Simultaneous $t\gamma q$+$t\bar{t}\gamma$ measurement

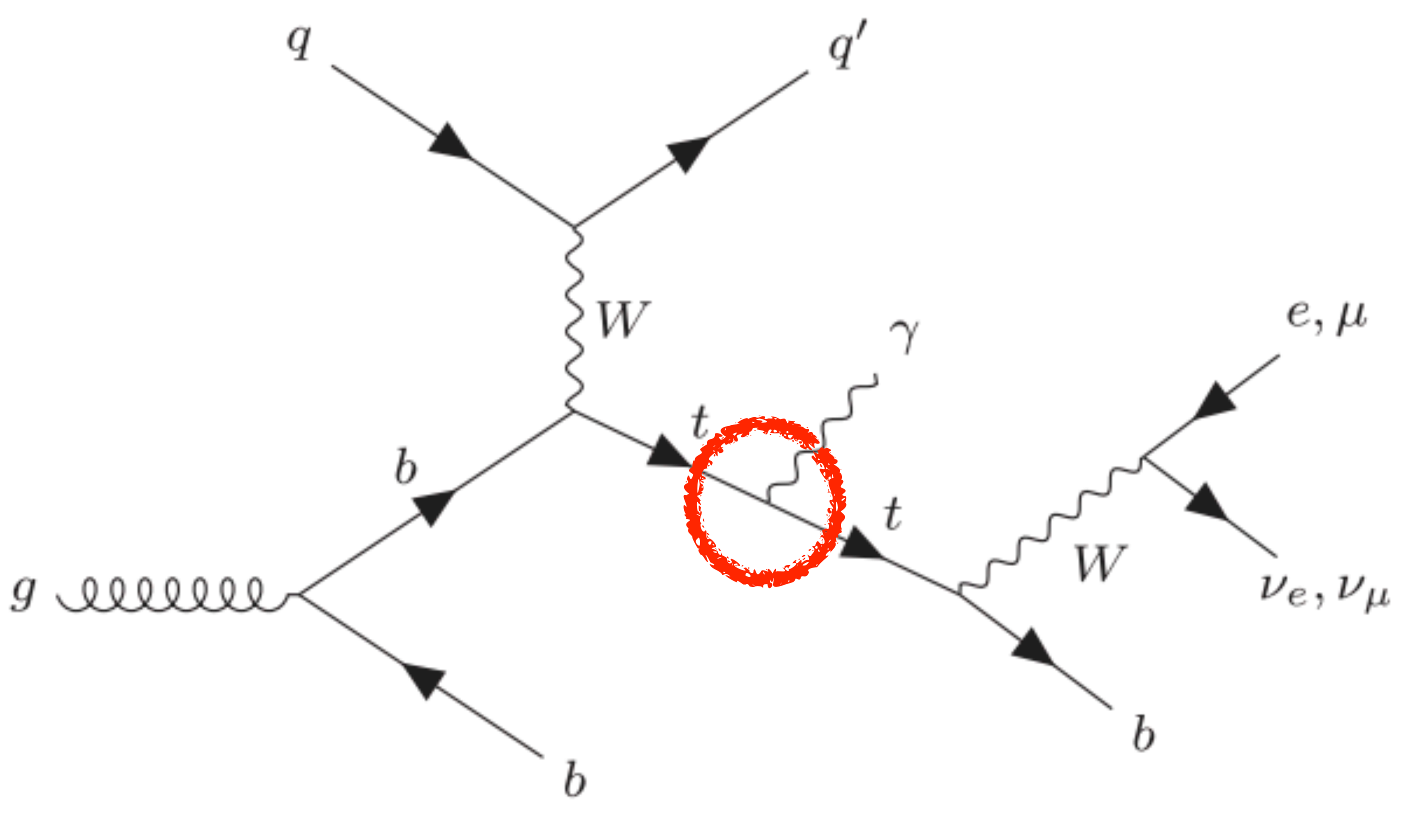
Ying AN, Hugo Becerril, Abideh Jafari, Maria Aldaya, Andreas Meyer

