpretation. We do not think such a major change could be reviewed only by the ARC.

- The samples are still not ready, a plan B for this part is to add additional uncertainty.
- For the EFT study, I gave two presentation in TOP PAG and TOP ETF groups separately before. Once the SM results are approved, I'll continue the EFT study

Summing differential cross section to inclusive gives similar uncertainty, only the first digit from the inclusive fit is shown, what is the second? Which one is more precise?

You can refer to the equation 20(a) in AN v4, the inclusive expected fit uncertainty for tγq is 1.0 -0.078/+0.081. They're almost the same (±0.08 from summing the differential photon p_T results)

-The result from the differential measurement is expected to have smaller theory uncertainties and larger statistical uncertainty It might be preferable to quote this as the final inclusive cross section and uncertainty instead of the inclusive fit. However, to make the uncertainty breakdown is more complicated. To be discussed

I prefer not to sum differential as the inclusive results. We measured many differential cross sections, which variable is used for the summing is difficult to decide and as you said it's complicated to do the uncertainty breakdown, so let's keep having single fit for the inclusive results

Removing jet requirement in fiducial cross section

- **The tqgamma sample contains an on-shell top quark and is thus well defined as t-channel single top, it is well distinguished from Wgamma. Same for ttgamma. There is no jet requirement for a ttZ measurement for example. The b jet requirement was already released making it different to TOP-18-010 so in our opinion there is no reason for the jet requirement.**
- **The treatment of the out of acceptance contribution is critical since the modelling is correlated to the signal and ideally it would be fit with one or more free floating parameters. This is currently not done, and likely also not feasible due to limited sensitivity to this part. This leaves a potential systematic issue. Reducing this out of acceptance contribution can be part of the solution.**
- **For these reasons, unless we overlooked something, we think there is an advantage of removing the jet requirement from the fiducial cross section definition. Please think about it.**

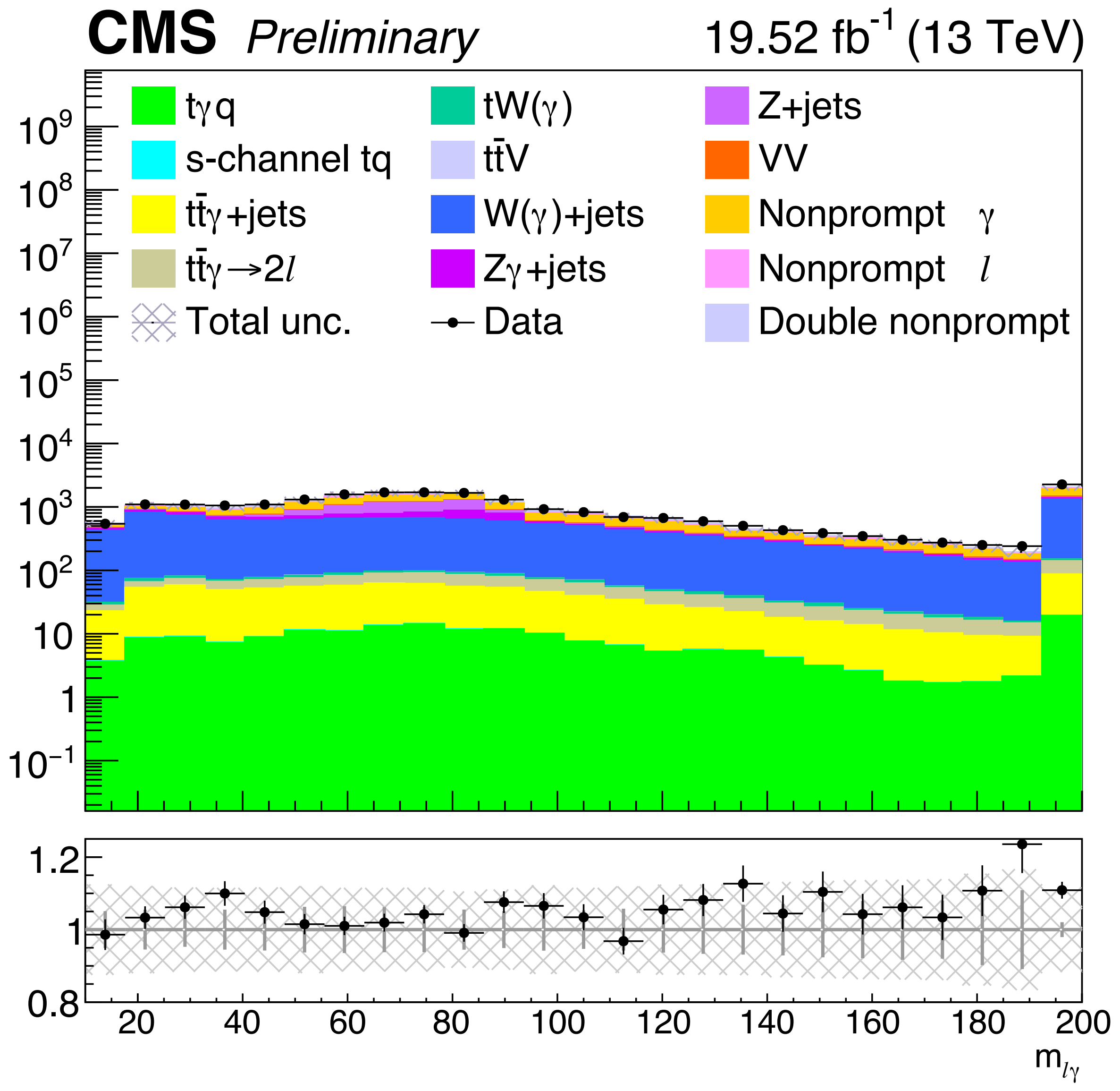
We plan to add the light jet η differential XS at particle level, which is kind of interesting in tγq process. So it's necessary to keep the jet requirement.

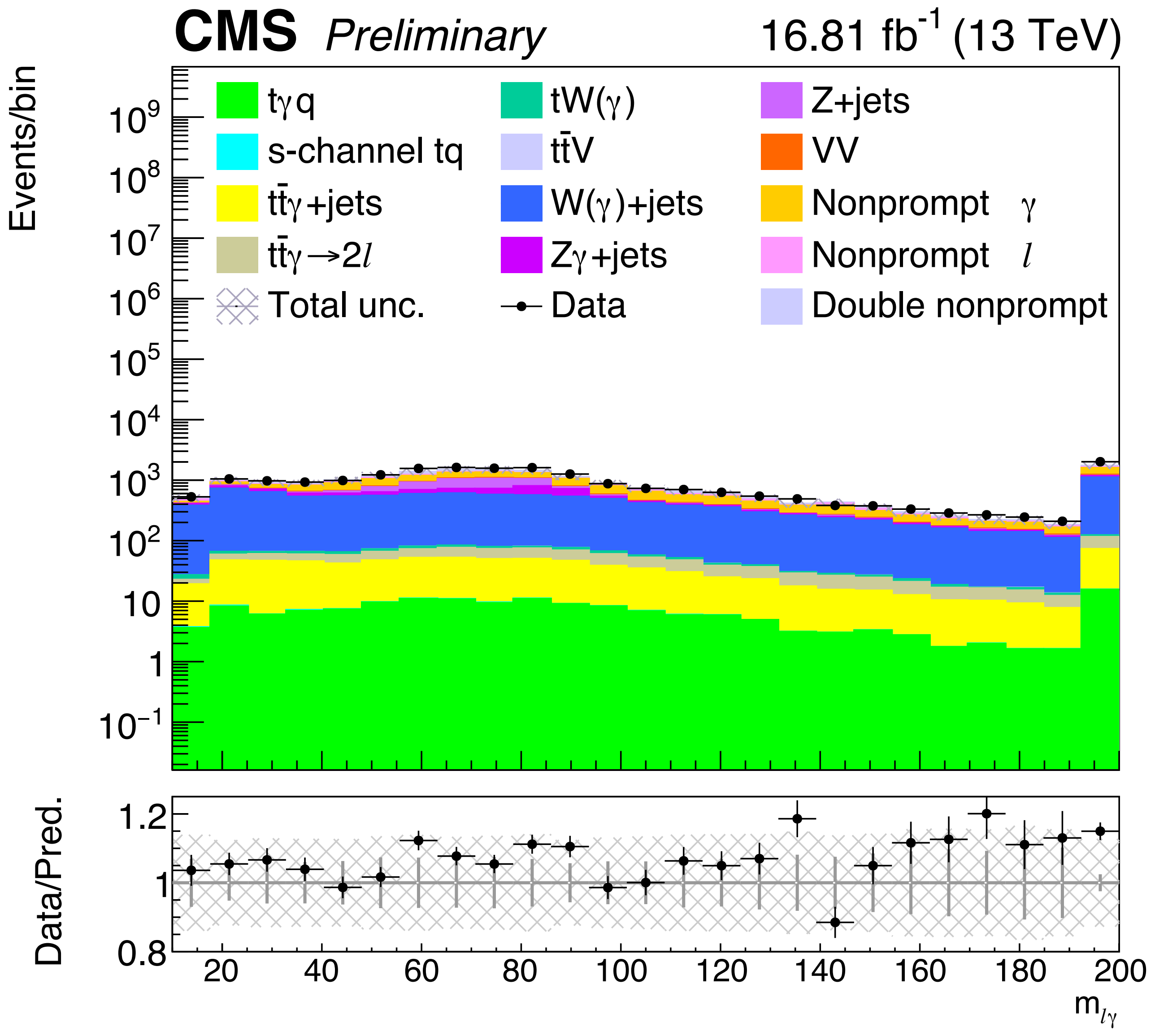
Double nonprompt closure: Did you do some kind of statistical test to check if the observed closure is not statistically significant? It looks like there are some trends visible or can at least not be excluded very well. Is there an uncertainty assigned to the double nonprompt to cover a potential bias?

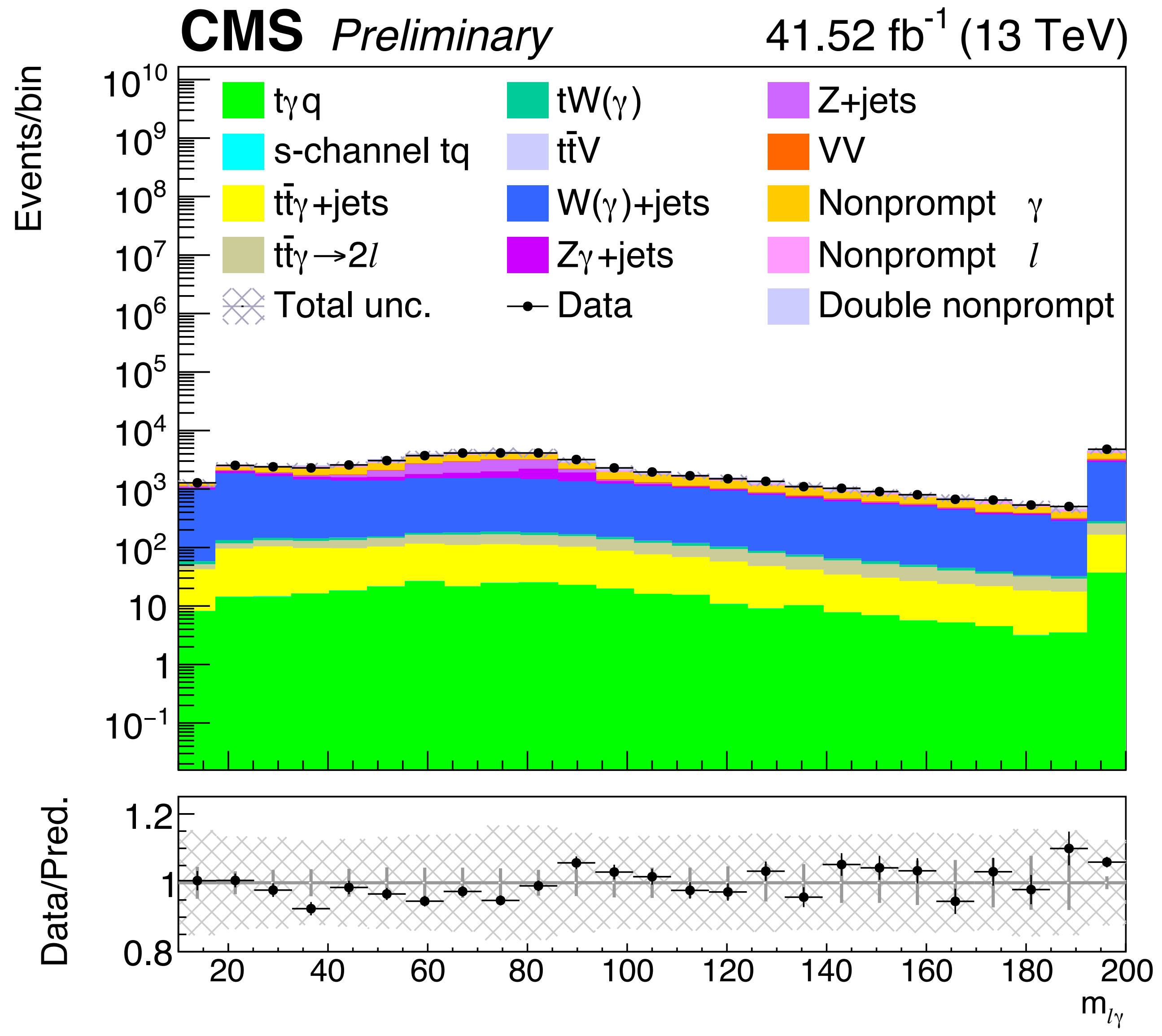
For these plots, only MC statistical uncertainty is considered.
The Kolmogorov–Smirnov (K-S) Test is performed for the comparisons between a.4 (b.4) and c.1, a.3 and c.3, and b.3 and c.4 in distributions of photon pT, except for the agreement between b.4 and c.1 with a p-value 0.02, the other comparisons have p-values larger than 0.05.

Fit: thank you for providing the inputs to the fit by era per process, but we would like to see the full “stacked” data/mc comparison plots per region that is treated separately in the fit. This will show if the agreement is similar per year. Already this is interesting to see in the CR, and will become important to see in the SR after unblinding.

The agreement per year in the muon channel has a similar trend. Looks reasonable within uncertainties.







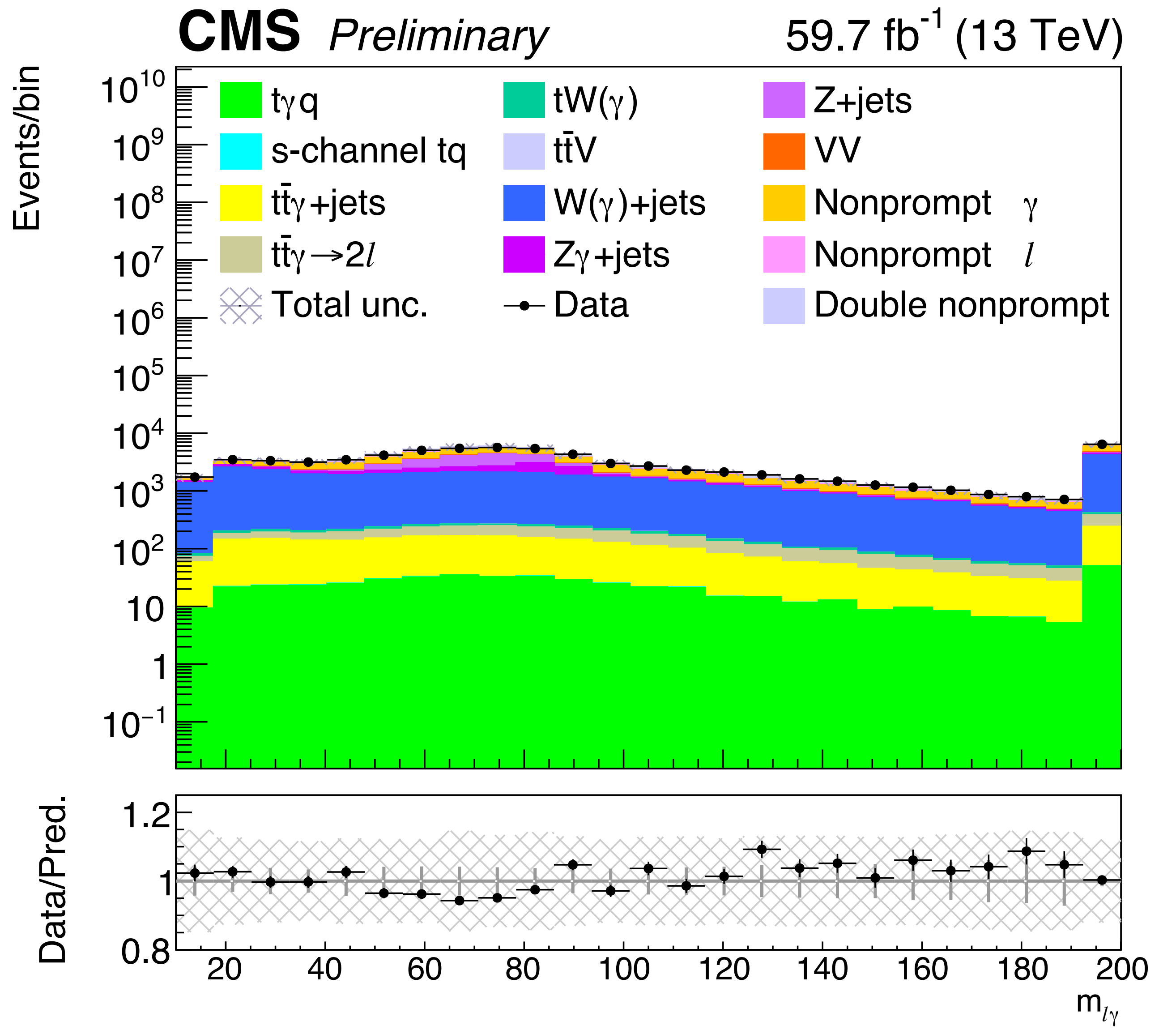


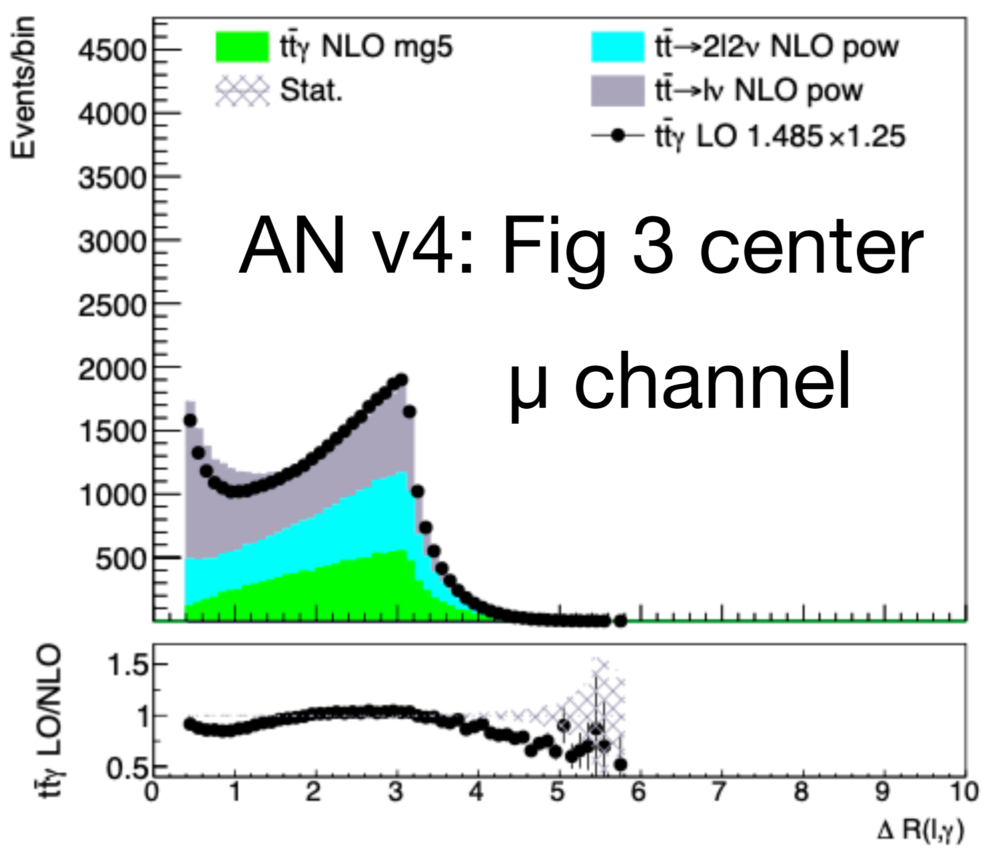
Figure 3/4:

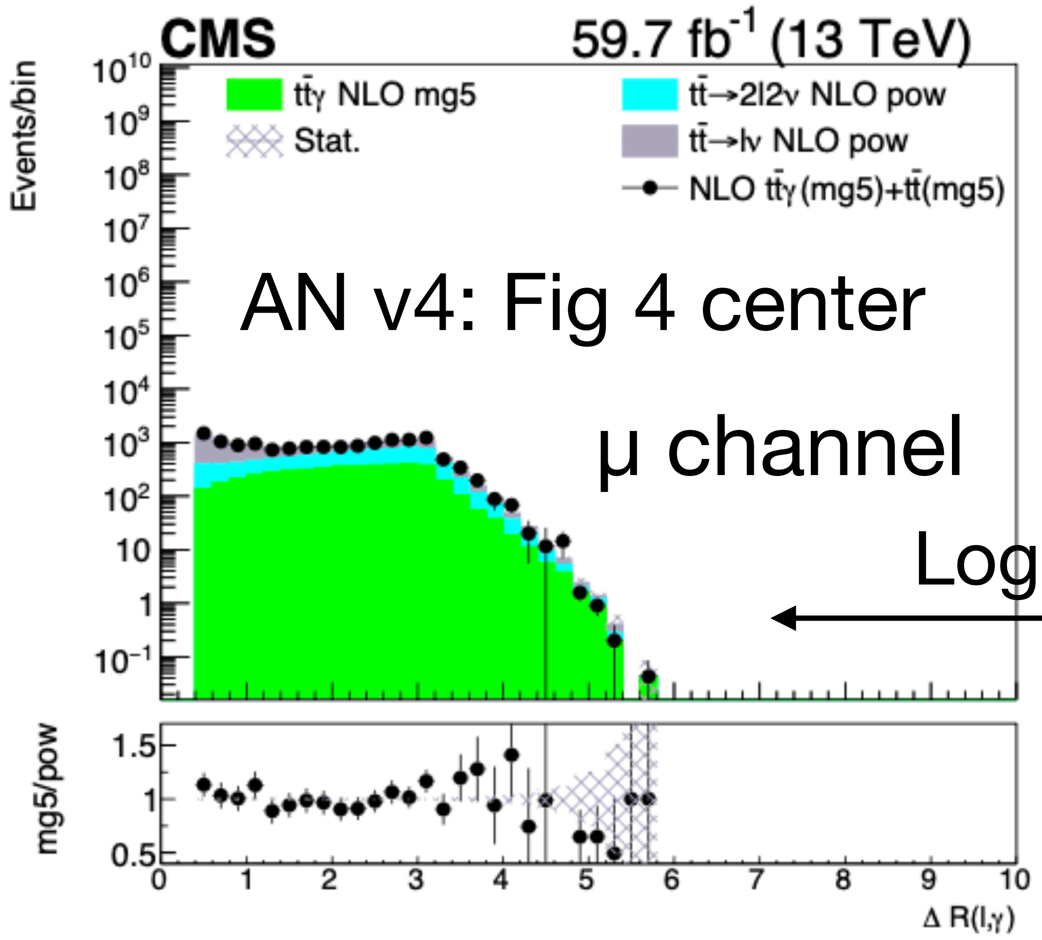
- One difference in these two figures are the data points, but what is the difference on the filled histograms?

- Is the grey contribution in Figure 3 center much larger compared to the lower plot, and if so why? They seem to have different style axes.

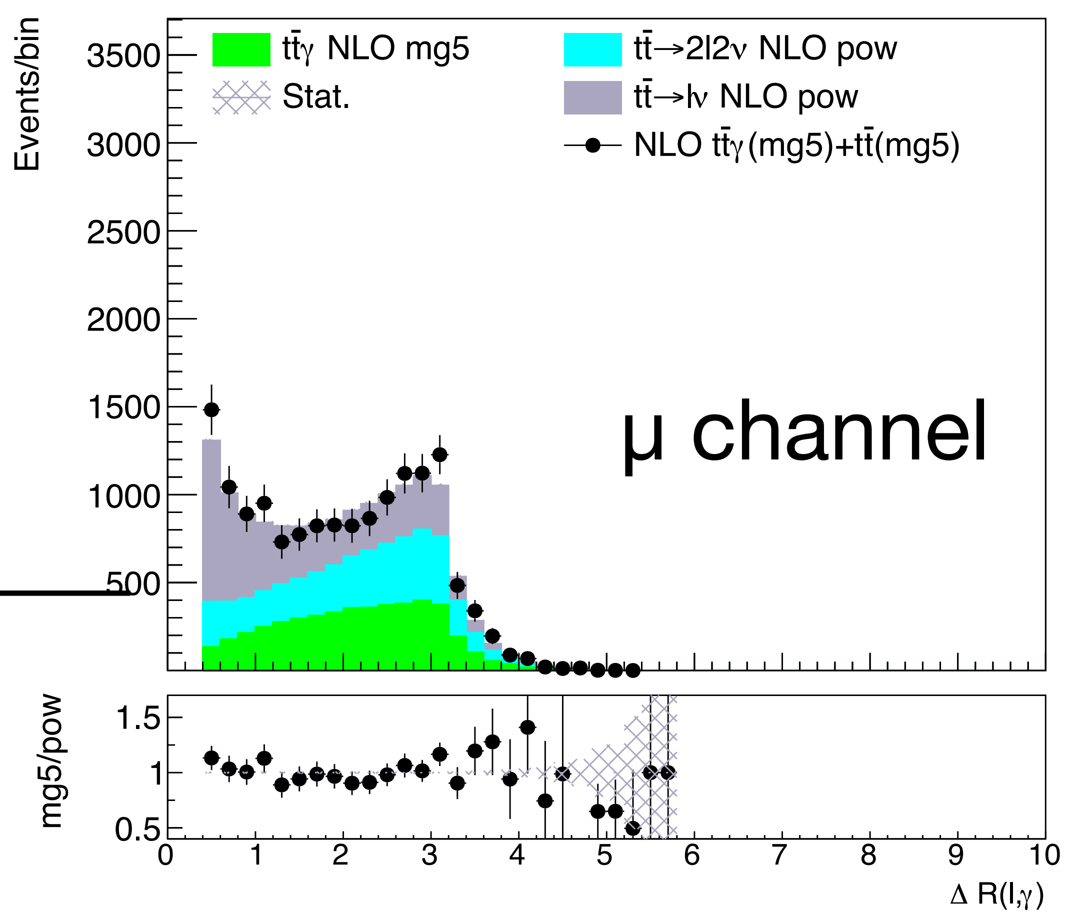
- Did you check the corresponding plots for electrons? How does it look like?

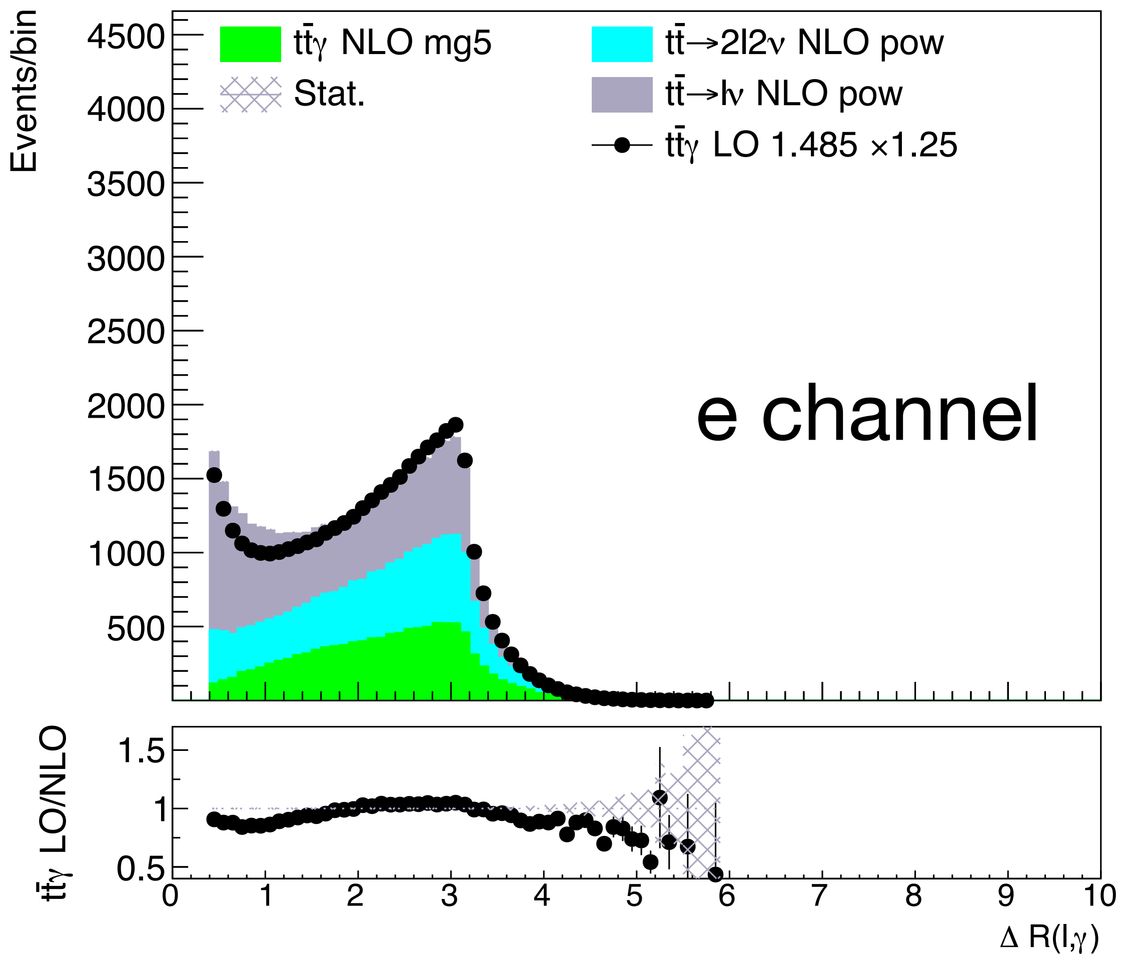
- ☐ Figure 3 is plot in fiducial region at particle level and Figure 4 is plot in SR at reconstruction level
- ☐ They are in different selection level that one is particle level and the other is reconstruction level. In reconstruction level,
- ☐ The performance in electron channel is similar to the muon channel





Log scale





Even with the sample with mll>10GeV there is a part of the phase space uncovered with mll<10GeV. Events from this phase space can enter your selection (in particular b veto CR) if one of the two leptons escapes acceptance/identification. Which could be enhanced in Zgamma as the photon can take away energy from the final state lepton. To ensure that it is not a problem, please provide a plot with Zgamma events selected in the b-veto CR as a function of mll. Is the contribution rising or falling when going towards mll of 10?

Events selected don't contain a second lepton, not possible to provide the Zγ events in CR in distribution of mll.

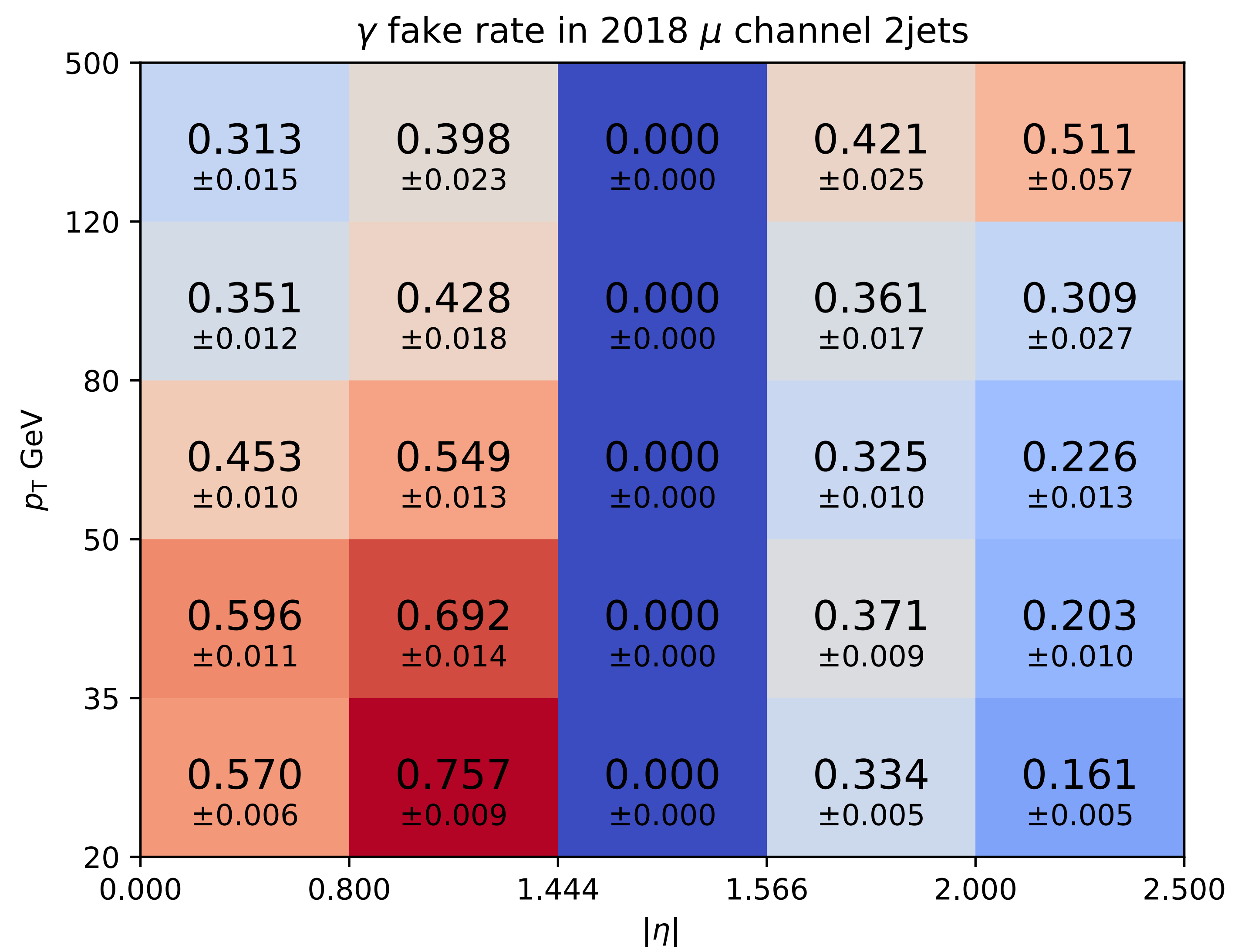
Moreover, our $m_{l\gamma}$ is required to be more than 10 GeV which was not mentioned in previous version of note, which can avoid this risk

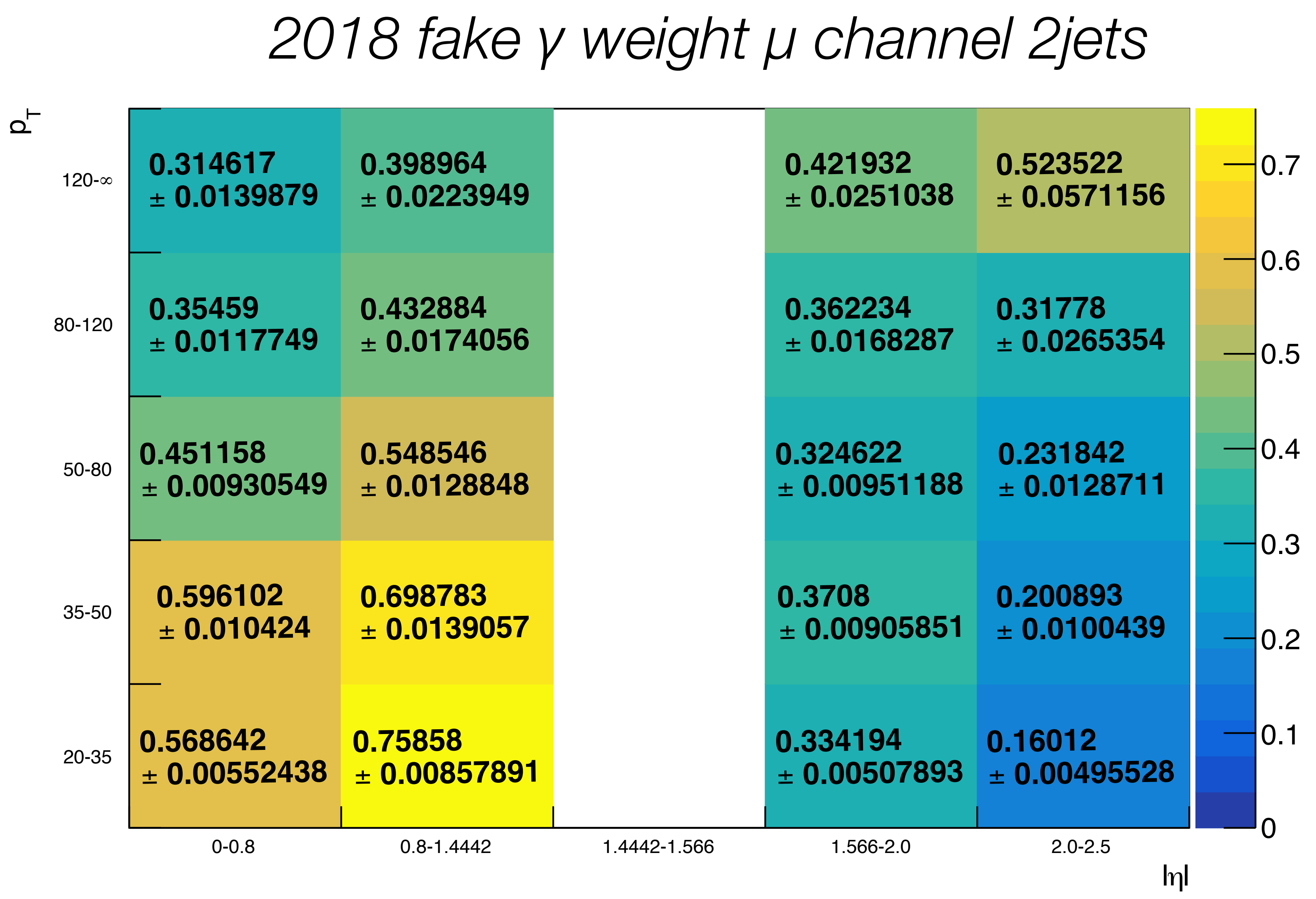
How many events are removed from the tqgamma and ttgamma samples where the photon is produced in production? I.e. those that don't have a good generator-photon. In figure 5 it looks like none are removed.

The requirements of the overlap removal is designed to keep photon production from Xγ sample as much as possible, and the contribution from X+jets sample after removal only includes the contribution from the phase space not covered by the Xγ sample. So almost more than 90% of Xγ events can be kept after the overlap removal in our signal region. For example, for 2018 muon channel, the tγq yields w/ and w/o removal are 1290 and 1305.