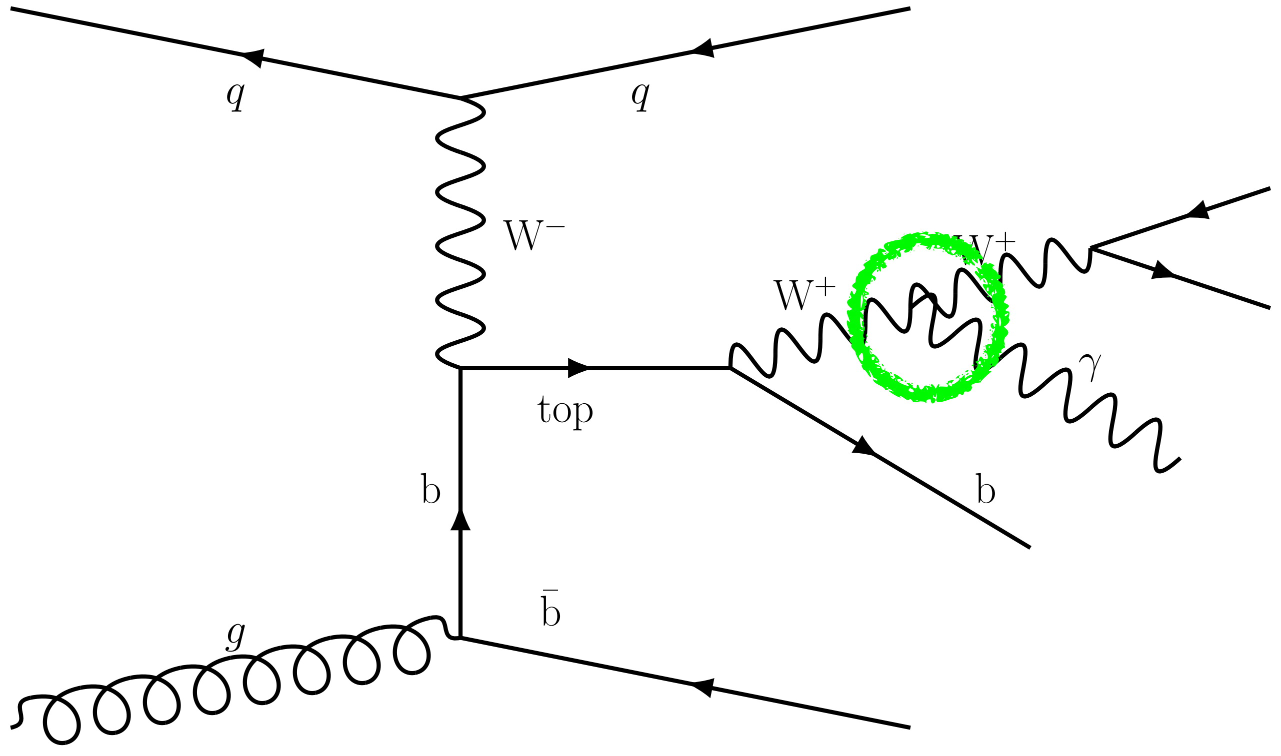
CMS TOP tX meeting

23 April 2024

HELMHOLTZ
SPITZENFORSCHUNG FÜR
GROSSE HERAUSFORDERUNGEN

Motivation

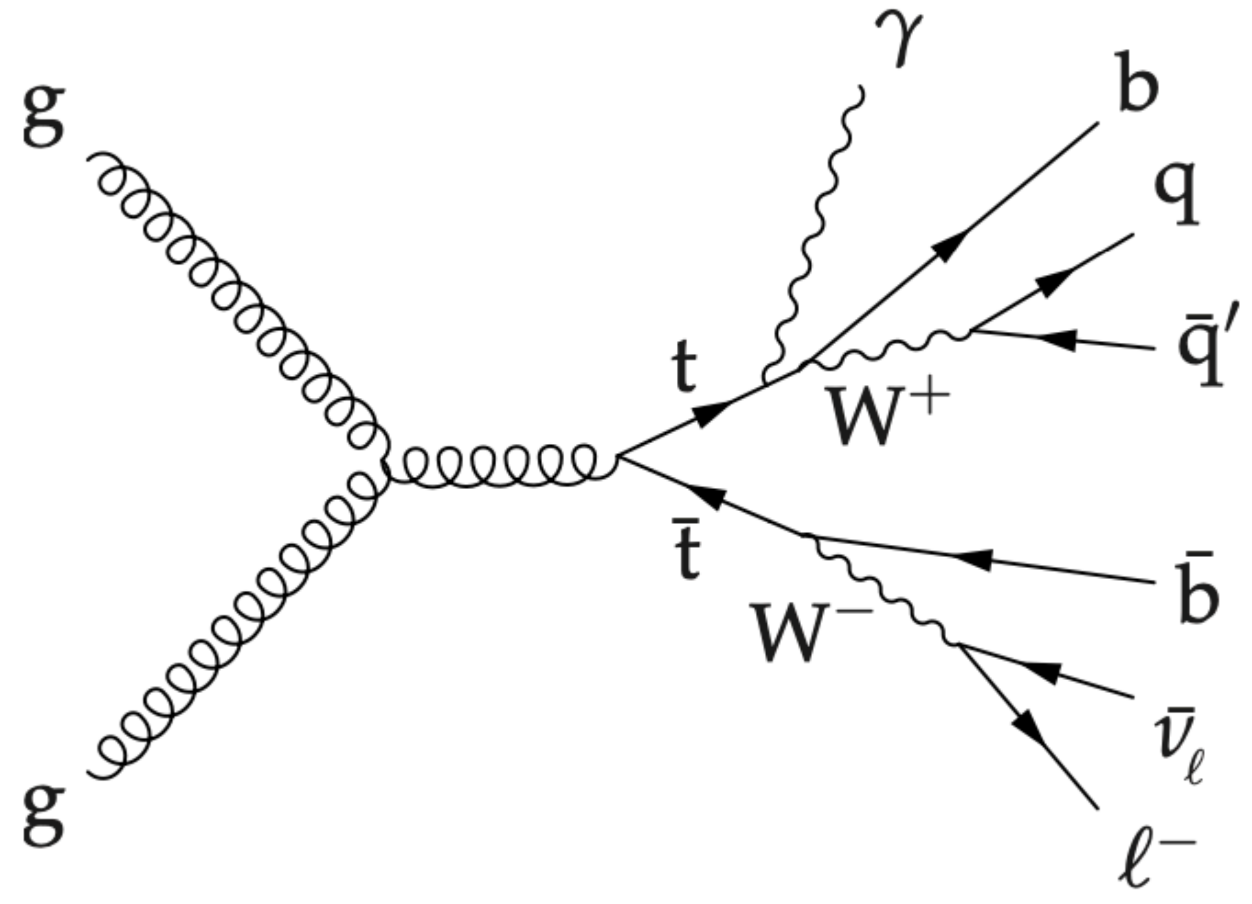




- The $t\gamma q$ process is observed by ATLAS. No differential cross section results to date. CMS has the evidence paper with partial data
- First-ever differential measurement
  - Fit tWγ at the same time if possible
- This process represents a direct probe of the top-photon coupling
  - Anomalous top coupling (top-photon electroweak) by top EFT fit is interesting (tγq+ttγ EFT)
  - Could perform the EFT interpretation for tγq+ttγ or even plus tWγ