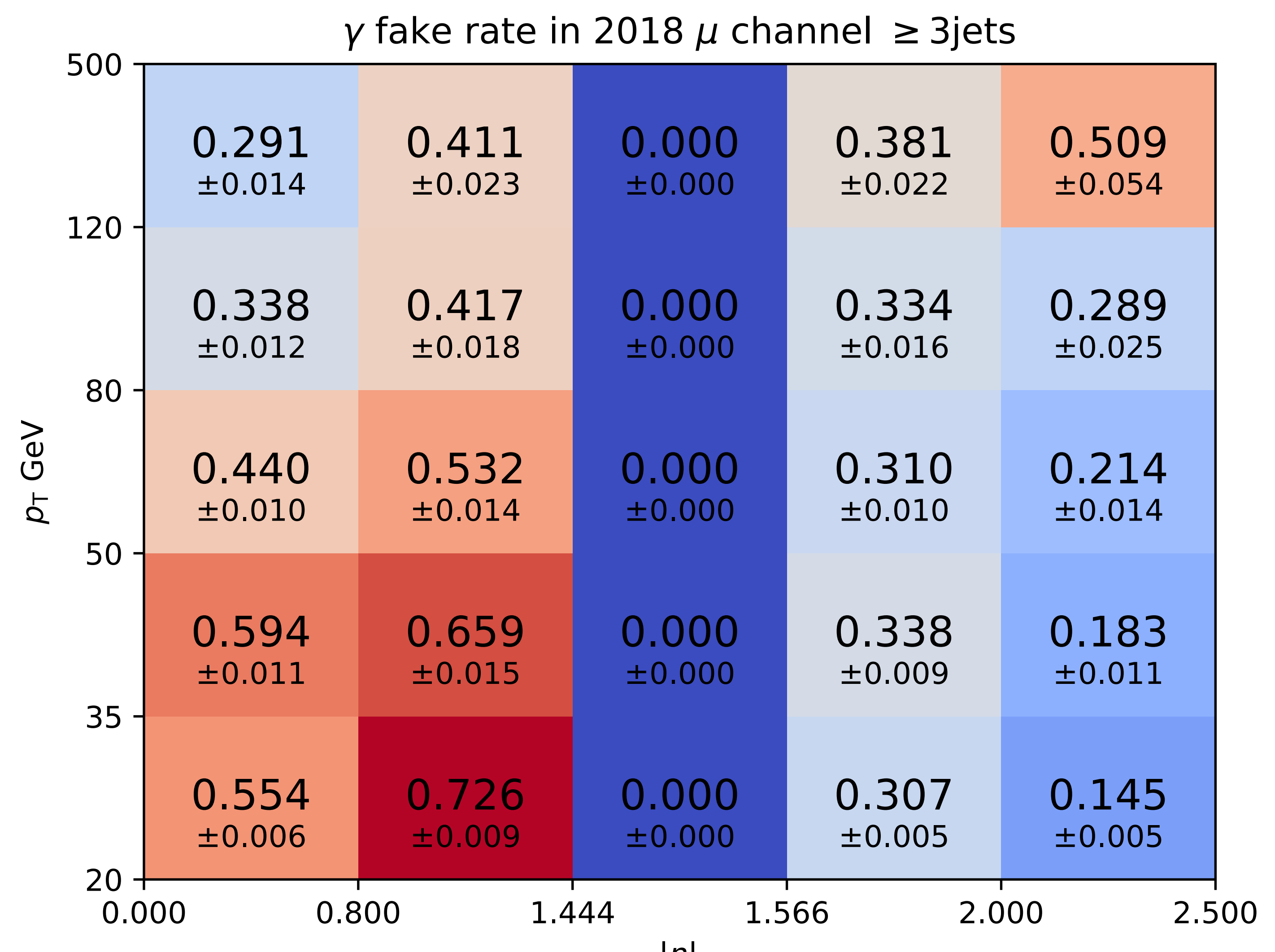
Is it possible to reduce the pT cut on the lepton veto ID? In particular to reduce the Zgamma events where 1 lepton escapes identification, e.g. to 10GeV. Also because Zgamma events with mll<10GeV are neglected.

Implemented. This cut should sync to nonprompt fake rate calculation, too. After some tests and checks, we found that the fake rate changing is small and within uncertainties. So we don’t re-calculate the nonprompt fake rate after changing this cut





- Decrease the p_T for second lepton rejection from 15 to 10 GeV has negligible for fake weights in data-driven estimation because the weight is a ratio, the change works in both numerator and denominator.



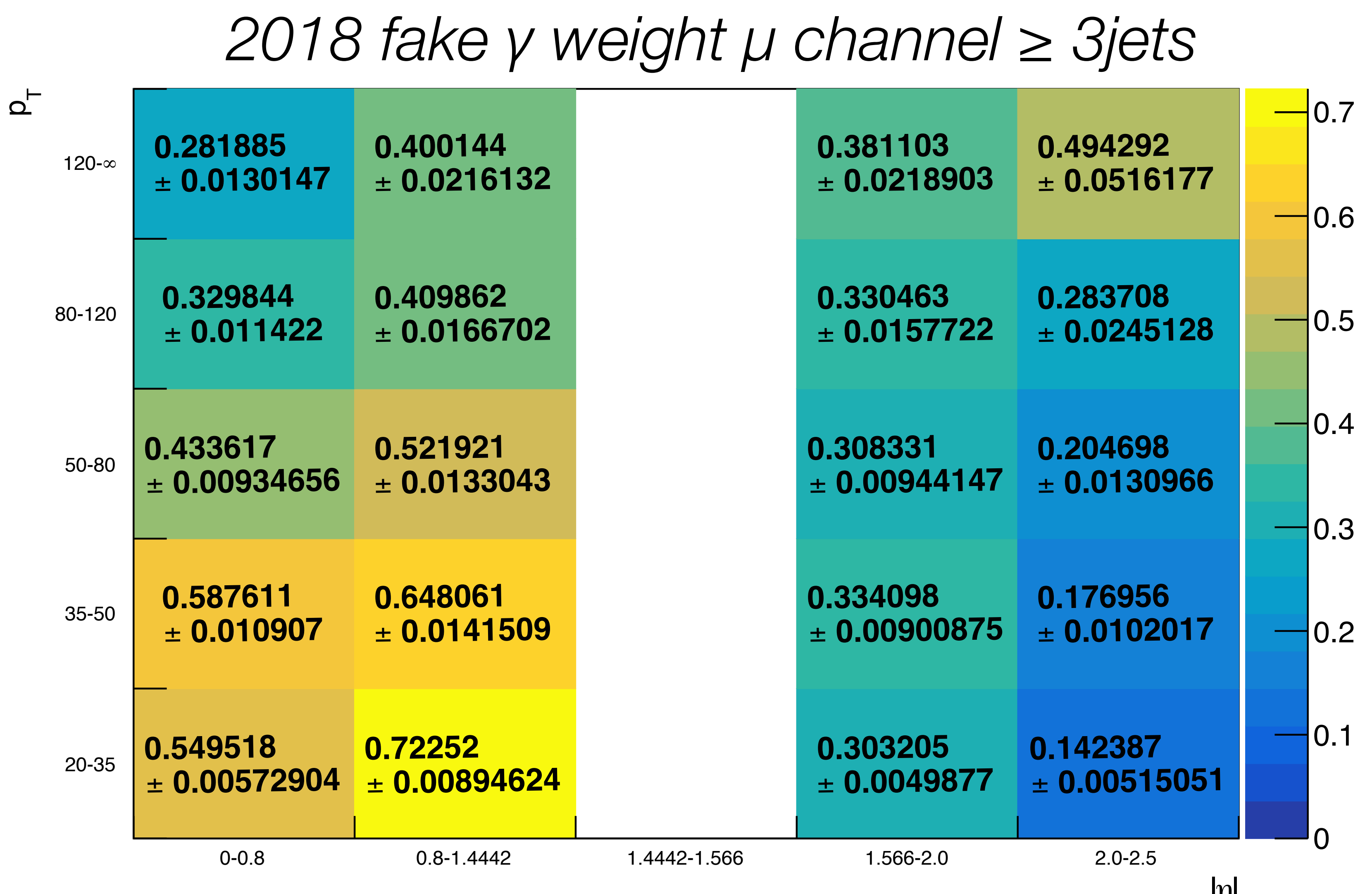
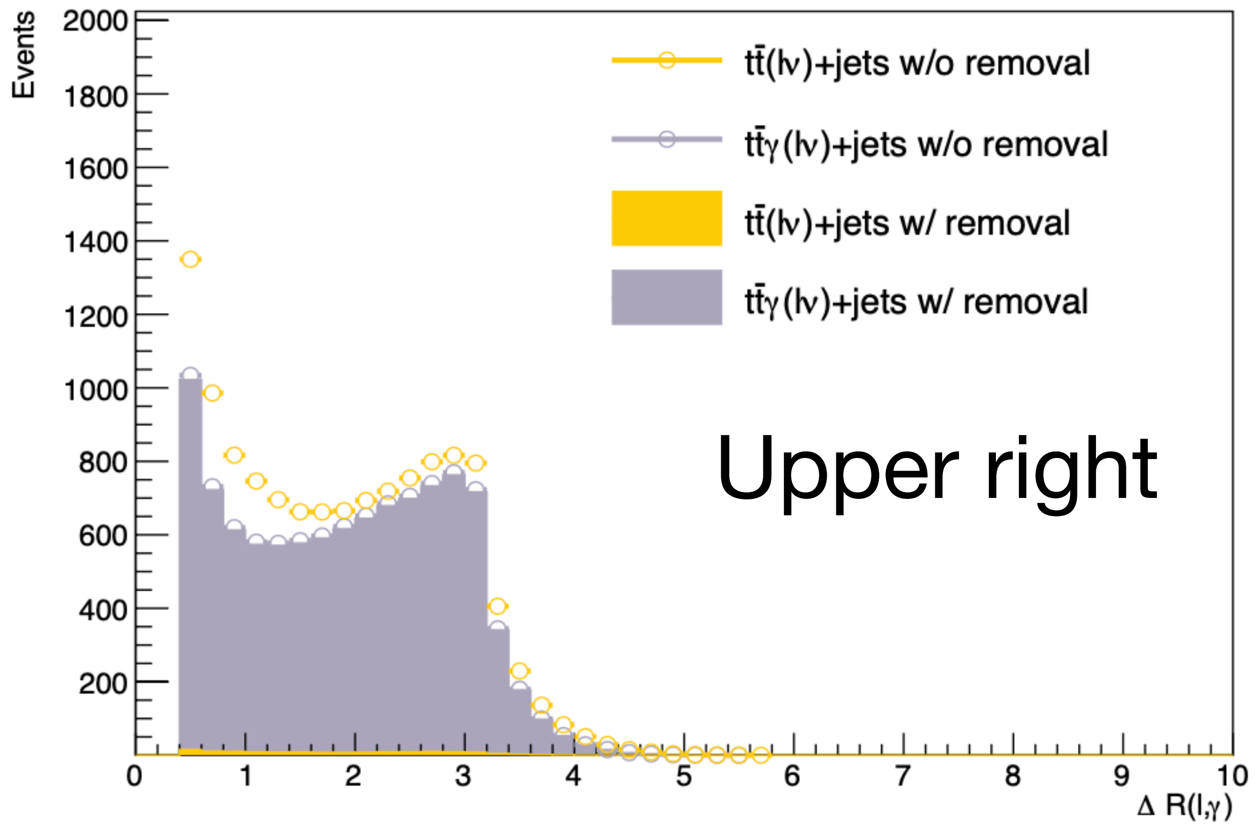
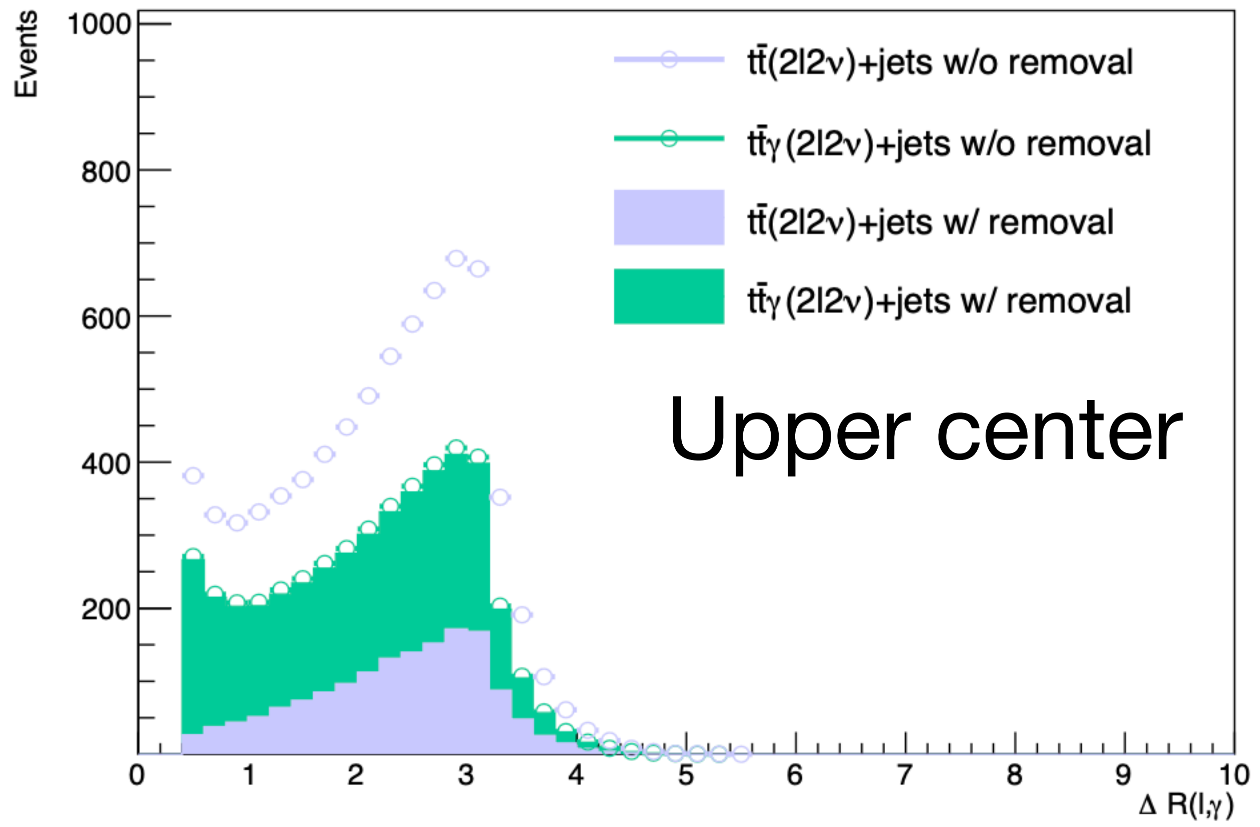


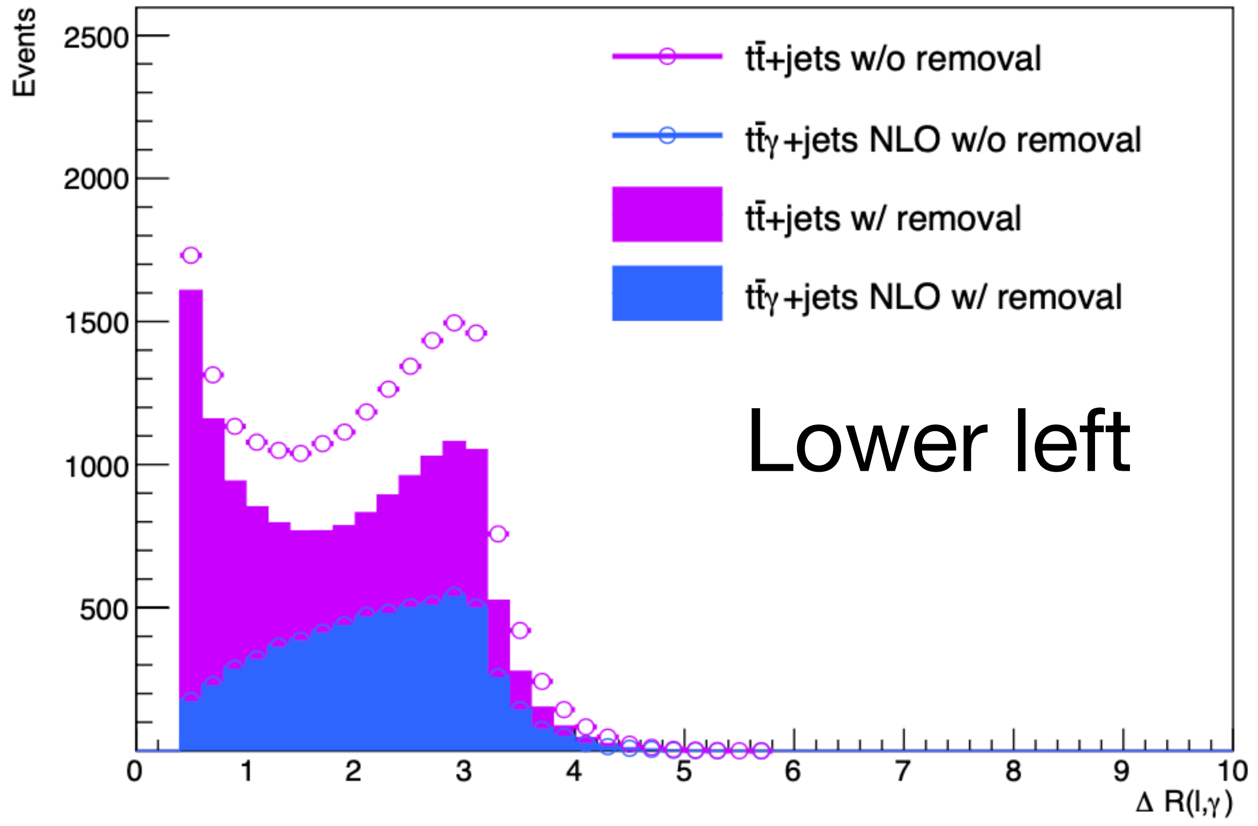
Figure 5:

- **What is the difference between the upper right and lower left?**

The upper center and right plots are with tt+jets from POWHEG in dilepton and single lepton channels.
The lower right plot is from the inclusive MG5 tt+jets sample.

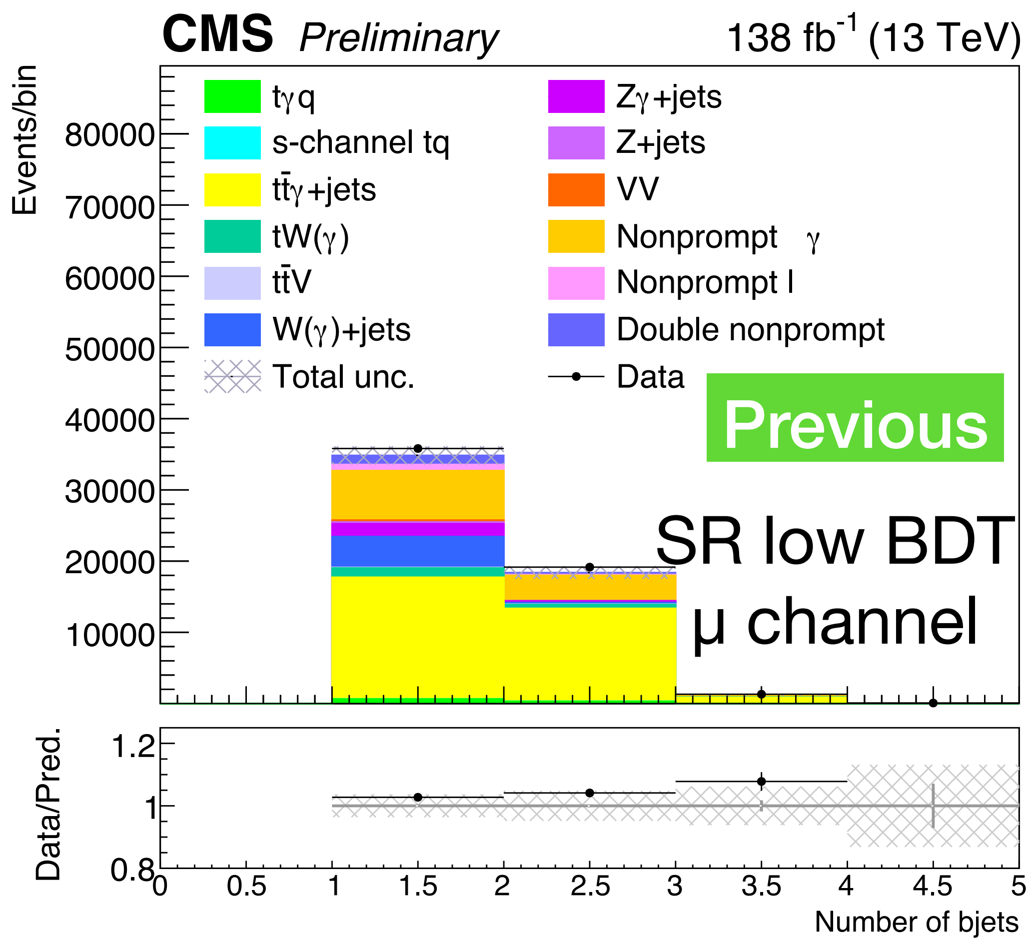


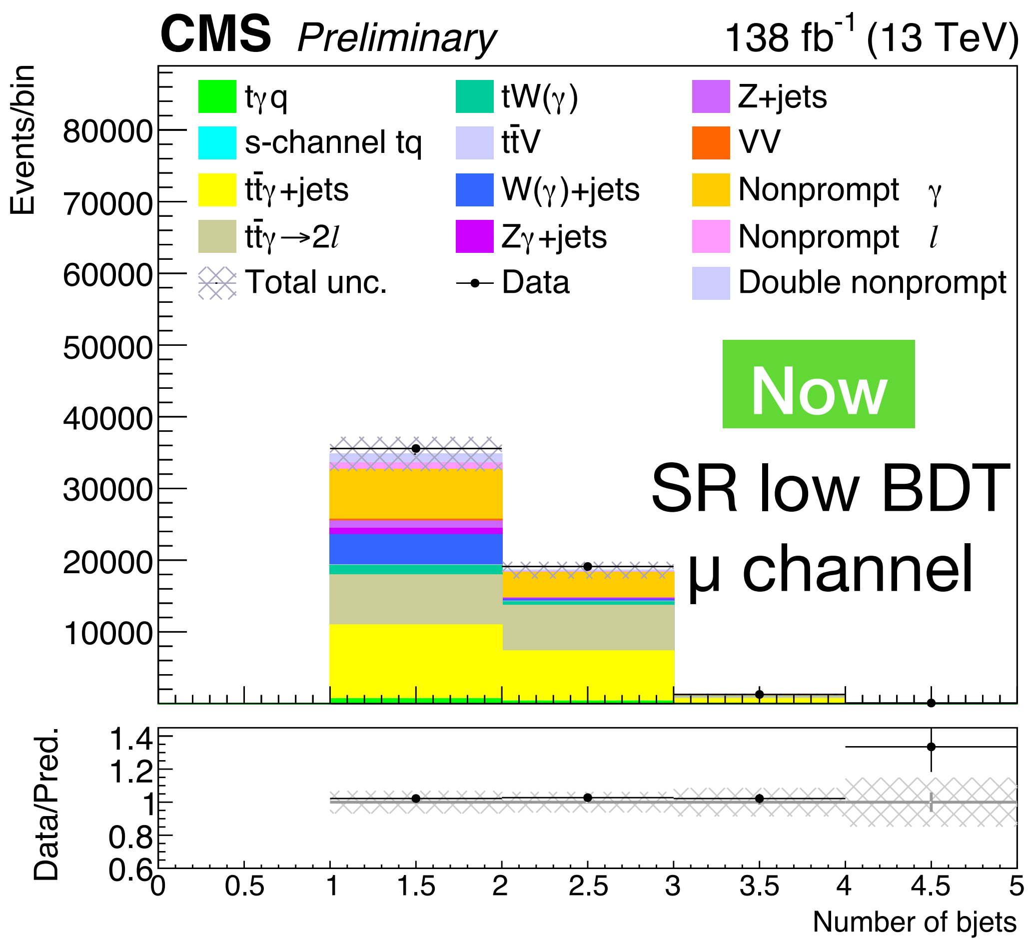




For ttgamma would it be possible to check the top quark decay mode instead?