Simultaneous fit for tγq+ttγ

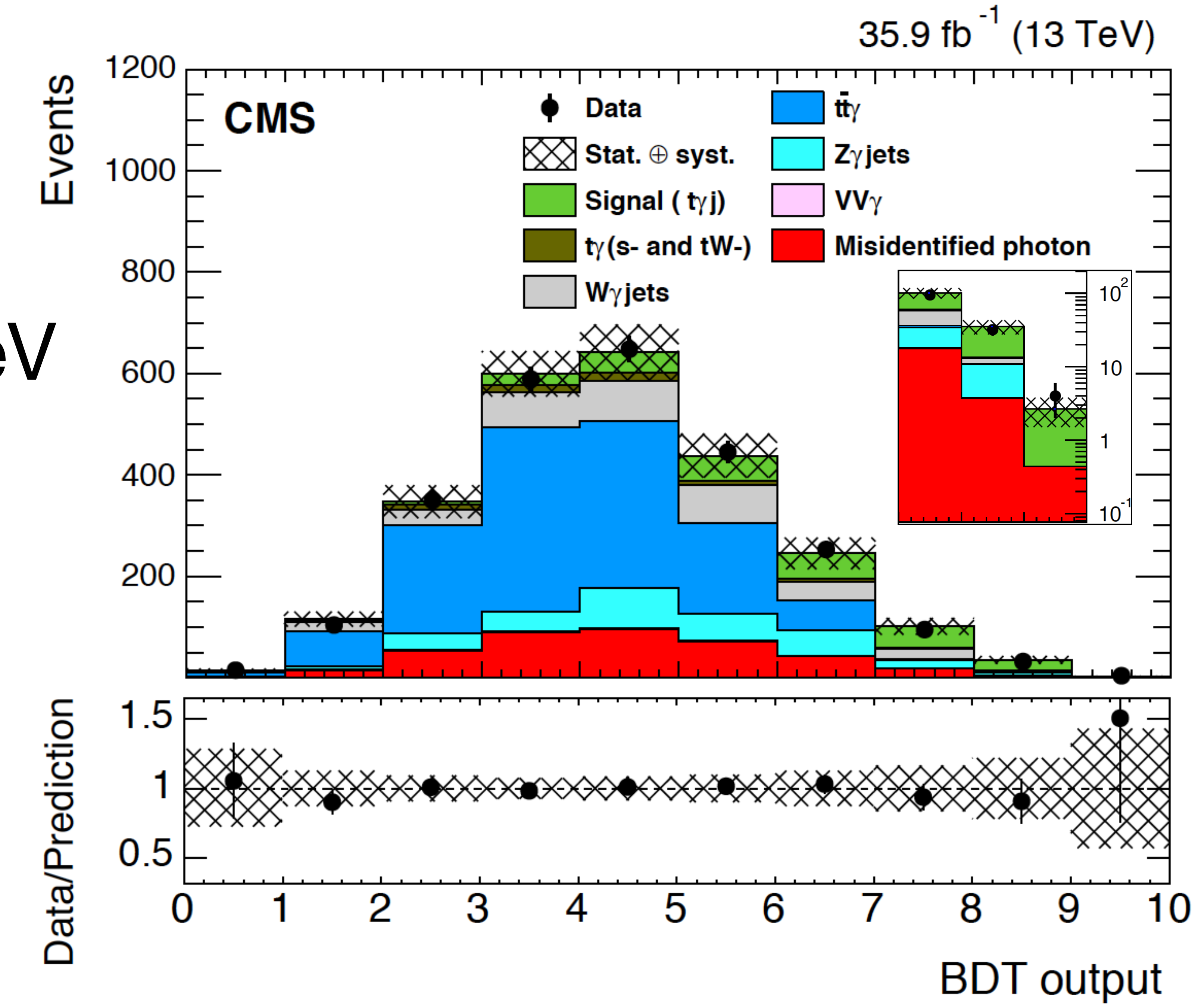
- Obtains full set of correlations between the two processes
- Possible for a more straightforward EFT interpretation
- High precision ttγ results



Overview

CMS result:

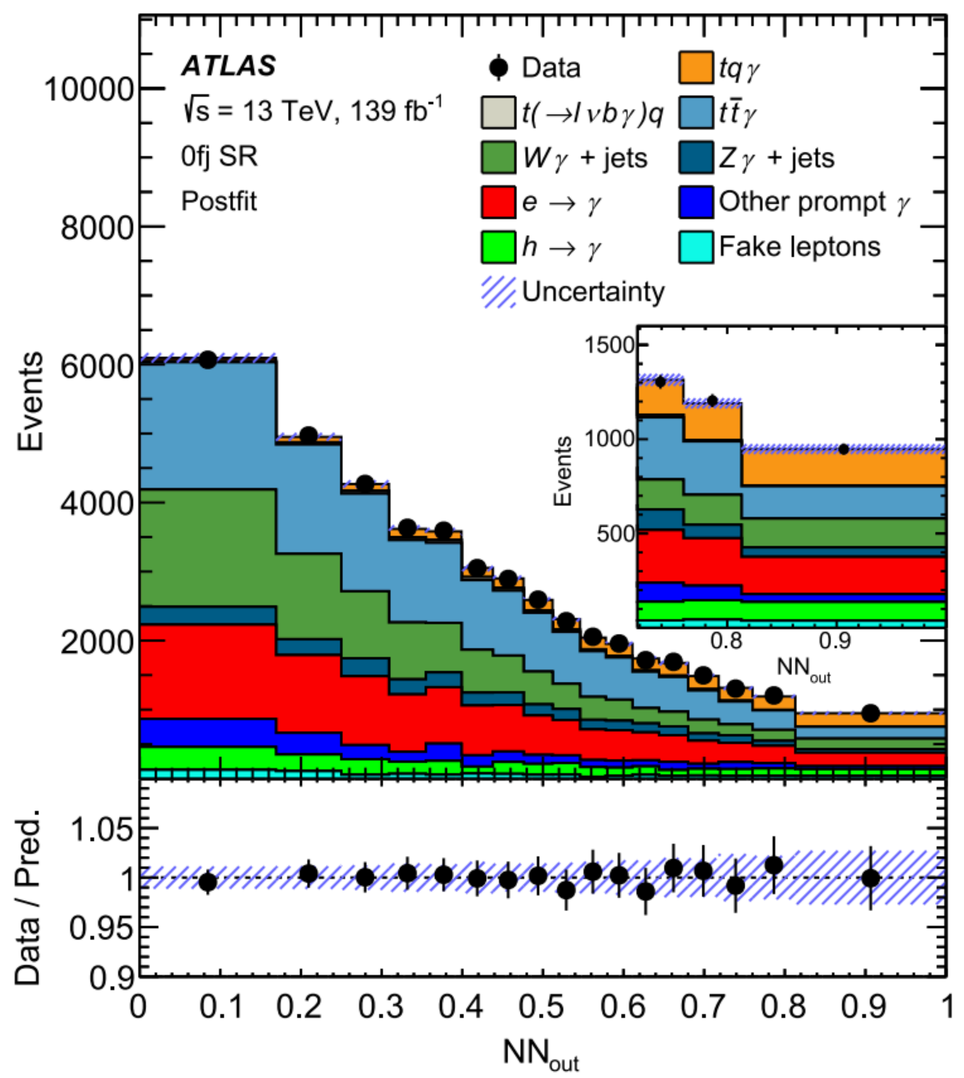
- Measure the μ channel making use of 2016 data only
  - Signal is with exactly 1γ, 1μ, 1b-jet, ≥1j, and satisfy MET > 30 GeV
- BDT is trained for tγq signal against the main background $t\bar{t}\gamma$
- Observed (Expected) significance is measured to **4.4 (3) σ**