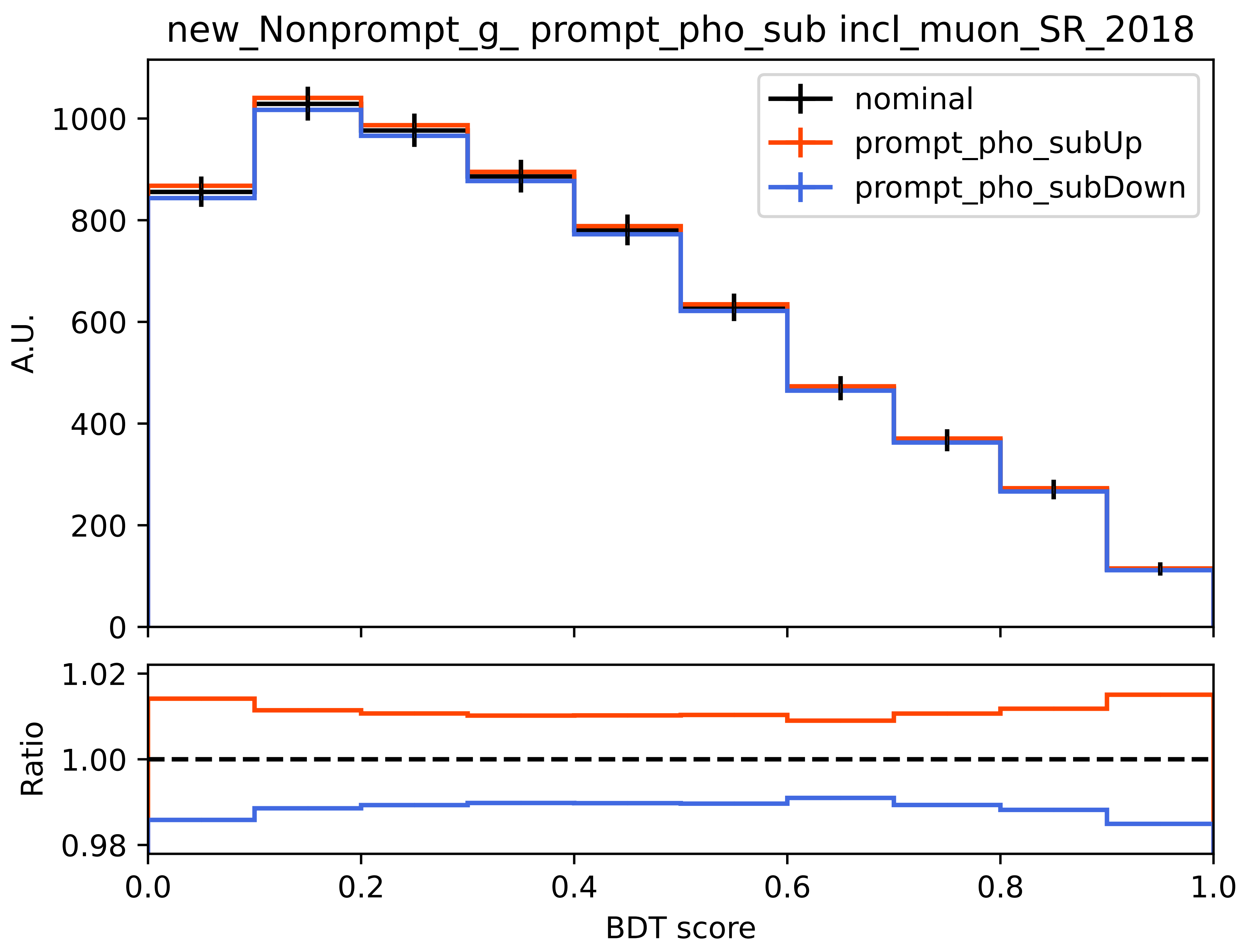
- We previously used the inclusive ttg as the signal sample, so it includes the ttg →2ℓ contribution
- Now, the inclusive ttg sample is changed to ttg →1ℓ and ttg →2ℓ samples where the former is as the signal and the latter is as the background

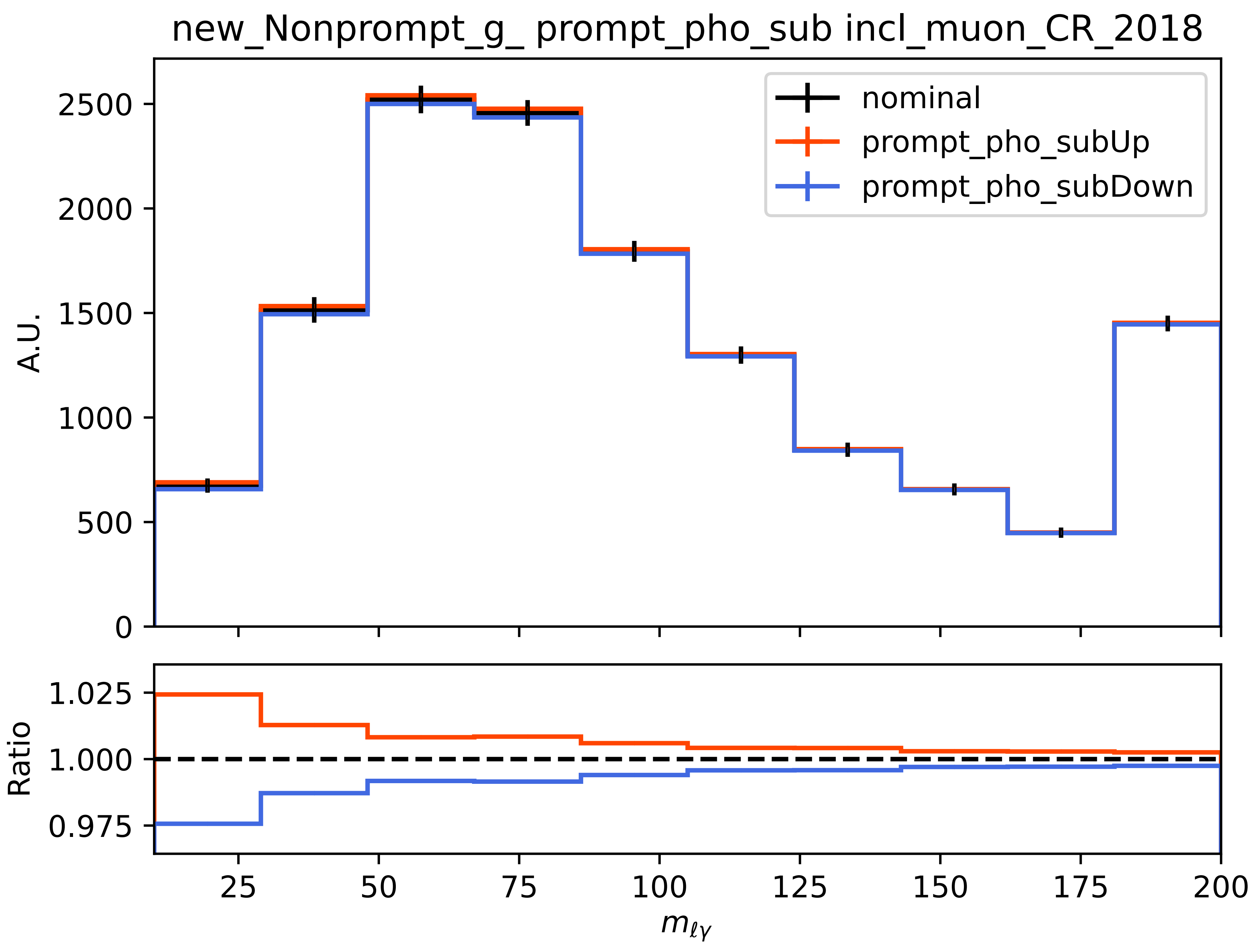




A (shape) uncertainty on the residual prompt MC should be assigned, e.g. by scaling the prompt MC up and down by 20% and repeating the prediction. Same for nonprompt lepton.

The variation from the prompt MC subtraction is from the MC cross section and the fake weights calculated.
When we change the fake weights up or down, it also propagates to the prompt MC subtraction and enters into the final nonprompt estimation. So if we want to consider the prompt MC subtraction from MC cross section, I think 20% is too large, 10% might be a proper value.

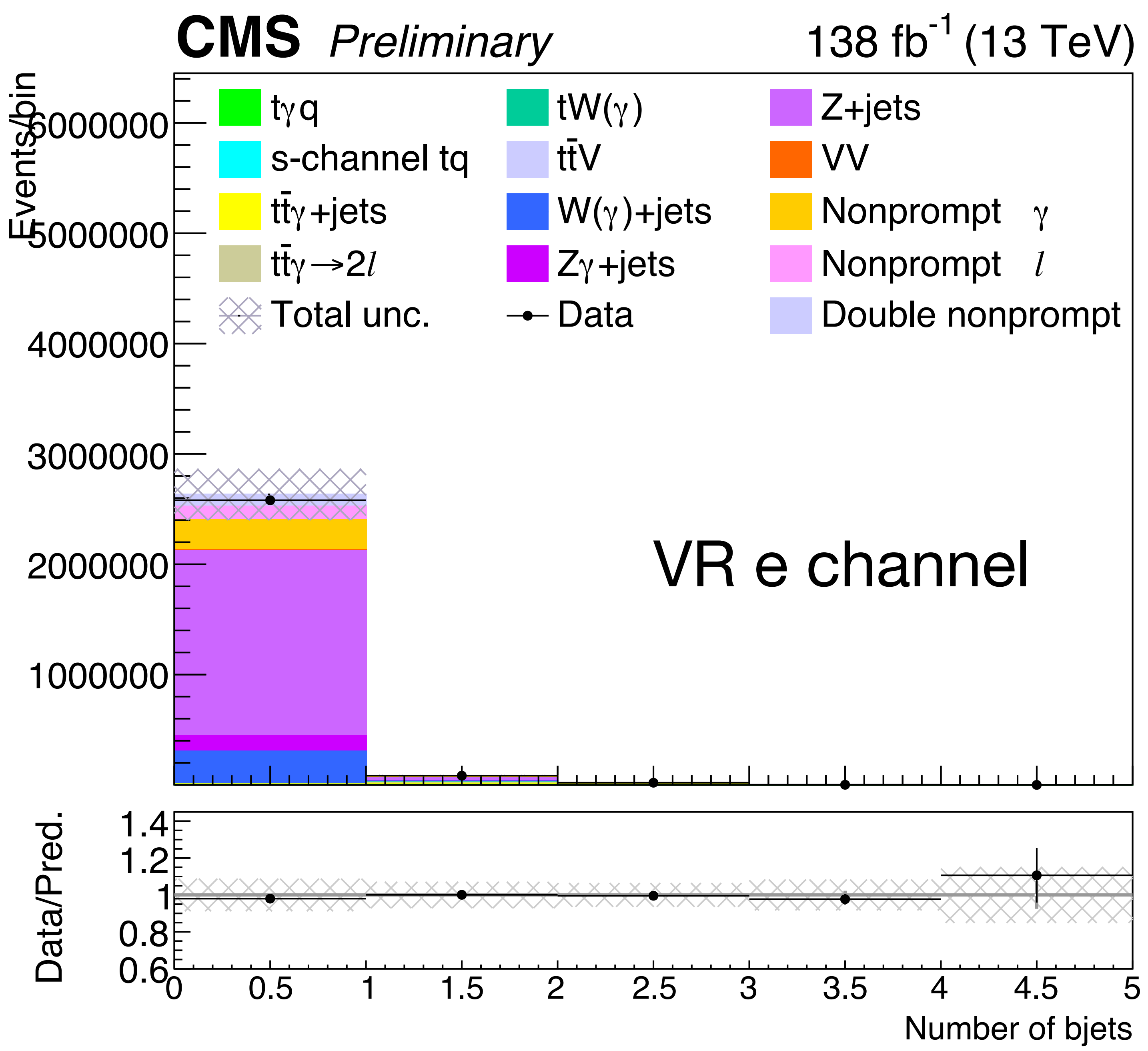


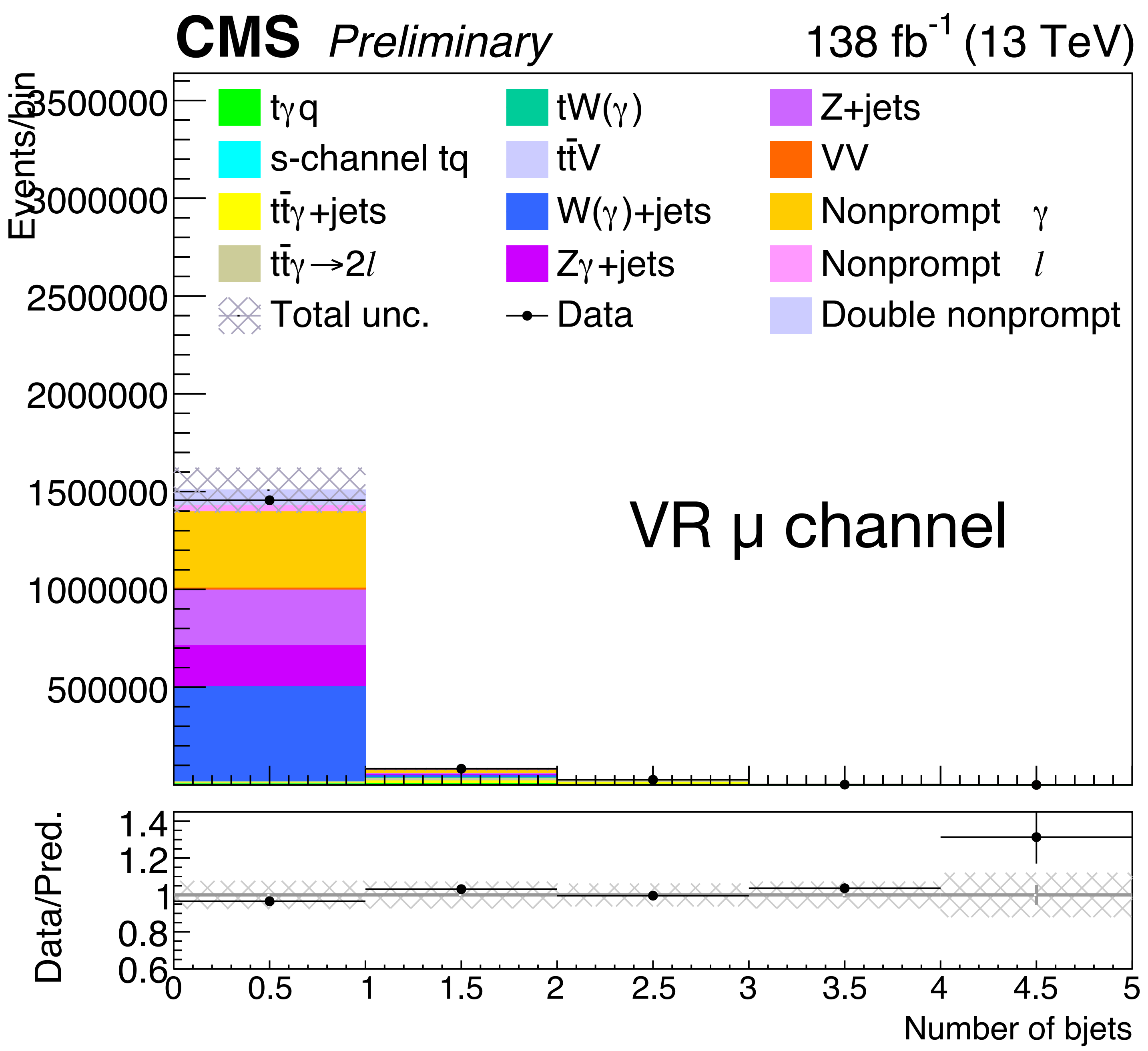


Given that the final fit is performed in b-jet veto CR and SR and has different b jet multiplicity, is it possible to extract the nonprompt photon also in 2 bins of b jet multiplicity? Or can you show the predictions are consistent between no b jet and b jet regions?

For data/MC agreement, we have the bjet multiplicity distribution in muon and electron channel where the Z+jets contribution in electron channel is scaled with a constant factor. The agreement looks good.

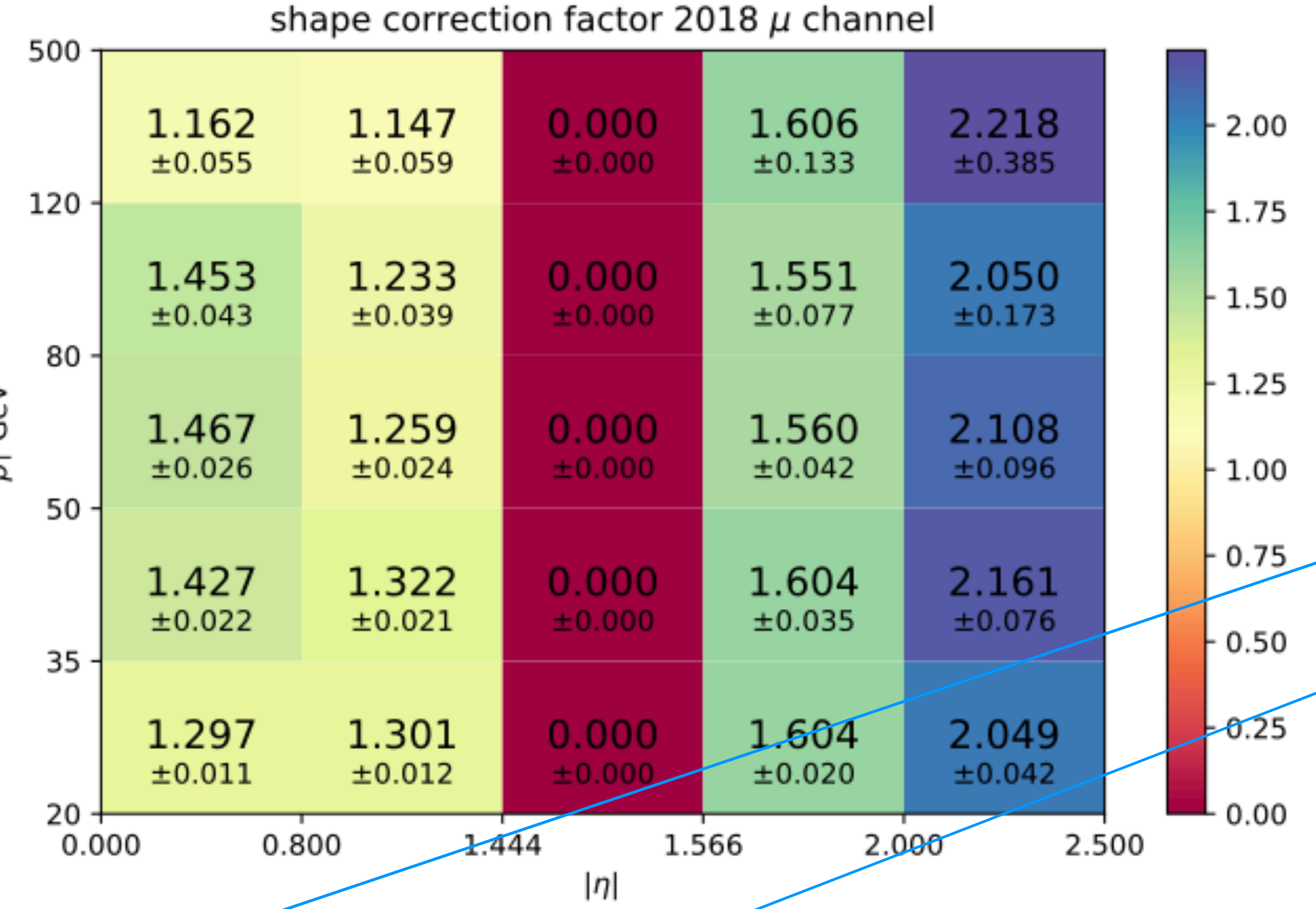
It's possible to get the nonprompt photon weights in bins of b jet multiplicity that calculate new fake rates as functions of photon pt, eta and number of bjets instead of number of jets used in current fake weight. But considering the good agreement in bjet multiplicity distribution and we already consider many NPs for nonprompt, the uncertainty for nonprompt photon is conservative enough.

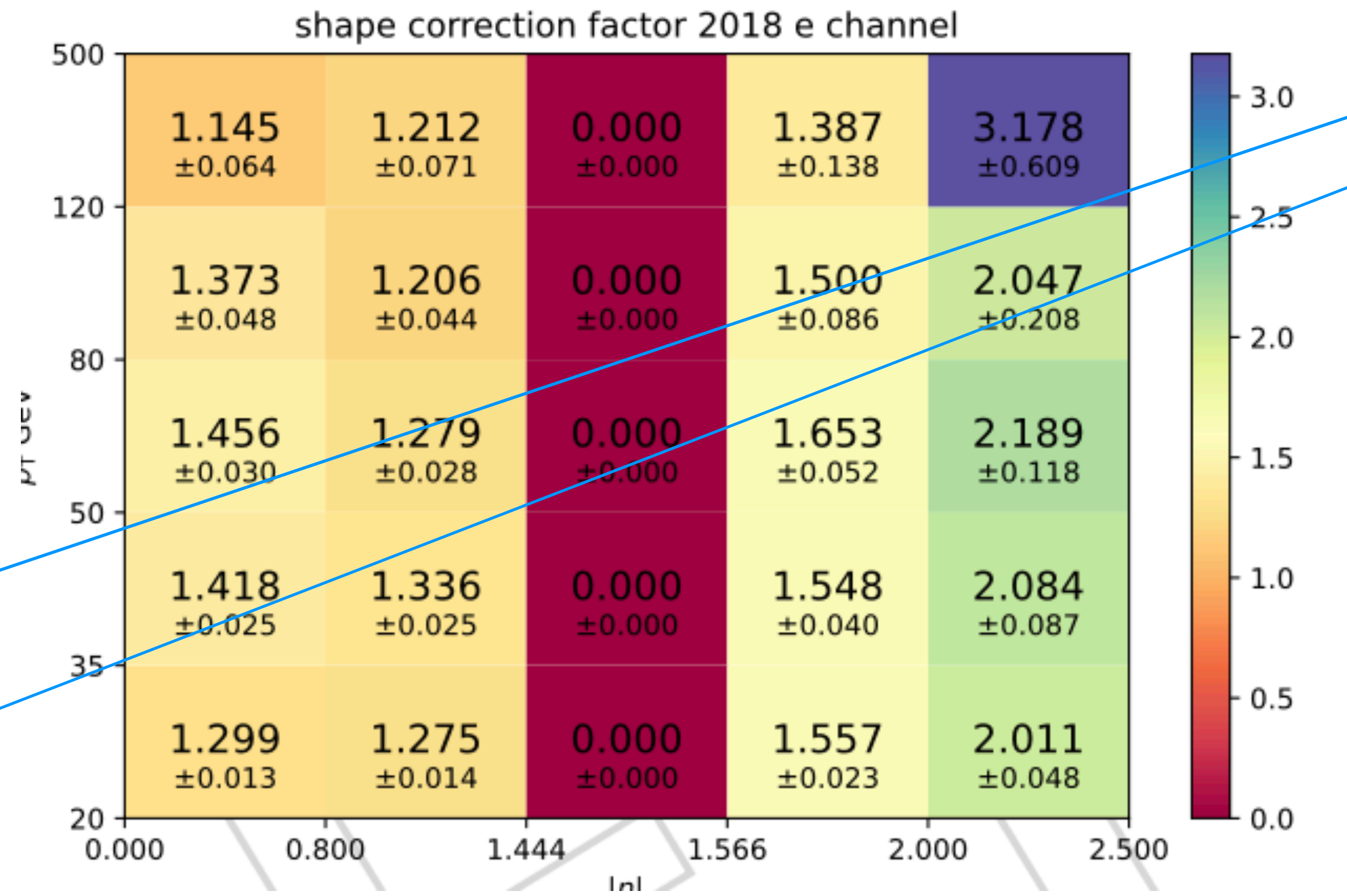




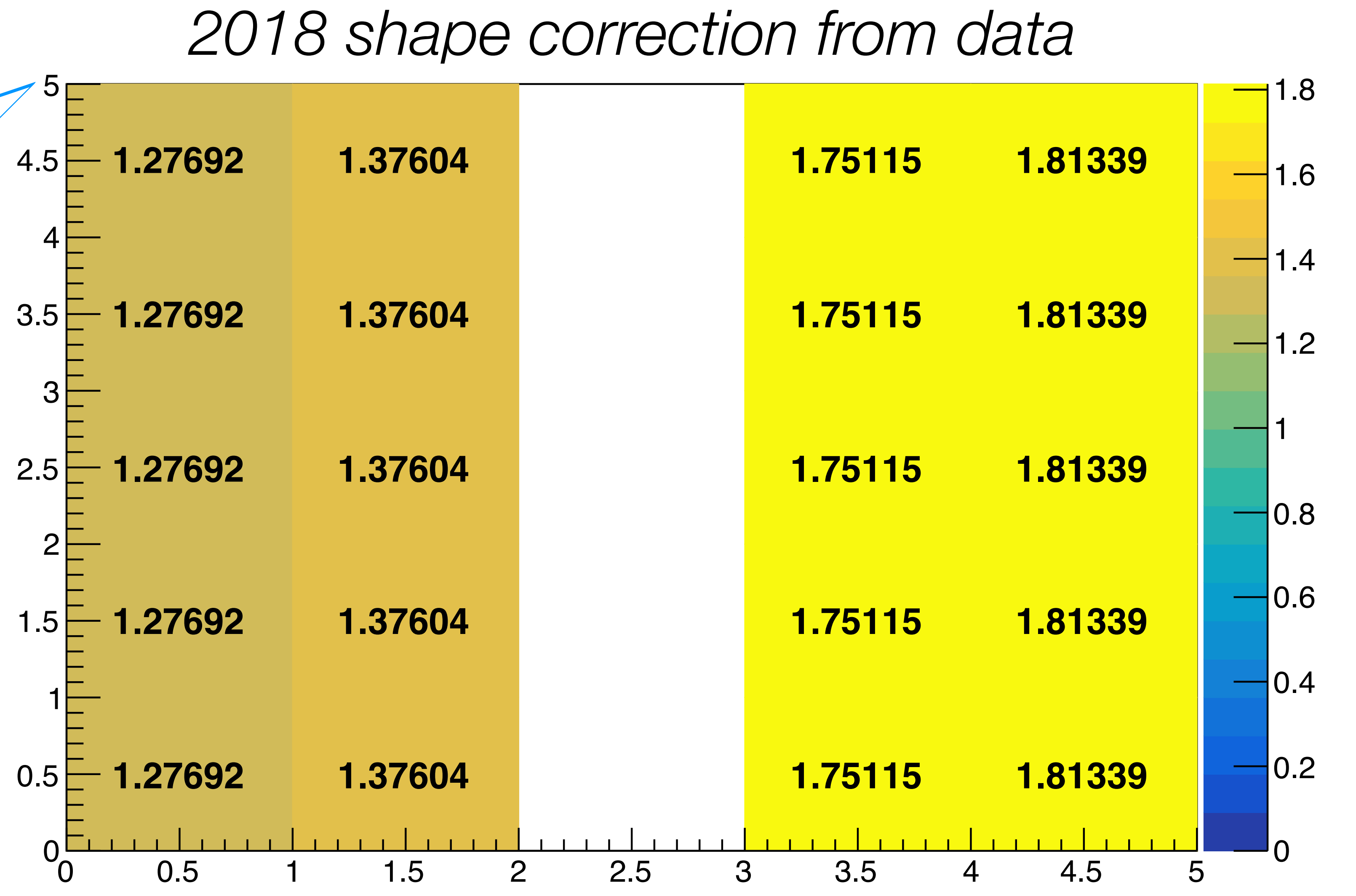
Is there an uncertainty assigned to the MC correction in equation 9 as data can be different from MC?

These kMC factors are derived by using nonprompt photons from MC. For data, such factors in fact can be calculated by doing prompt MC subtraction from data in ABCD regions. respectively. But we'll have two difficulties including **limited statistics** in data and the **electron mis photon in electron channel**. So we can only have these factors in muon channel with limited bins. That make us use the factors calculated from data muon channel to calculate uncertainty for both muon and electron channel. Moreover, the factors calculated in MC also have non-negligible statistical uncertainty, combined what I got as follows, I think we can consider 10% (barrel photon) and 20% (endcap photon) uncertainty for these k-factors treated as uncorrelated across years, but correlated across lepton and jet flavour channels





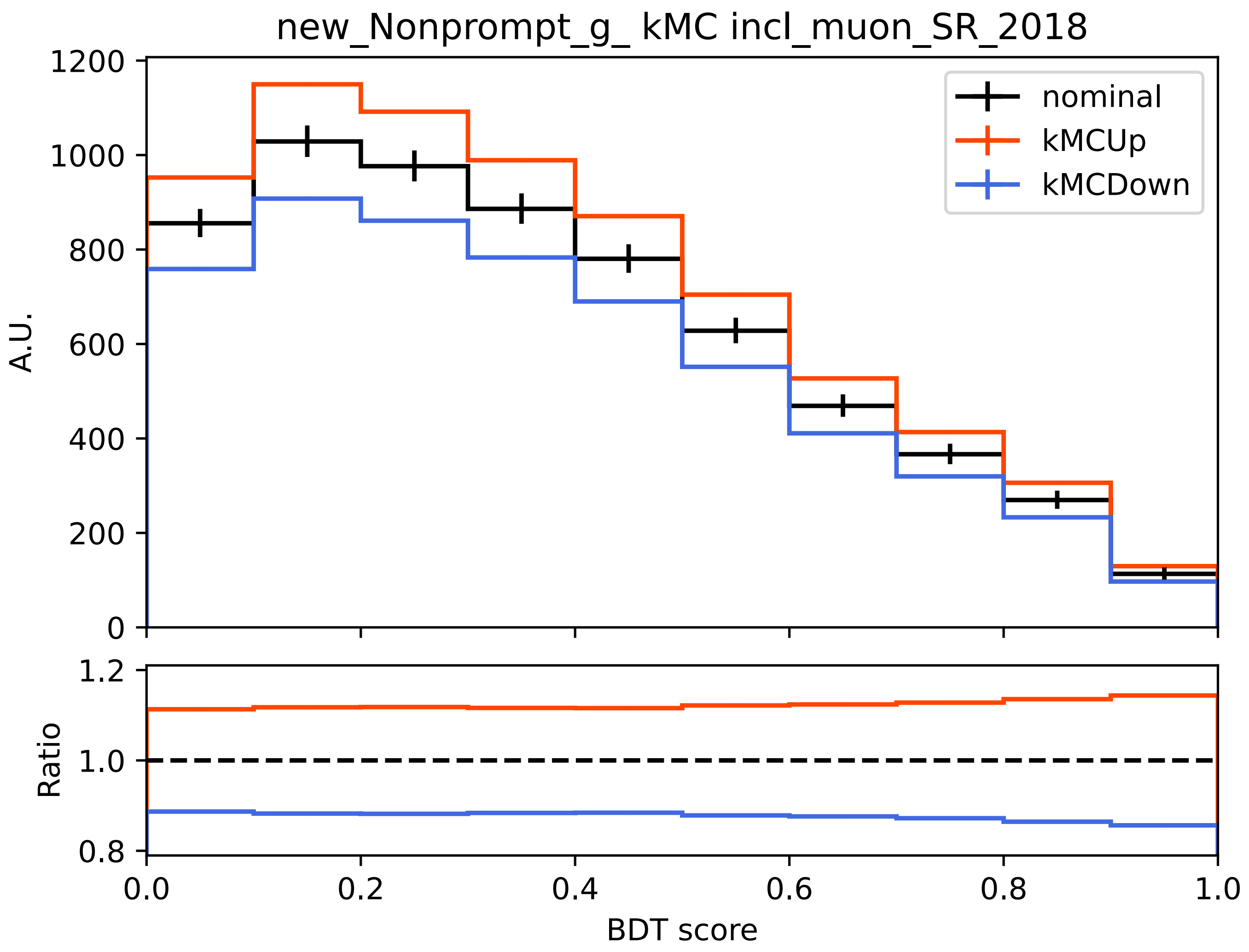
from Data

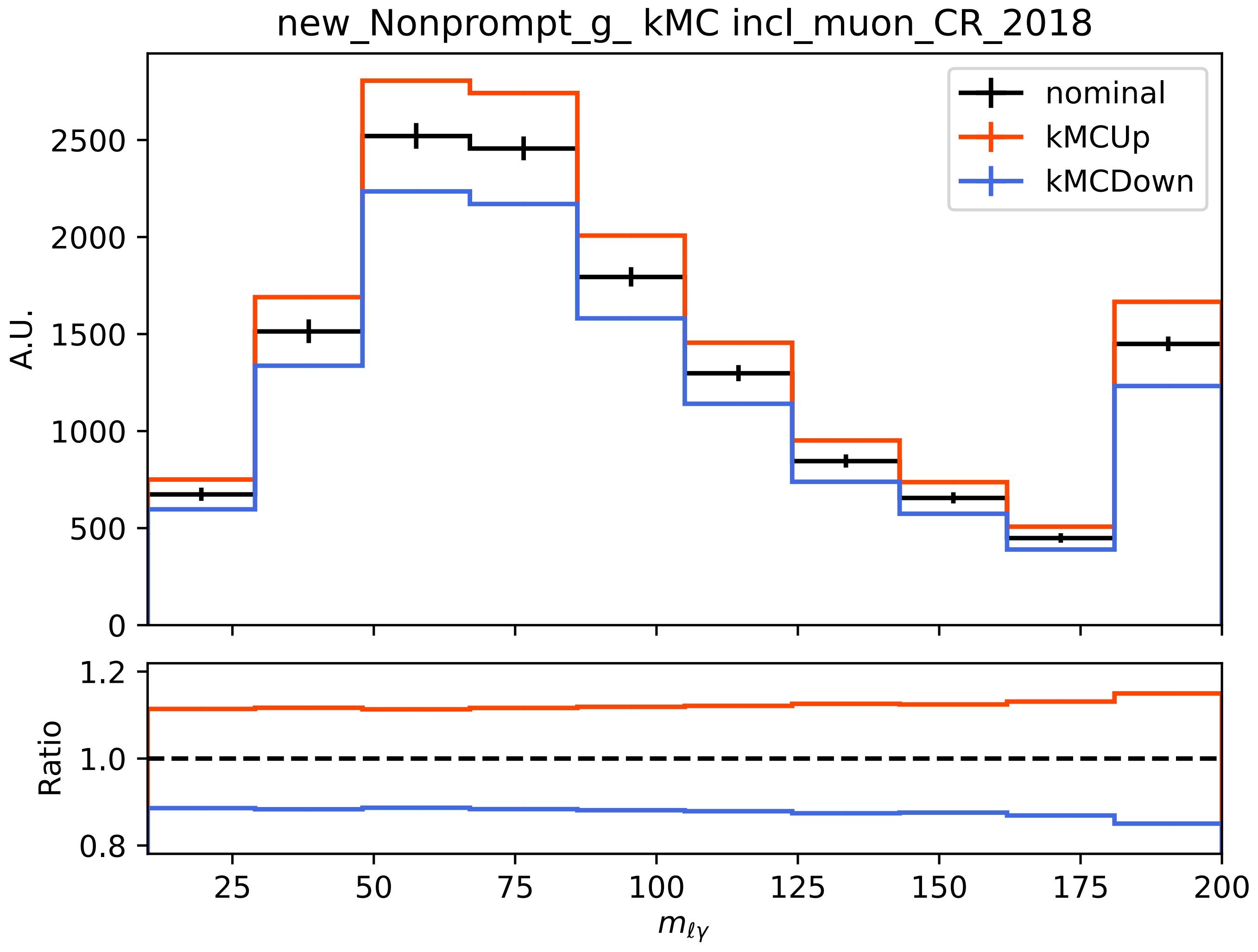


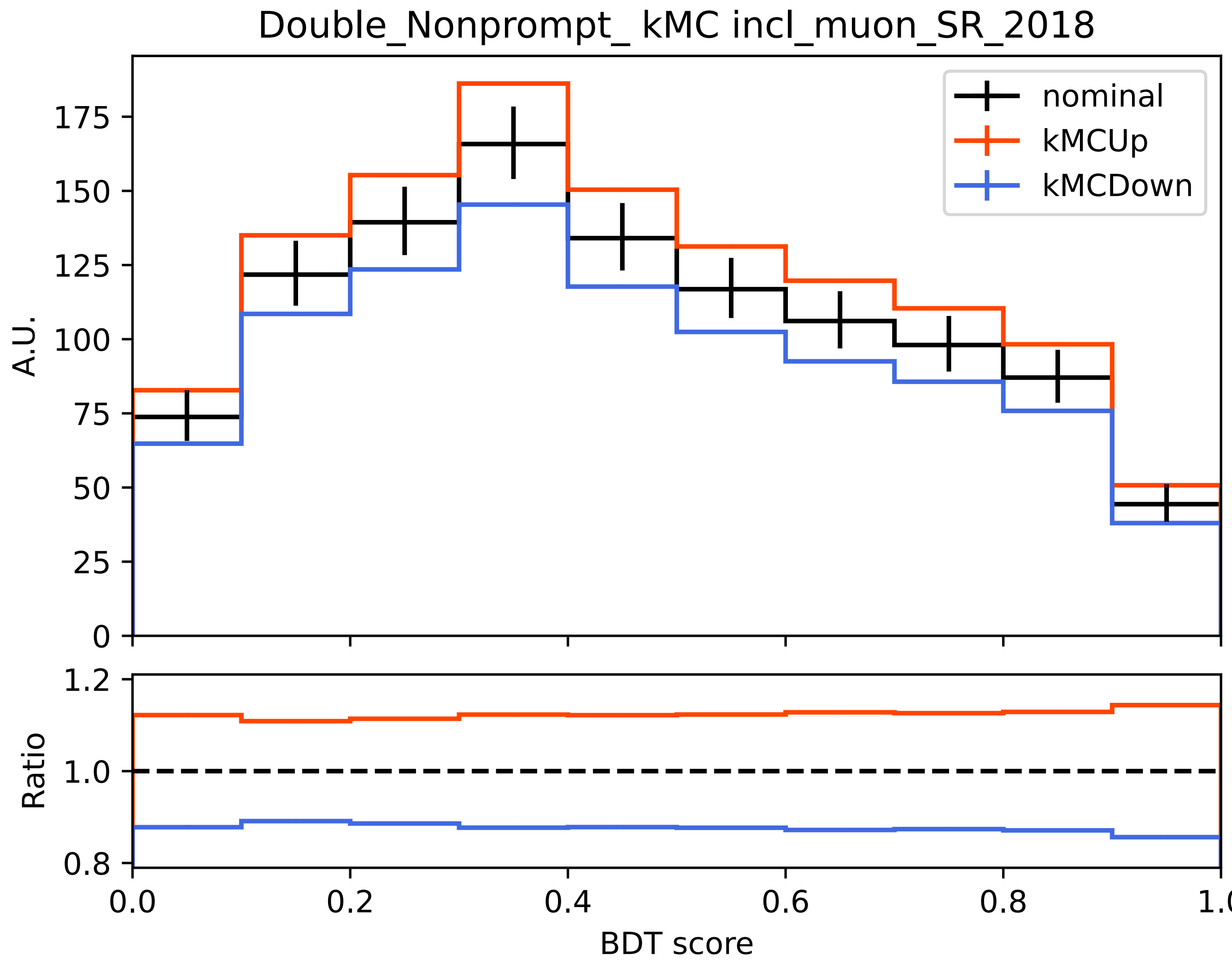
Compare shape corrections calculated from MC (left) and data (right), we consider **10% (barrel** photon) and **20%** (**endcap** photon) uncertainty