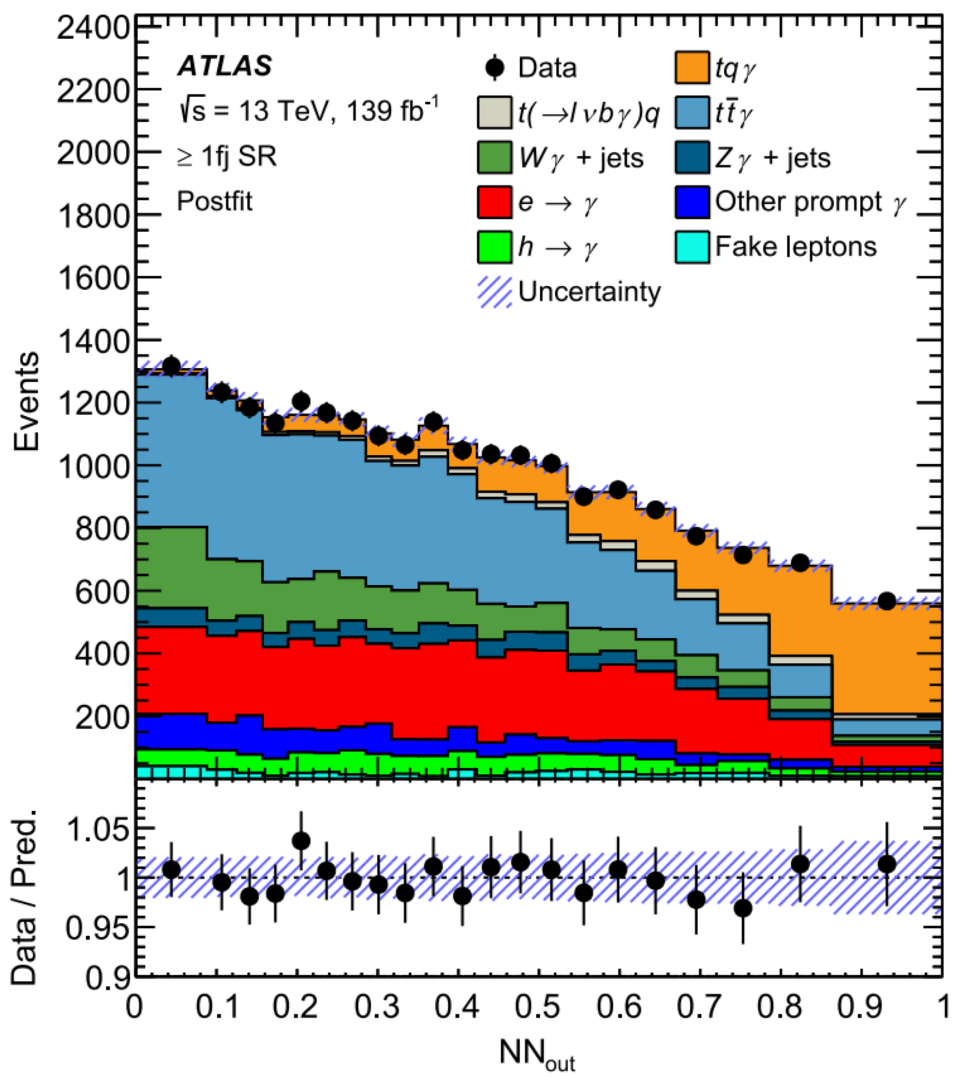


◆ Ongoing inclusive tγq measurement using full run 2 data by IPM

ATLAS result:

- Measure both the μ and e channel with full run 2 data
  - Signal is with exactly 1γ, 1ℓ, 1b-jet, and MET > 30 GeV
  - Categorise signal to 0fj and ≥1fj (number of forward jets)
- NNs are trained in the SRs
- Observed (Expected) significance is measured to **9.3 (6.8) σ**