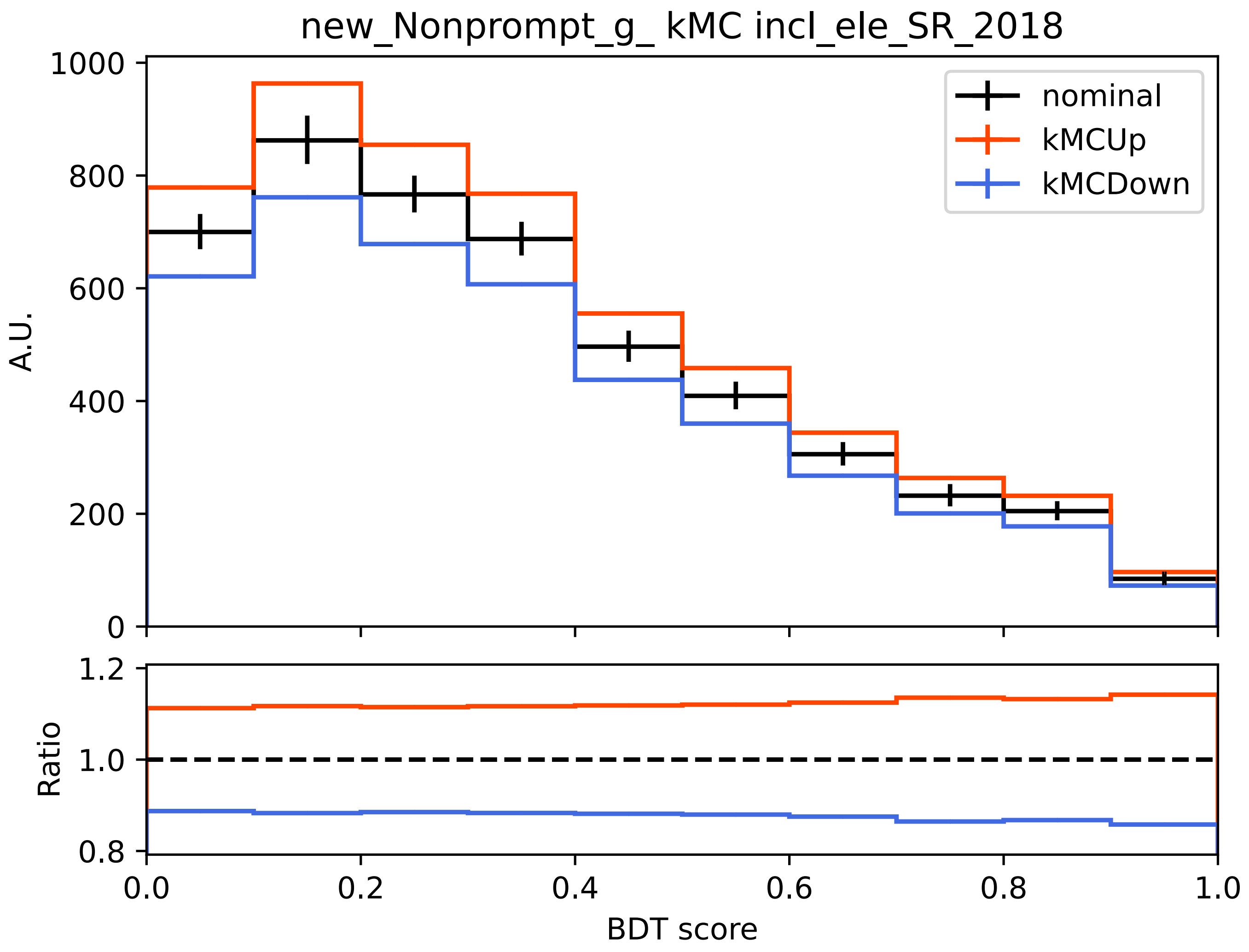
Is there an uncertainty assigned to the MC correction in equation 9 as data can be different from MC?

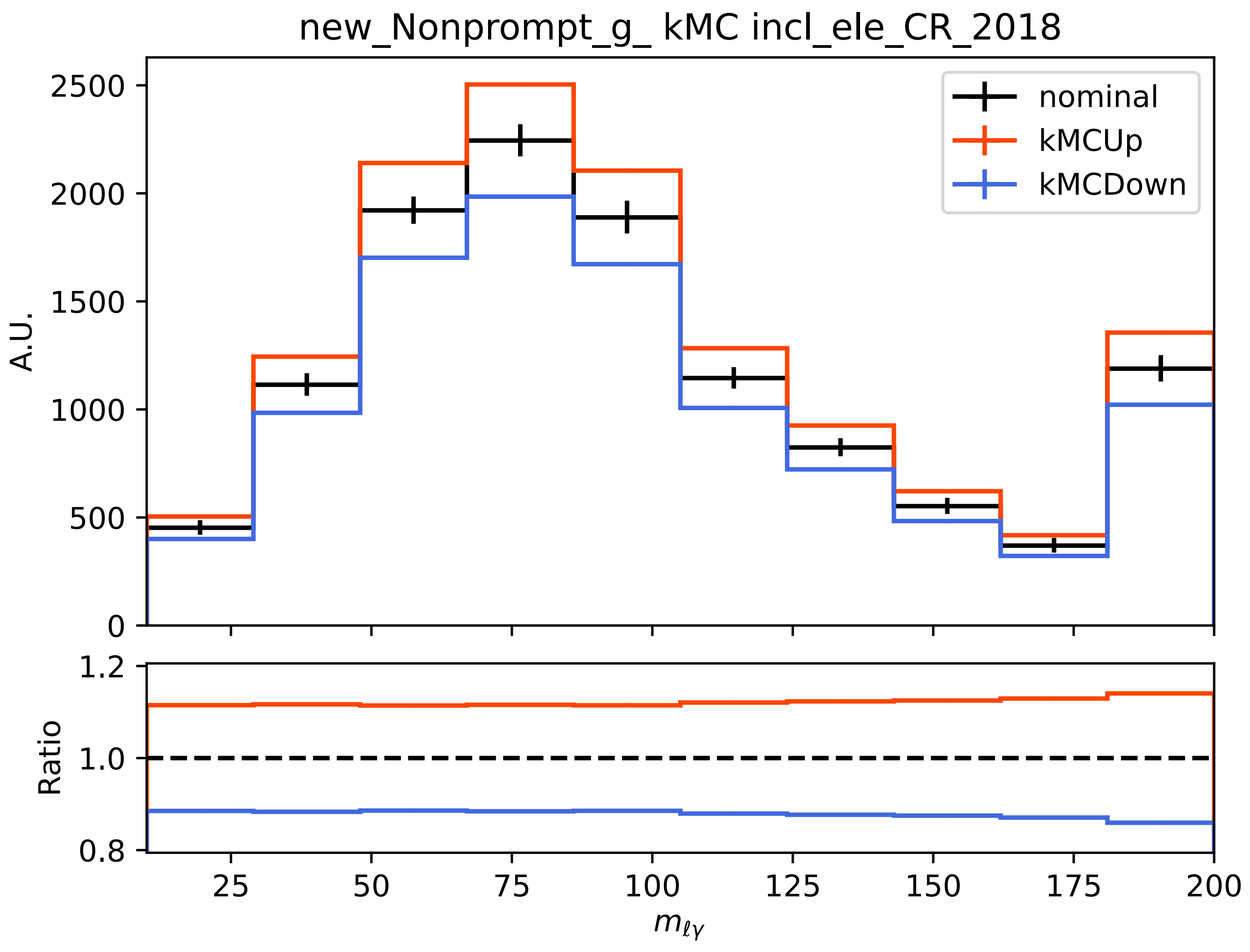
Compare shape corrections calculated from MC (left) and data (right), we consider **10% (<u>barrel</u>** photon) and **20%** (**<u>endcap</u>** photon) uncertainty











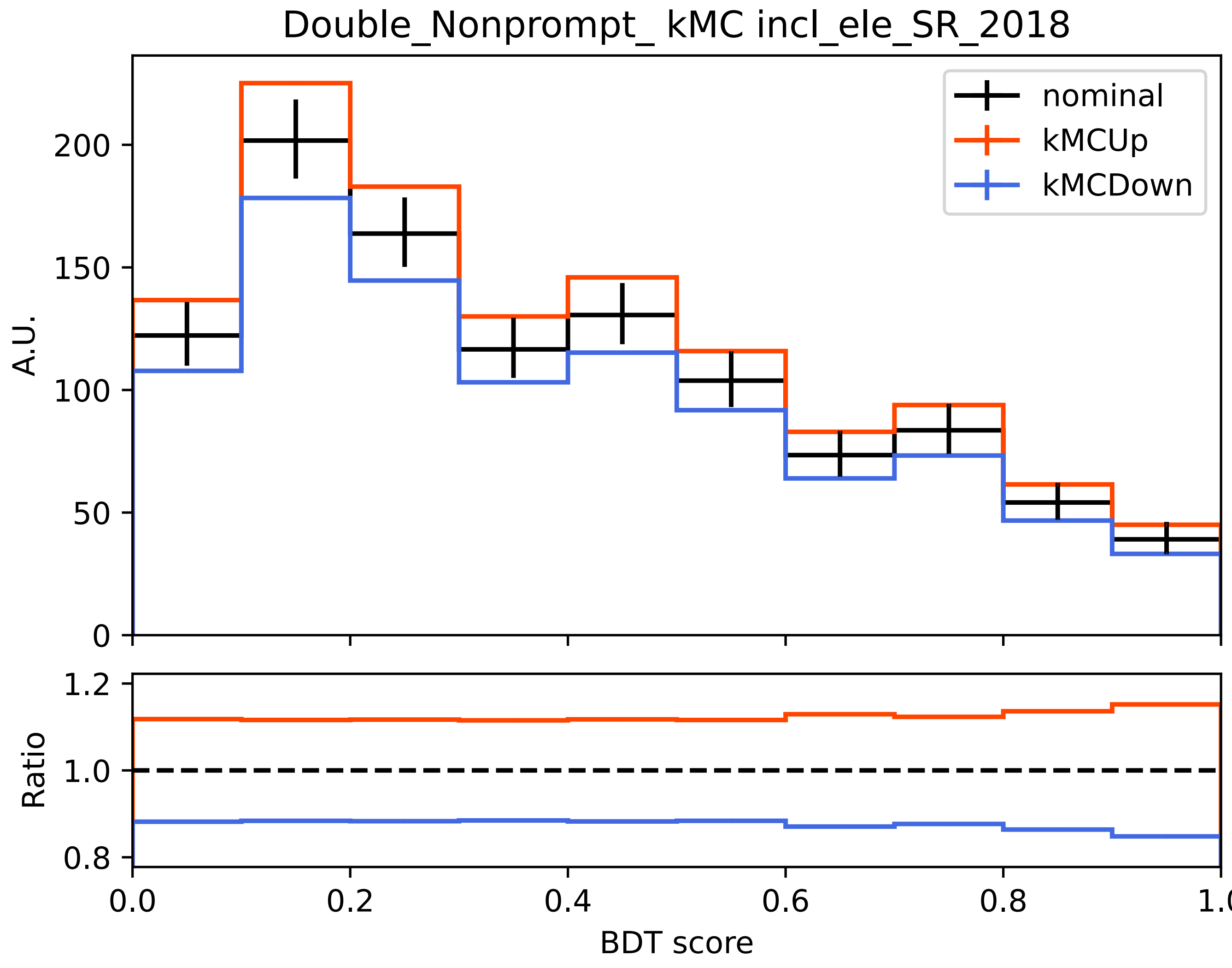


Figure 11 and 17, what is the normalization of the histograms?

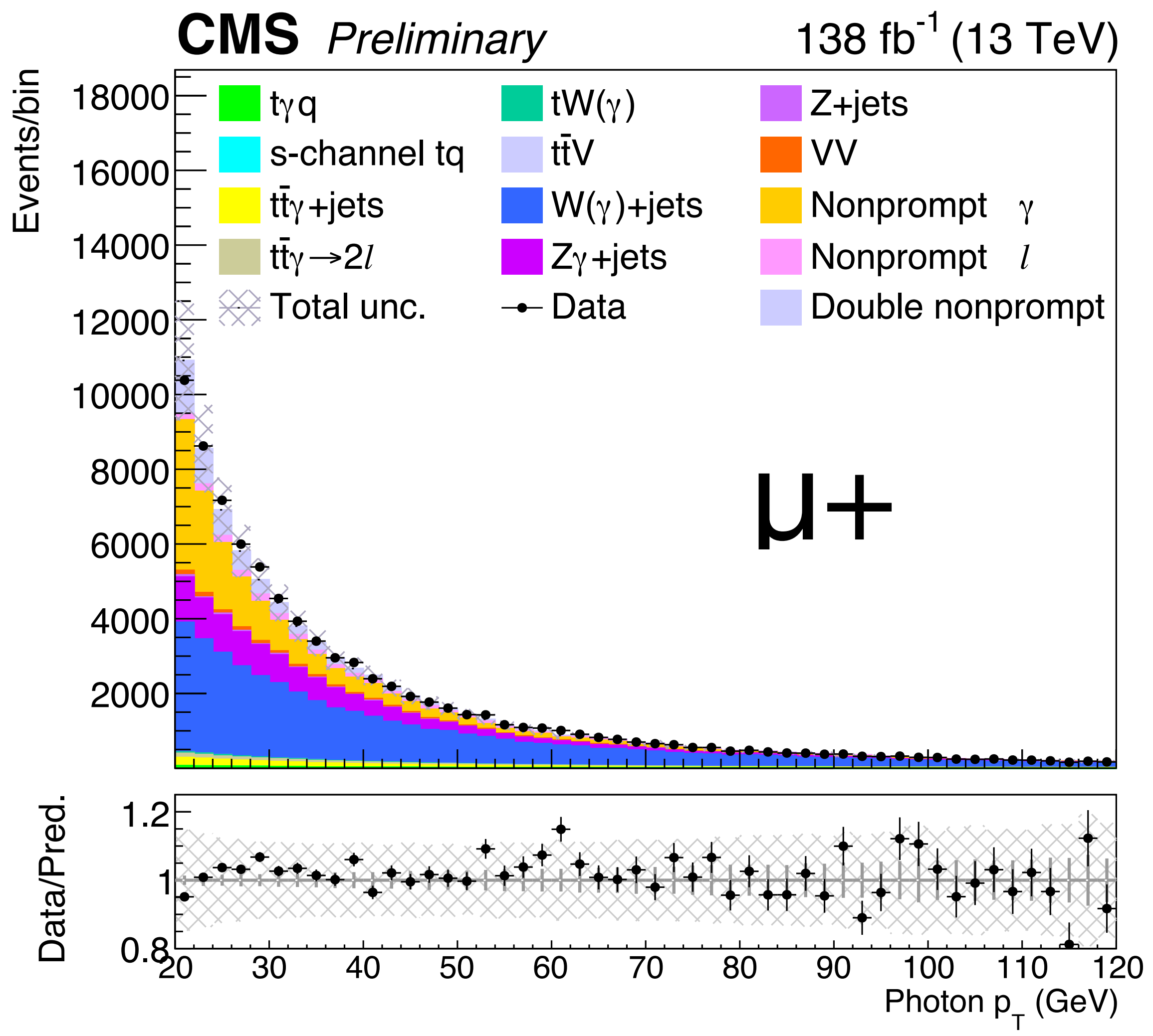
Figure 11 is showing the event yields. Figure 17 was in fact stolen from Beatriz and it's a normalised distribution that shows the shape.

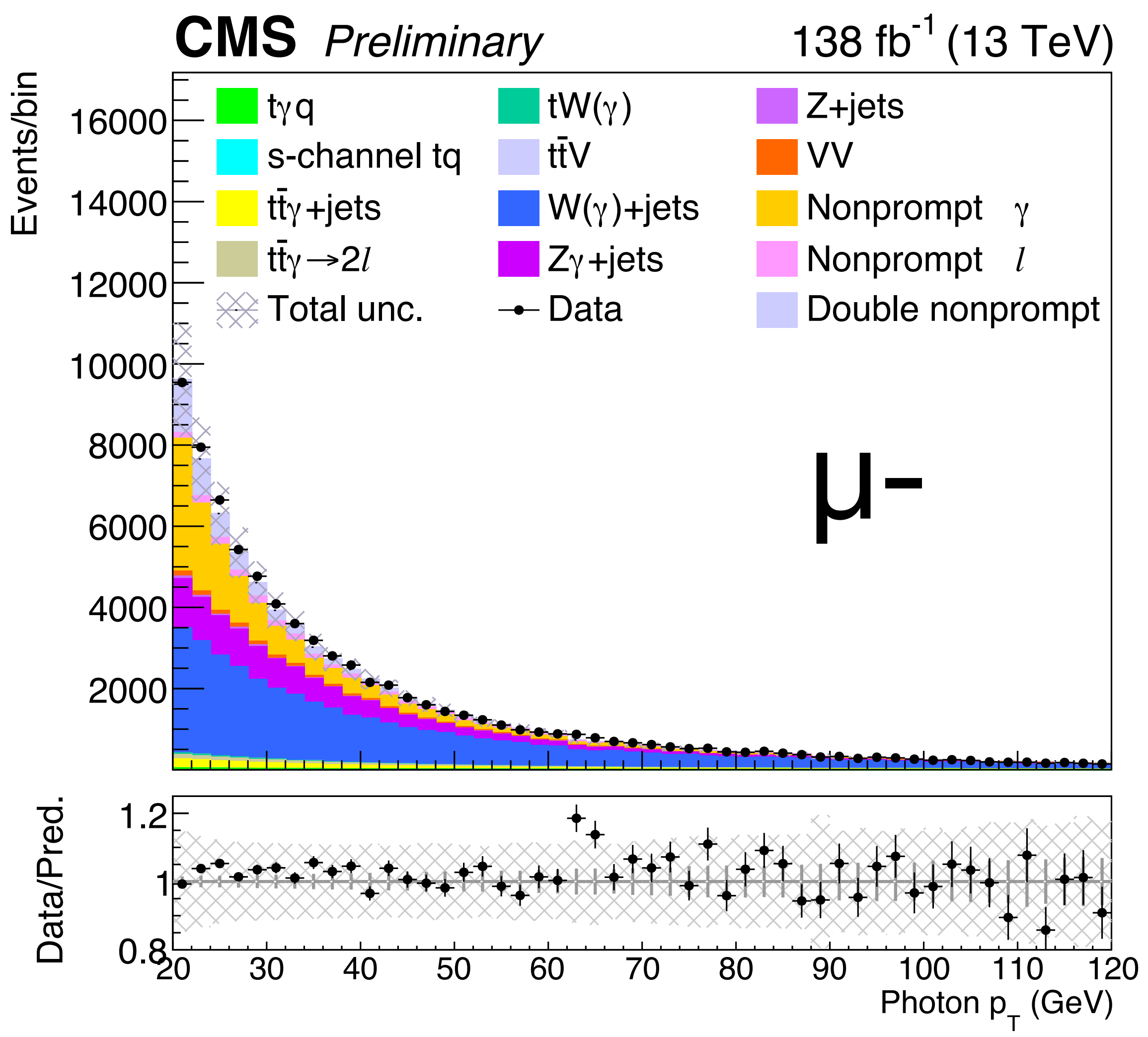
Figure 25: It does not look like the uncertainty on the data-driven nonprompt is being propagated correctly, the size of some error bars is much larger than the bin to bin fluctuation.

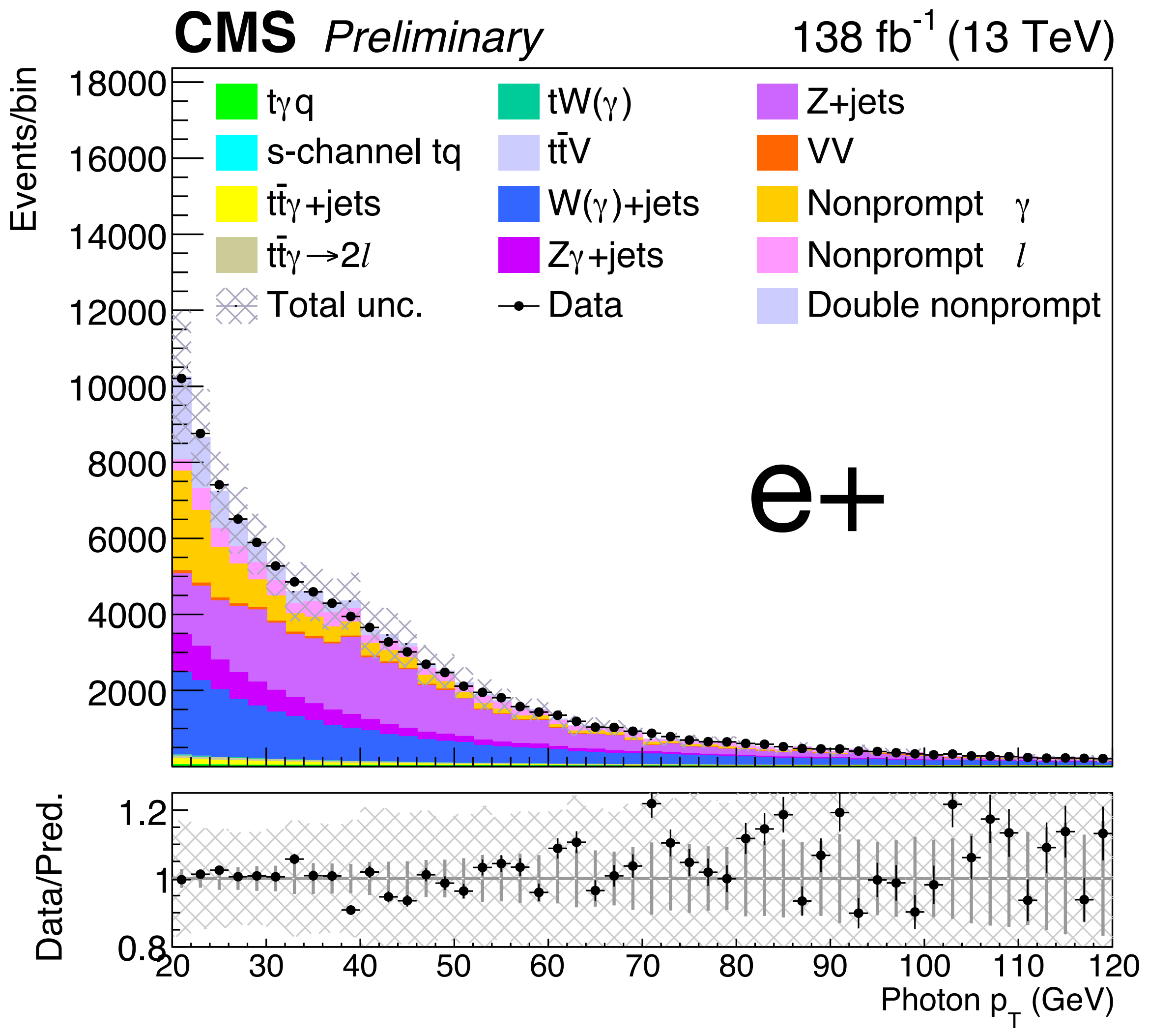
For these closure plots, I don't apply shape uncertainties but the constant factor of 30% for every bin

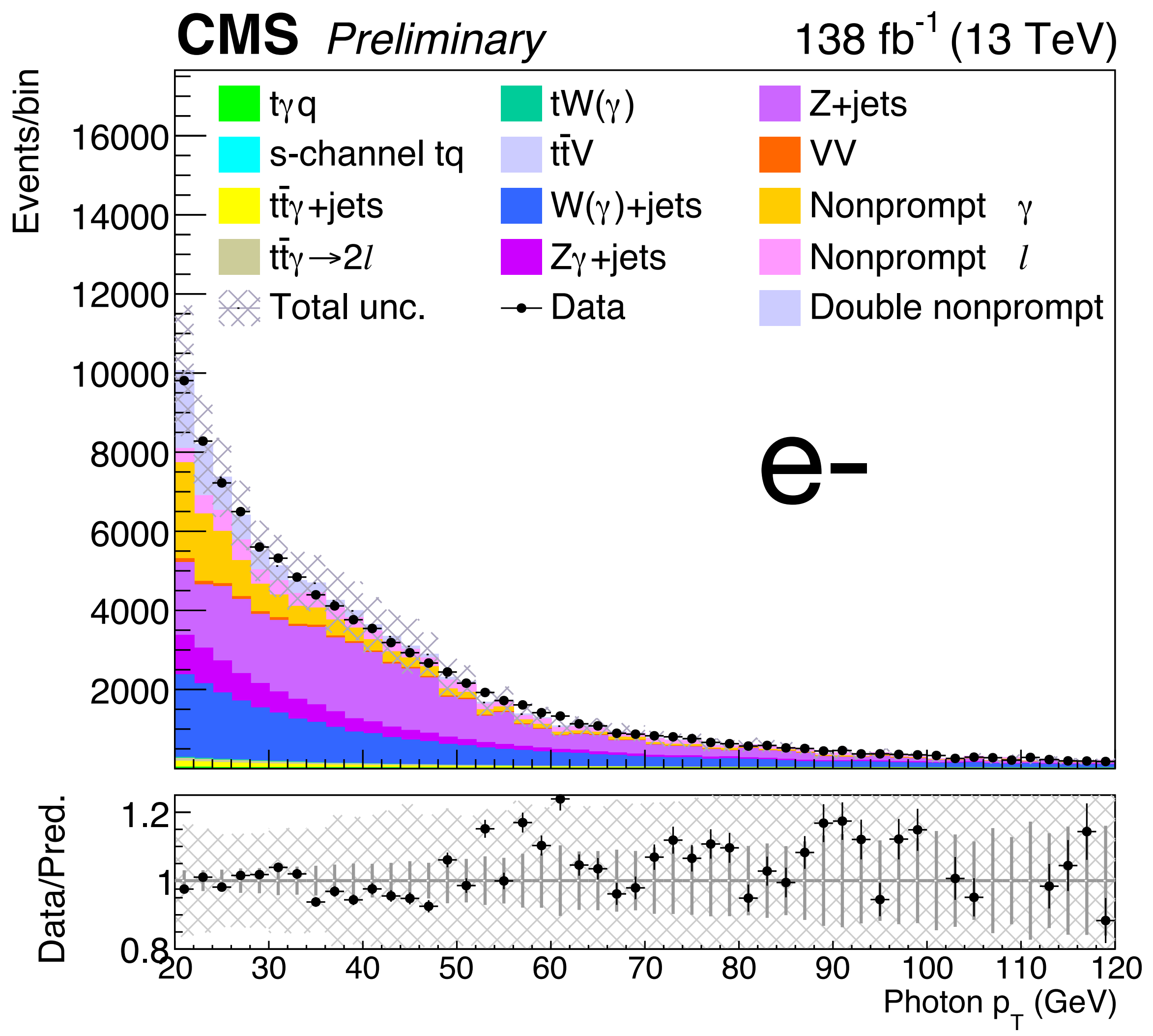
Did you ever check the non-prompt estimation separately for different lepton charges?

Check results are as follows, the agreement in the b-veto CR looks reasonable in different lepton charges.









The binning of figures 45-48 does not make sense in many cases, reduce the number of bins.

Ongoing

L1 prefiring:
-This is among the leading uncertainties and may need more care.

-There are two types of L1 prefiring, the L1 ECAL prefiring in 2016 and 2017 and the L1 muon prefiring in all years of Run 2, including 2018. In the text it reads as if only the ECAL prefiring is considered, could you check and clarify?

-This uncertainty has to be split in L1 muon and ECAL prefiring. And each one split in statistical (uncorrelated by era) and systematic (correlated). The full breakdown is available in NanoAODv9, see: https://cms-nanoaod-integration.web.cern.ch/autoDoc/NanoAODv9/2017UL/docTTToSemiLeptonic_TuneCP5_13TeV-powheg pythia8_RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1.html#L1PreFiringWeight

- NanoAOD v9 provide L1 ECAL, L1 Muon, and their products called L1PreFiringWeight. I am using their products as the final L1 prefiring weights.
- The splitting is implemented.

Row	L1PreFiri	L1PreFiri	L1PreFiri
0	1	1	1
1	0.9759036	1	0.9759036
2	1	1	1
3	0.9655726	1	0.9655726
4	0.9241010	1	0.9241010
5	0.9627832	0.9935401	0.9690431

How is the 10% uncertainty on Wgamma and Zjets justified?

It's just a conservative value to cover any deviations. If shape uncertainty is used, this log-normal uncertainty is not needed anymore

Check the trigger scale factor section and improve clarity.

Ongoing

Table 23, is the eta binning in the "Nonprompt γ from η binning" category correct?

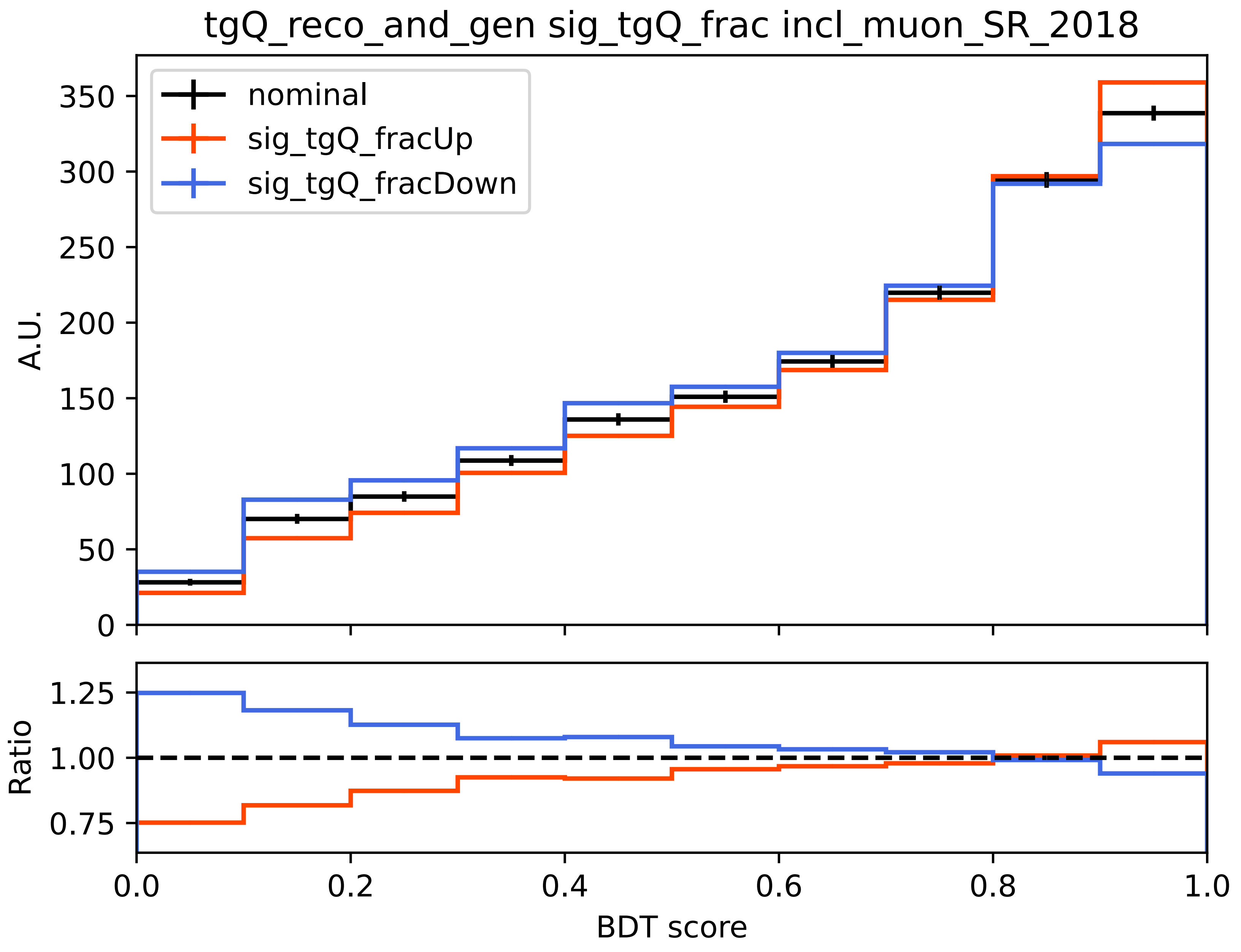
Fixed

Each ttgamma and tqg signal is made of 2 simulations, with gamma from ME level and from PS.
Is there an uncertainty assigned to the fraction of these two simulations?

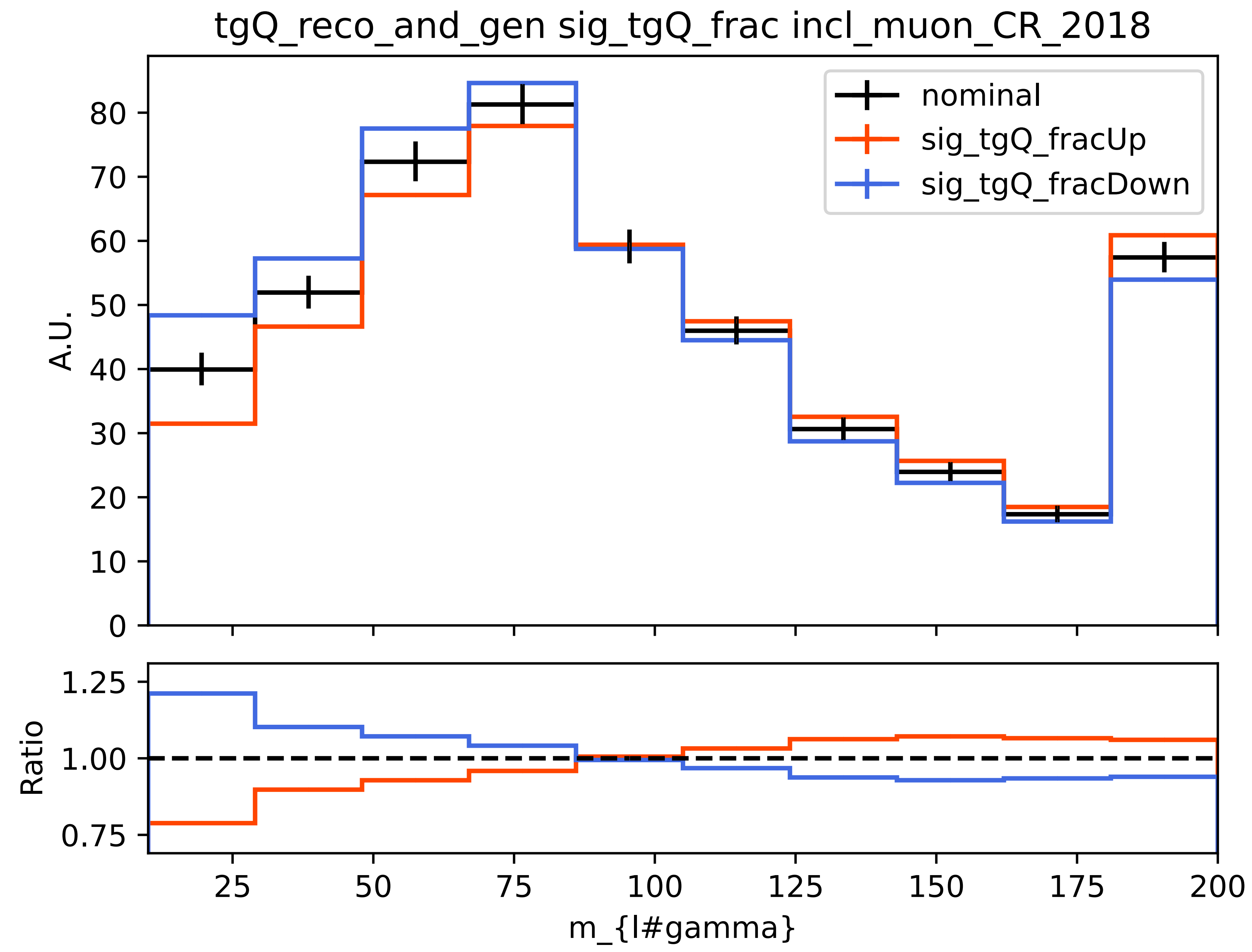
-There should be, already because the samples are produced with different generators (aMC@NLO and powheg) but also because the photon modelling is different.

-An uncertainty can be assigned by scaling one part of the 2 simulations up and the other down while keeping the total cross section constant. The size of the variation should be motivated, but as a start, to see how important it is, one can use e.g. 10% and check the impact. Both for tqgamma and ttgamma.