backup

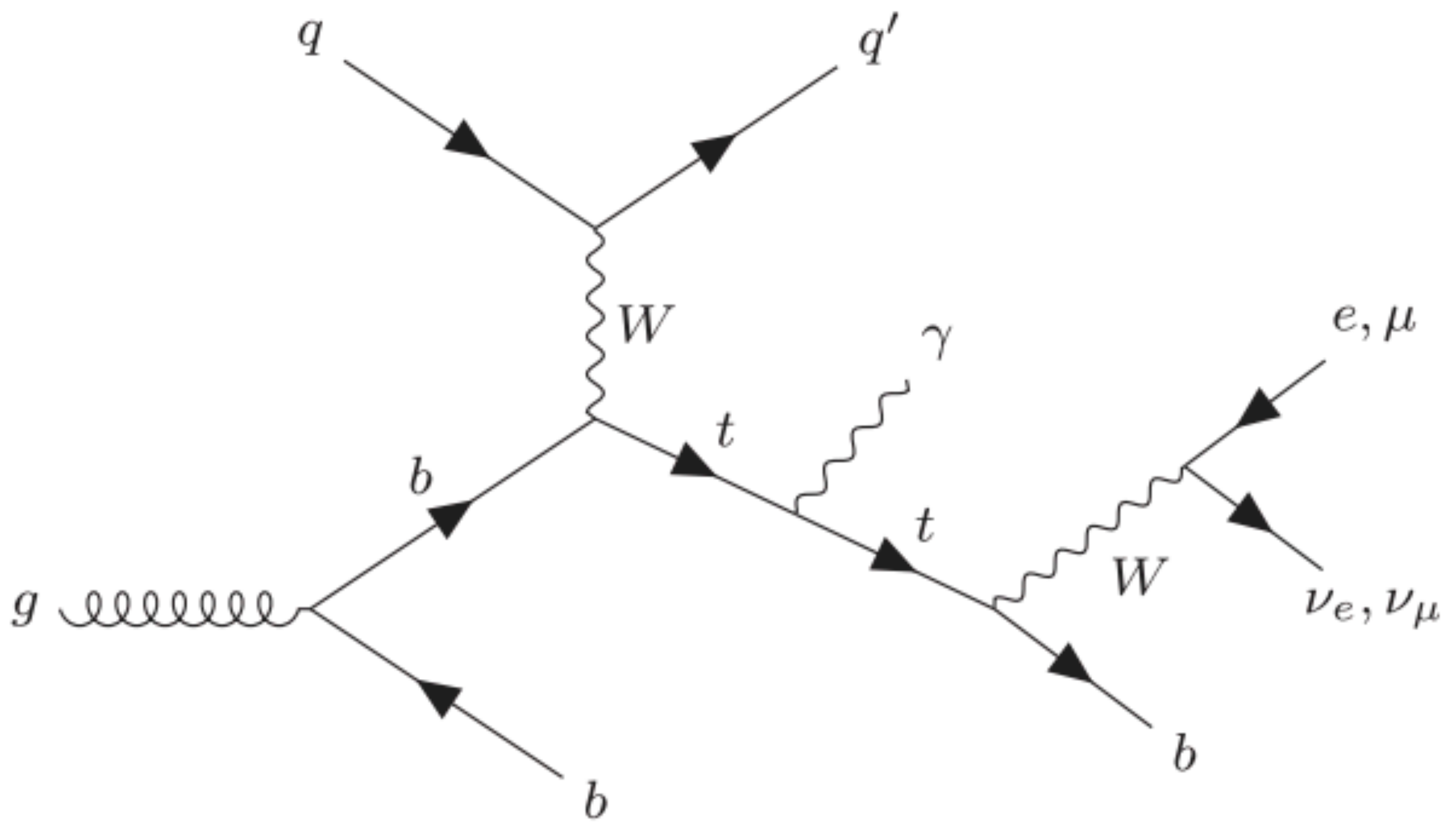
Goal and strategy

Separate signal and background

- Train **DNN** to separate tγq, ttγ, and others

Background estimation/constraint ($t\bar{t}\gamma$ as signal):

- Simulation: $t\bar{t}$, V+Jets/Vγ+Jets, tW/tWγ, TTV, VV
- Data-Driven backgrounds:
  - j→γ, j→ℓ, e→γ (mainly in e channel)
- Define proper control regions
  - Constrain main and data-driven background normalisations



Events of interest: **N_γ =1, N_ℓ=1, N_j≥2, N_b≥1**

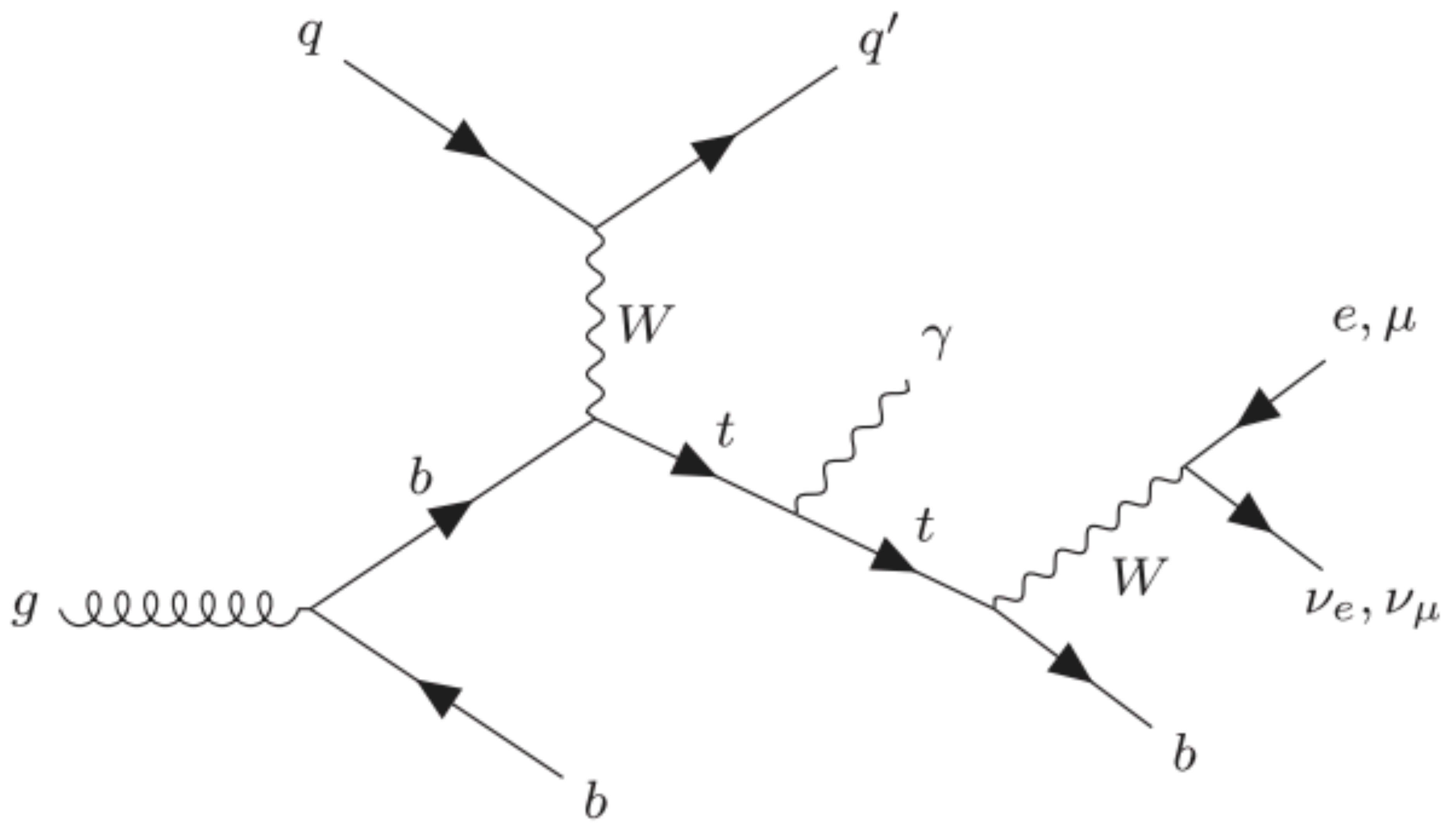
Goal and strategy

Separate signal and background

- Train **DNN** to separate tγq, ttγ, and others

Background estimation/constraint ($t\bar{t}\gamma$ as signal):

- Simulation: $t\bar{t}$, V+Jets/Vγ+Jets, tW/tWγ, TTV, VV
- Data-Driven backgrounds:
  - j→γ, j→ℓ, e→γ (mainly in e channel)
- Define proper control regions
  - Constrain main and data-driven background normalisations



Events of interest: **N_γ =1, N_ℓ=1, $N_j \geq 2$, $N_b \geq 1$**

- Inclusive/Differential xs for both tγq+$t\bar{t}\gamma$
- EFT interpretation of tγq+$t\bar{t}\gamma$

backup

Data and trigger

- Data: Full Run-II dataset
  - 2016 & 2017: SingleMuon, SingleElectron
  - 2018: SingleMuon, EGamma
- Triggers

Year	2016	2017	2018
Muon	IsoMu24 or IsoTkMu24	IsoMu27	IsoMu24
Electron	Ele27_WPTight_Gsf	Ele32_WPTight_Gsf_L1DoubleEG*	Ele32_WPTight_Gsf

Trigger SFs are ready for these HLT paths

✦ Analysis is based on the NanoAOD v9 UL campaign

✦ Results shown today are from 2018 data and MC

backup

Object selection

backup

Electron	Good
p_T/GeV	> 35
$\vert\eta\vert$	< 2.5 not in ECAL gap
ID	cut-based medium ID
Others	Impact (d_{xy},d_z)

Muon	Good
p_T/GeV	> 30
$\vert\eta\vert$	< 2.4
ID	cut-based tight ID
Iso	Tight Iso (<0.15)

Photon	Good
p_T/GeV	> 20
$\vert\eta\vert$	< 2.5 not in ECAL gap
ID	cut-based medium ID
Electron-veto	pixel seed veto

JetMET	Jet	b-jet	MET
p_T/GeV	> 30	> 30	> 20
$\vert\eta\vert$	< 4.7	< 2.5	—
Type	AK4CHS	AK4CHS	PFMET
ID	tight jet ID	medium deepjet ID	—

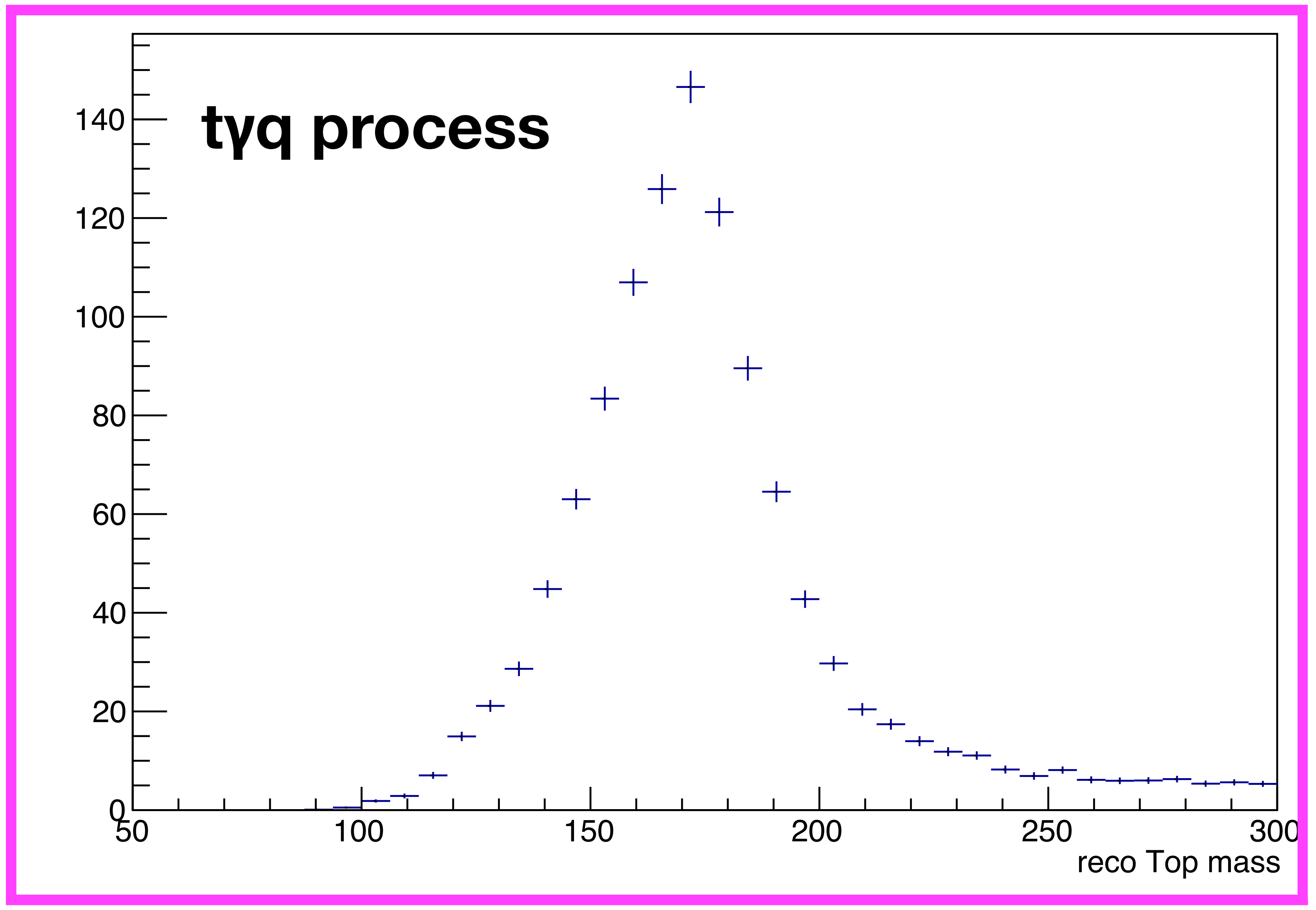
- Pileup reweighing
- Lepton energy correction
  - *muon Rochester and electron energy scale/smearing*
- Lepton ID/ISO/RECO/HLT scale factors
- e/γPhoton energy scale/smearing
- Photon ID/Pixel Seed Veto scale factors
- Jet energy correction
- ~~Jet pileup ID scale factors~~
- b-jet ID scale factors → to be updated