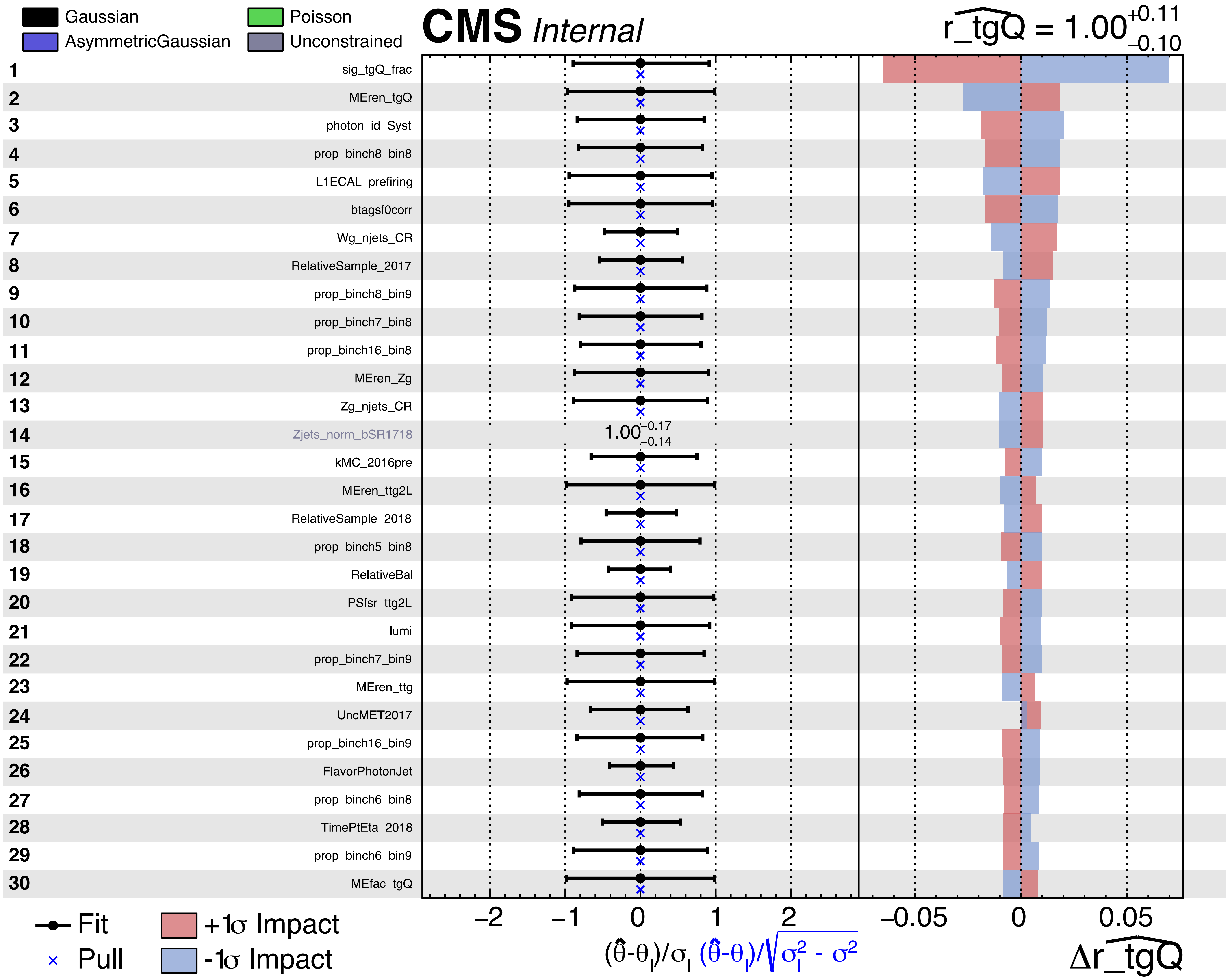
As you suggest, this kind of uncertainty is calculated and applied. This uncertainty has large impact for tgQ and less important for ttg



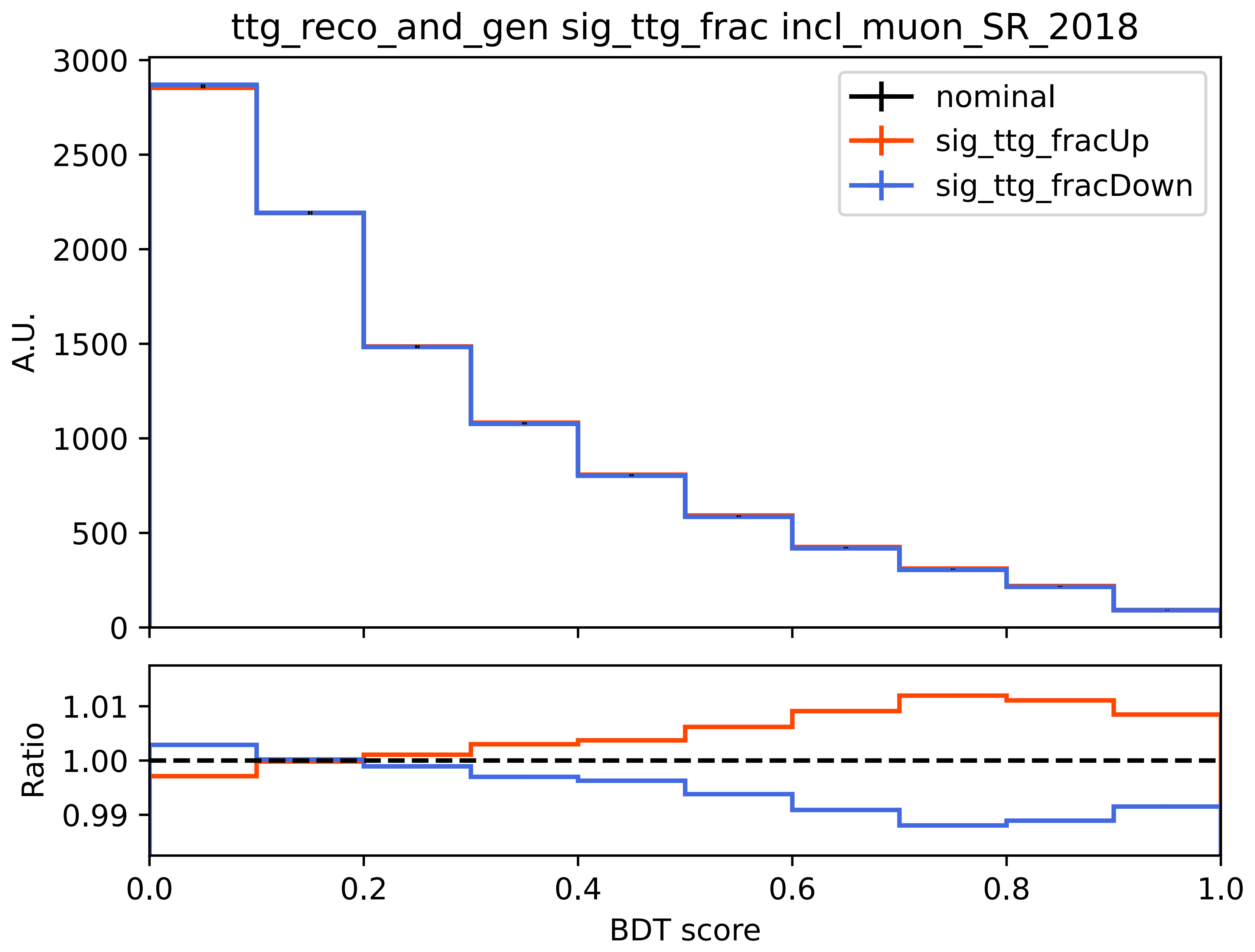
tgQ_reco_and_gen sig_tgQ_frac incl_muon_SR_2018
nominal
sig_tgQ_fracUp
sig_tgQ_fracDown
A.U.
Ratio
BDT score



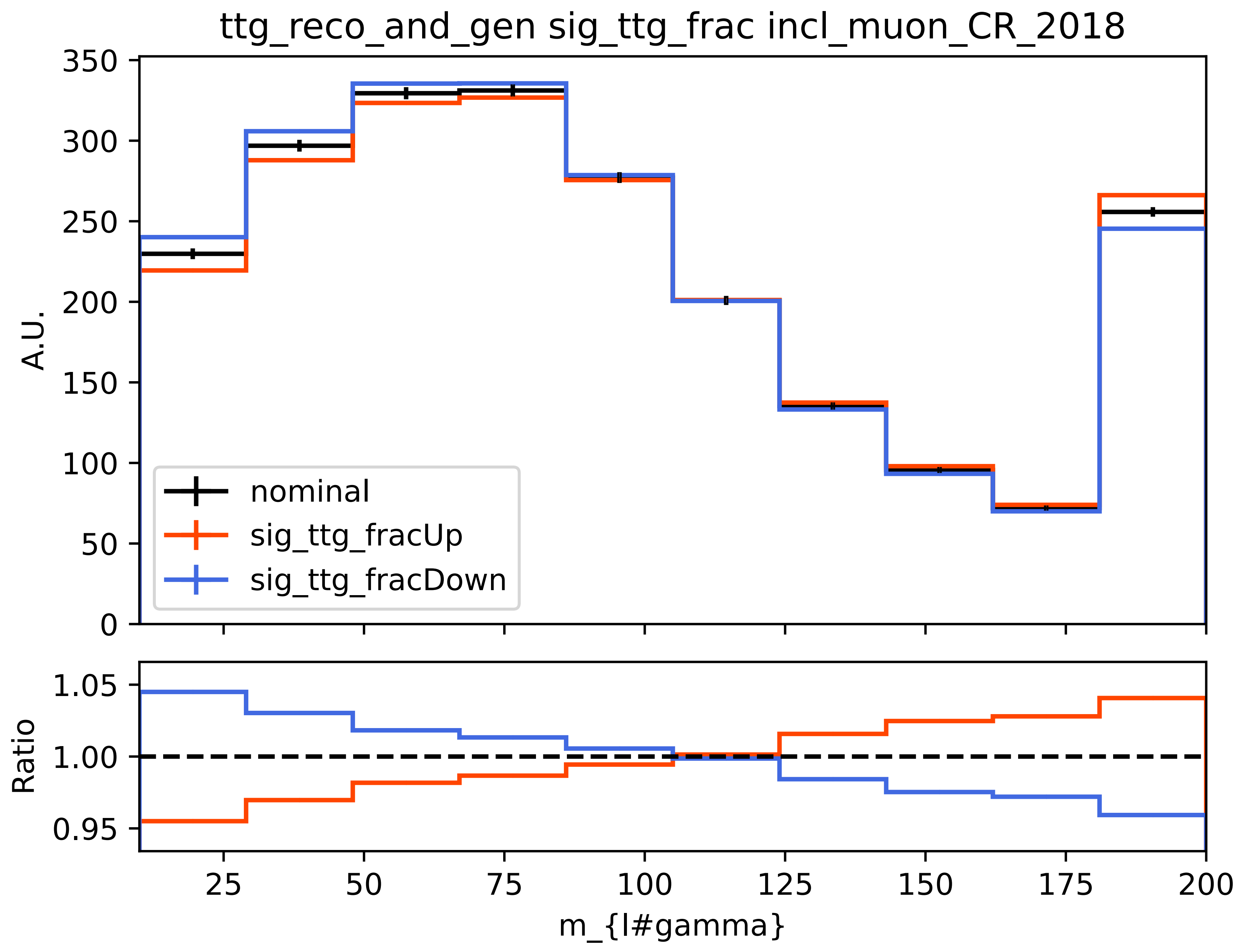
tgQ_reco_and_gen sig_tgQ_frac incl_muon_CR_2018
nominal
sig_tgQ_fracUp
sig_tgQ_fracDown
A.U.
Ratio
m_{l#gamma}



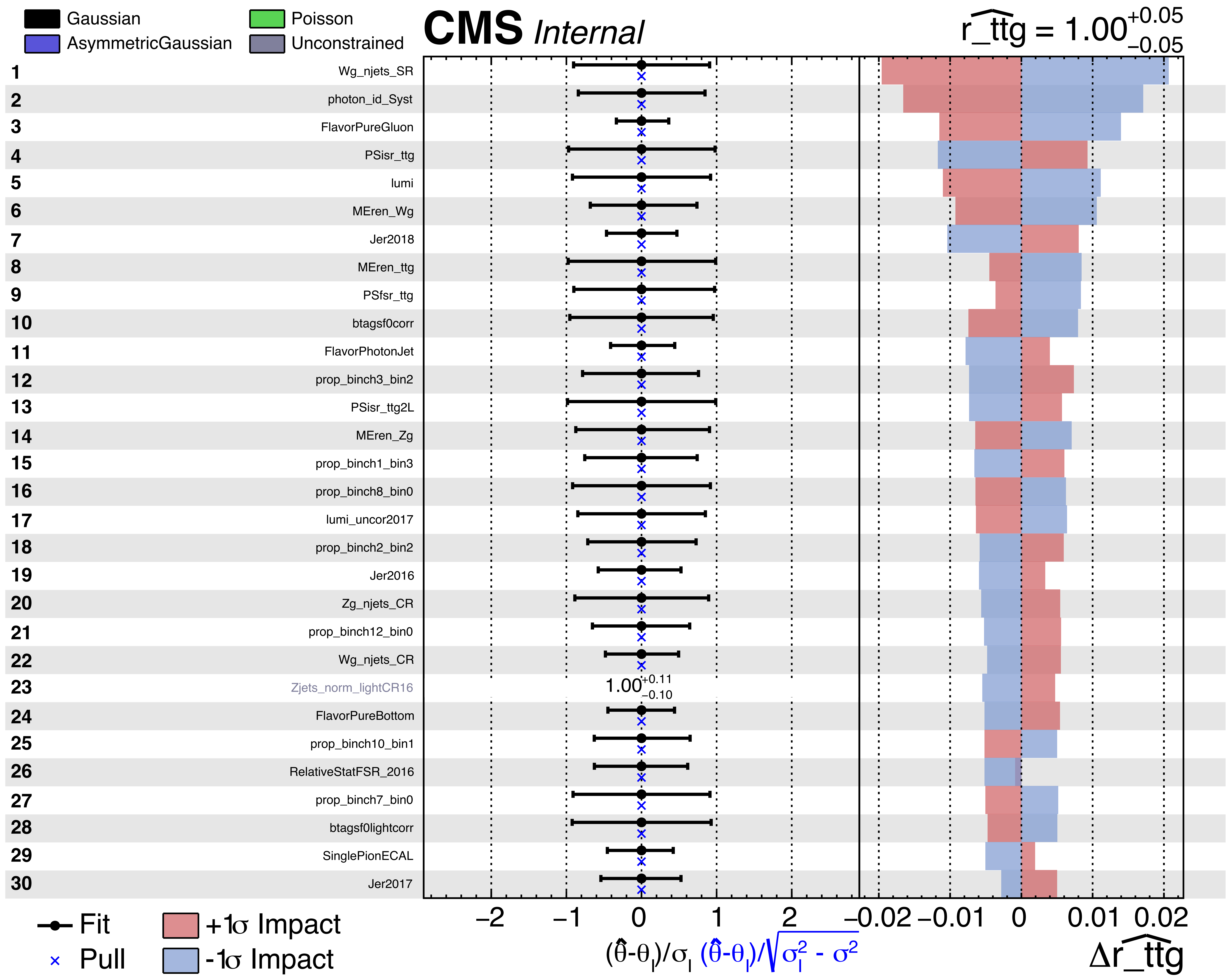
Gaussian
Poisson
AsymmetricGaussian
Unconstrained
CMS Internal
r_tgQ = 1.00 +0.11 −0.10
1 sig_tgQ_frac
2 MEren_tgQ
3 photon_id_Syst
4 prop_binch8_bin8
5 L1ECAL_prefiring
6 btagsf0corr
7 Wg_njets_CR
8 RelativeSample_2017
9 prop_binch8_bin9
10 prop_binch7_bin8
11 prop_binch16_bin8
12 MEren_Zg
13 Zg_njets_CR
14 Zjets_norm_bSR1718
1.00 +0.17 −0.14
15 kMC_2016pre
16 MEren_ttg2L
17 RelativeSample_2018
18 prop_binch5_bin8
19 RelativeBal
20 PSfsr_ttg2L
21 lumi
22 prop_binch7_bin9
23 MEren_ttg
24 UncMET2017
25 prop_binch16_bin9
26 FlavorPhotonJet
27 prop_binch6_bin8
28 TimePtEta_2018
29 prop_binch6_bin9
30 MEfac_tgQ
Fit
Pull
+1σ Impact
-1σ Impact
Δr_tgQ



ttg_reco_and_gen sig_ttg_frac incl_muon_SR_2018
nominal
sig_ttg_fracUp
sig_ttg_fracDown
A.U.
Ratio
BDT score



ttg_reco_and_gen sig_ttg_frac incl_muon_CR_2018
nominal
sig_ttg_fracUp
sig_ttg_fracDown
A.U.
Ratio
m_{l#gamma}



Gaussian
Poisson
AsymmetricGaussian
Unconstrained
CMS Internal
r_ttg = 1.00 +0.05 −0.05
1 Wg_njets_SR
2 photon_id_Syst
3 FlavorPureGluon
4 PSisr_ttg
5 lumi
6 MEren_Wg
7 Jer2018
8 MEren_ttg
9 PSfsr_ttg
10 btagsf0corr
11 FlavorPhotonJet
12 prop_binch3_bin2
13 PSisr_ttg2L
14 MEren_Zg
15 prop_binch1_bin3
16 prop_binch8_bin0
17 lumi_uncor2017
18 prop_binch2_bin2
19 Jer2016
20 Zg_njets_CR
21 prop_binch12_bin0
22 Wg_njets_CR
23 Zjets_norm_lightCR16
1.00 +0.11 −0.10
24 FlavorPureBottom
25 prop_binch10_bin1
26 RelativeStatFSR_2016
27 prop_binch7_bin0
28 btagsf0lightcorr
29 SinglePionECAL
30 Jer2017
Fit
Pull
+1σ Impact
-1σ Impact
Δr_ttg

What is the uncertainty on the out of acceptance ttgamma and tqgamma contributions? As they are not scaled with the signal strength parameters there should probably be a separate (lnN) uncertainty assigned.

The types of the uncertainty for events out of acceptance is the same as the signal events. And uncertainty for signal events is only with shape effect; uncertainty for out of acceptance events is with both normalzation and shape effects.

The impact of JetFlavorQCD is with rank 7 on ttgamma rather high and significantly constrained. In this case it is recommended to split this source into the different flavor components: b, c, light, gluon.

Splitting is implemented for FlavourQCD and other sources

What about things like hdamp or top mass variations, their impact should be checked given that the precision of ttgamma reaches 4%

Currently, these two uncertainties are not applied

Fig 55: Is the disagreement in these plots (particularly in the Z-mass peak) problematic, or can the fit accommodate it?

It's because the plot is the electron prefit plots, the ele mis. γ from Zjets doesn't really float in prefit

Are underflow/overflow events in the CR included in the fit?

$m_{l\gamma}$ is from 10 to 200 GeV with overflow events included

Please explicitly state which categories are passed to the fit: contributions are separated by era and lepton flavor? Yielding 160 bins in the inclusive fit? And they are not separated by lepton charge?

For SR: distribution of BDT in 10 bins from 0 to 1
For CR: distribution of mlγ in 10 bins from 10 to 200 including the overflow events
Distributions are separate to electron and muon channel without considering the lepton charge

Details will be added in next version of note

Figure 56 shows the same upper and lower plots

Fixed

Why are three variables measured at particle level and one at parton level?

We plan to add more variables, the light jet η at particle level, and top charge at parton level

A measurement of the top quark-antiquark cross sections separately and their ratio should be added for tqgamma.

The fit should be binned in positive and negative charge to facilitate the measurement of top quark and antiquark. It can also help to reduce the correlation between tqgamma and ttgamma cross sections since ttgamma is charge symmetric while tqgamma is not (Complementary to the BDT score since the charge information is not included in the BDT).

This might require some checks/updates on the nonprompt estimation since it's not clear that the nonprompt/fakerates are the same in both charges.

- As shown in previous slides, the nonprompt estimation in different lepton charge performs well.
- The measurement of the top quark-antiquark cross sections can be performed with current nonprompt estimation and will be perform after we get converge for all comments

Once unblinded, it would be nice to add some goodness of fit test to quantify the agreement.

For example using a saturated likelihood test. When using the same distribution that is fit for unfolding also in an inclusive fit, the likelihood ratio is approximately chi2 distributed with the number of degrees of freedom from the difference in free parameters, i.e. 8 for a fit with 5 gen bins. Then Chi2(saturated)/ndf = 2*(log(L_inclusive) - log(L_differential)). As an alternative, a chi2 test including the covariance can be done. This can already be set up and tested on toys, otherwise later after unblinding is fine.

The goodness fit for Asimov data always give me the observed test statistic around 0, so we'll always have p-value around 1. We'll of course do the goodness fit if we get the GL.

Quantities such as a measurement of the top quark polarization could be added (see for example TOP-20-010)

Are there any other (differential) theory predictions that the result can be compared to? E.g. the top quark-antiquark ratio could be compared to predictions with different PDF sets. Are different PDF sets available in the current samples? If not, we could request gen events. Also e.g. for 5FS tqgamma maybe.

For the top polarisation, we don’t have the plan to include it in the paper, it can be left for Run3

In current tγq NLO sample, it has PDF sets
- NNPDF31_nnlo_as_0118_nf_4_mc_hessian
- NNPDF31_nnlo_as_0118_nf_4
- CT10nlo_nf4 (2 weights)
- MSTW2008lo68cl_nf4
- …
- MMHT2014nlo68cl_nf4
- …
- PDF4LHC15_nlo_nf4_30
- NNPDF31_nnlo_hessian_pdfas
- …

Summary

Update on selection:

- Decrease the p_T for second lepton rejection from 15 GeV to 10 GeV
- Sync the removal strategy to TOP-23-002
- Discard e/γ in an affected region in 2016preVFP data and MC
- Small bugs fixed like SFs application in overflow bins and so on

Cross check:

- Check the non-Zjets fake γ from electrons
- Z+jets normalization factor in VR and CR unblind fit
- Nonprompt γ fake rate in more fine p_T bins
- Significance results after some modifications
- Data/MC agreement in different lepton charges with the uniform nonprompt estimation

Update on uncertainty:

- Uncertainty assigned to the fraction of simulations tγq (ttγ) and single-top (ttbar)?
- Split uncertainties on photon/muon SFs
- Split uncertainties on JECs
- Split btagging uncertainties
- Split L1 pre-firing uncertainty
- Add Zγ/Wγ/Zjets shape uncertainties as a function of the number of jets
- Add more uncertainties on nonprompt background estimation
  - 2NPs for fake rates
  - 1NP for prompt subtraction
  - 1NP for kMC factor for nonprompt γ
- Add light jet η differential XS in particle level
- Add top quark charge differential XS in Parton level