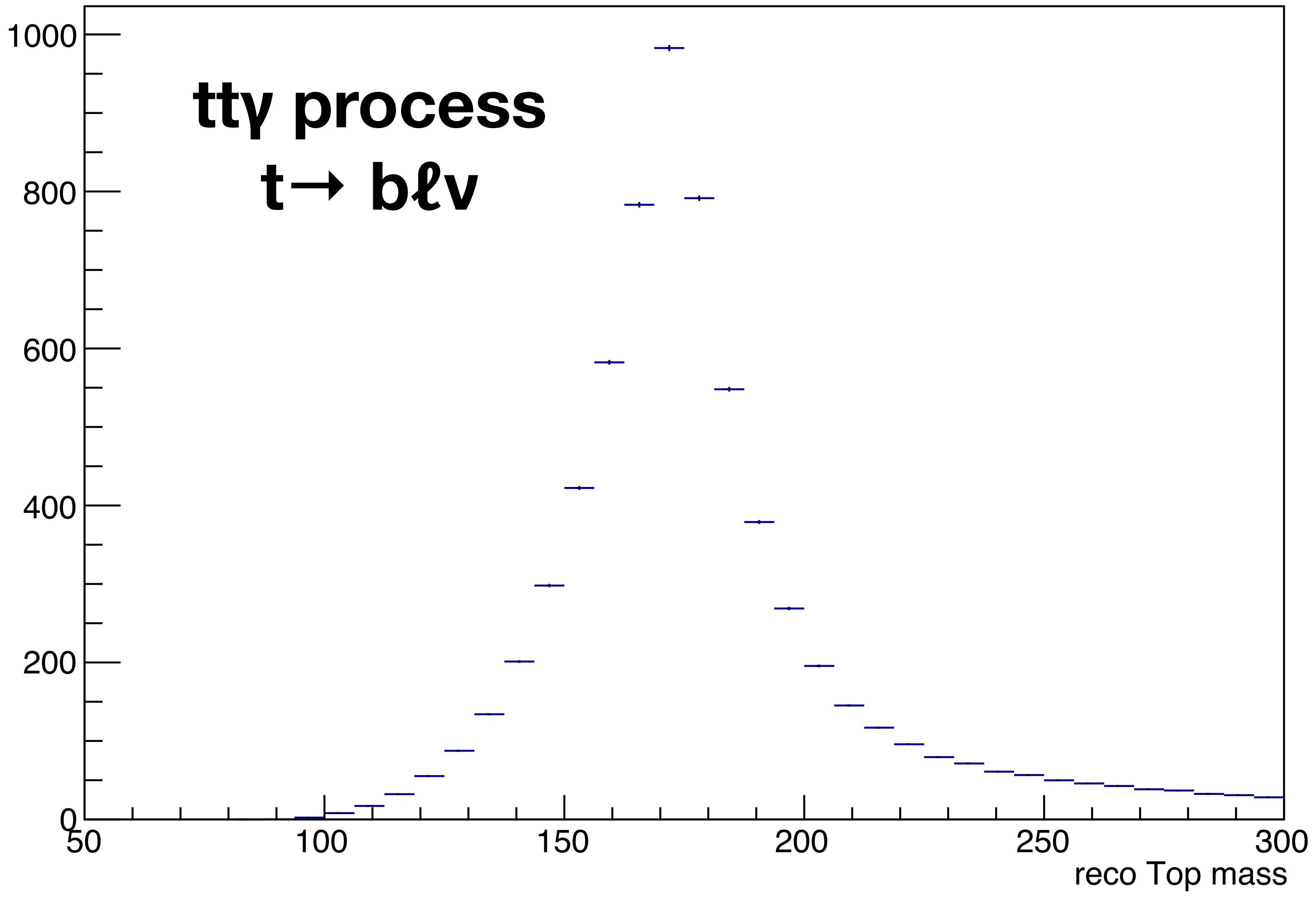
- Event ≥ 1 good PV and pass MET Filters and pass high-level trigger
- Exactly one lepton
  - Reject events containing extra ℓ with veto lepton requirement
- At least one photon
- At least two jet with one at least one being b-jet
- $\Delta R(\ell,\gamma) > 0.4, \Delta R(\ell,j) > 0.4, \Delta R(\gamma,j) > 0.4$
- MET p_T > 20 GeV

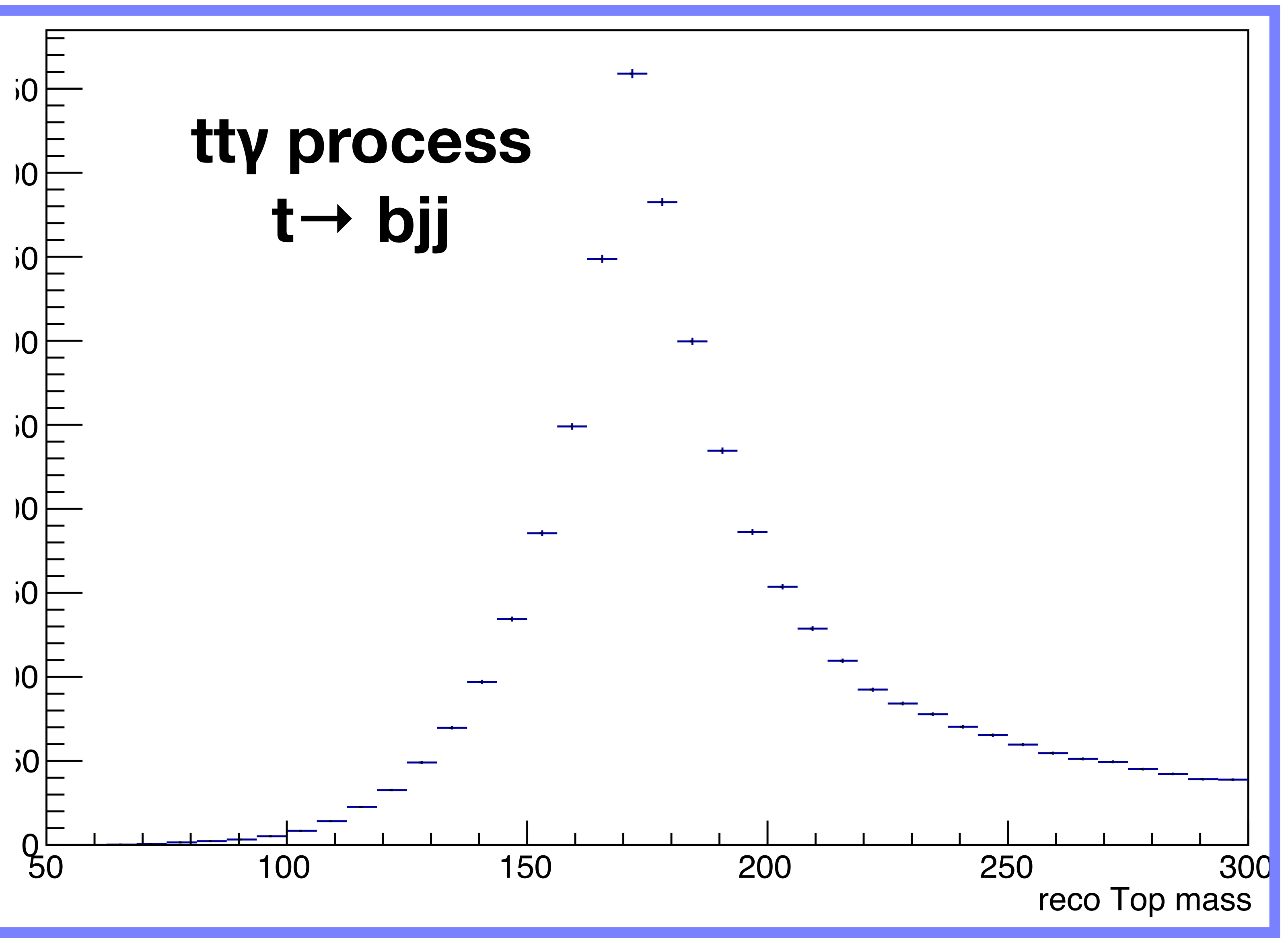
Top reconstruction

- Chi-square minimisation is performed
- Leptonic and hadronic top quarks are reconstructed depending on the the available objects (details in backup)
- If the reconstruction is not possible, give a default value -10

$$\chi^2_{\mathrm{t,lep}} = \left(\frac{m_{\ell\nu\mathrm{b}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,lep}}}\right)^2 \qquad \chi^2_{\mathrm{t,had}} = \left(\frac{m_{\mathrm{bjj}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,had}}}\right)^2 \qquad \chi^2_{\mathrm{t}} = \left(\frac{m_{\ell\nu\mathrm{b}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,lep}}}\right)^2 + \left(\frac{m_{\mathrm{bjj}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,had}}}\right)^2$$







Simulation — removal strategy

Due to photons added from hadronization and parton showering, there are phase space **overlap between** samples w/ and /w/o a Madgraph matrix element γ (e.g. $t\bar{t}\gamma$ /$t\bar{t}$, DY/Zγ) → take care of double-counting when running over samples **X+jets** and **Xγ+jets**

- Overlap definition : events contain at least one good genphoton*
- The good genphoton* is from NanoAOD *GenPart* with:
  - *|pdgId| = 22*
  - *status = 1* → stable particle
  - $p_T > 20$ GeV, $|\eta| < 2.5$
  - *isPrompt* → not from hadron, μ or τ decay → different for different production mode
  - *ΔR(γ, part.) > R0gamma* for every status=1 genparticle (not γ or ν)

Sample/ cut	LO ttγ/tt	NLO ttγ/tt	tγq/single top (t-channel)	Zγ/DY	Wγ/ W+jets	tWγ/tW (t-channel)
R0gamma	0.1	0.05	0.05	0.05	0.05	0.05
Remarks	—	γ from top or ISR	γ from top or ISR	—	—	—

Removal for X+jets and Xγ+jets

- Xγ+jets events have at least one genphoton* as above
- X+jets events don't have the genphoton* as above

backup

Signal simulation — *tγq*

tγq NLO production with 4-flavour scheme:

Process card

generate p p > t b~ j a $$ w+ w- [QCD] @0

add process p p > t~ b j a $$ w+ w- [QCD] @1

Madspind card

decay t > w+ b, w+ > ell+ vl

decay t~ > w- b~, w- > ell- vl~

run card

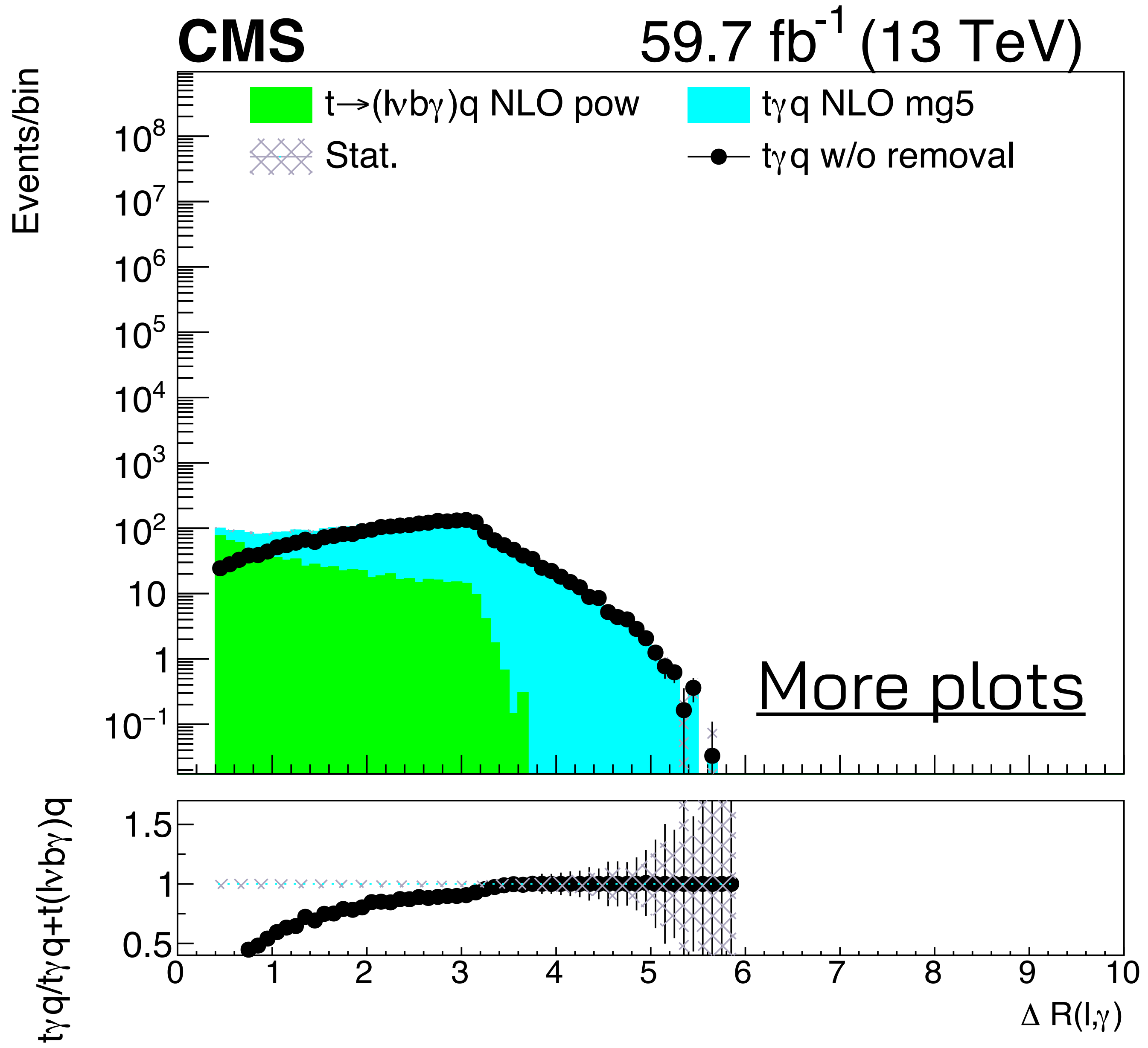
mll_sf = 30

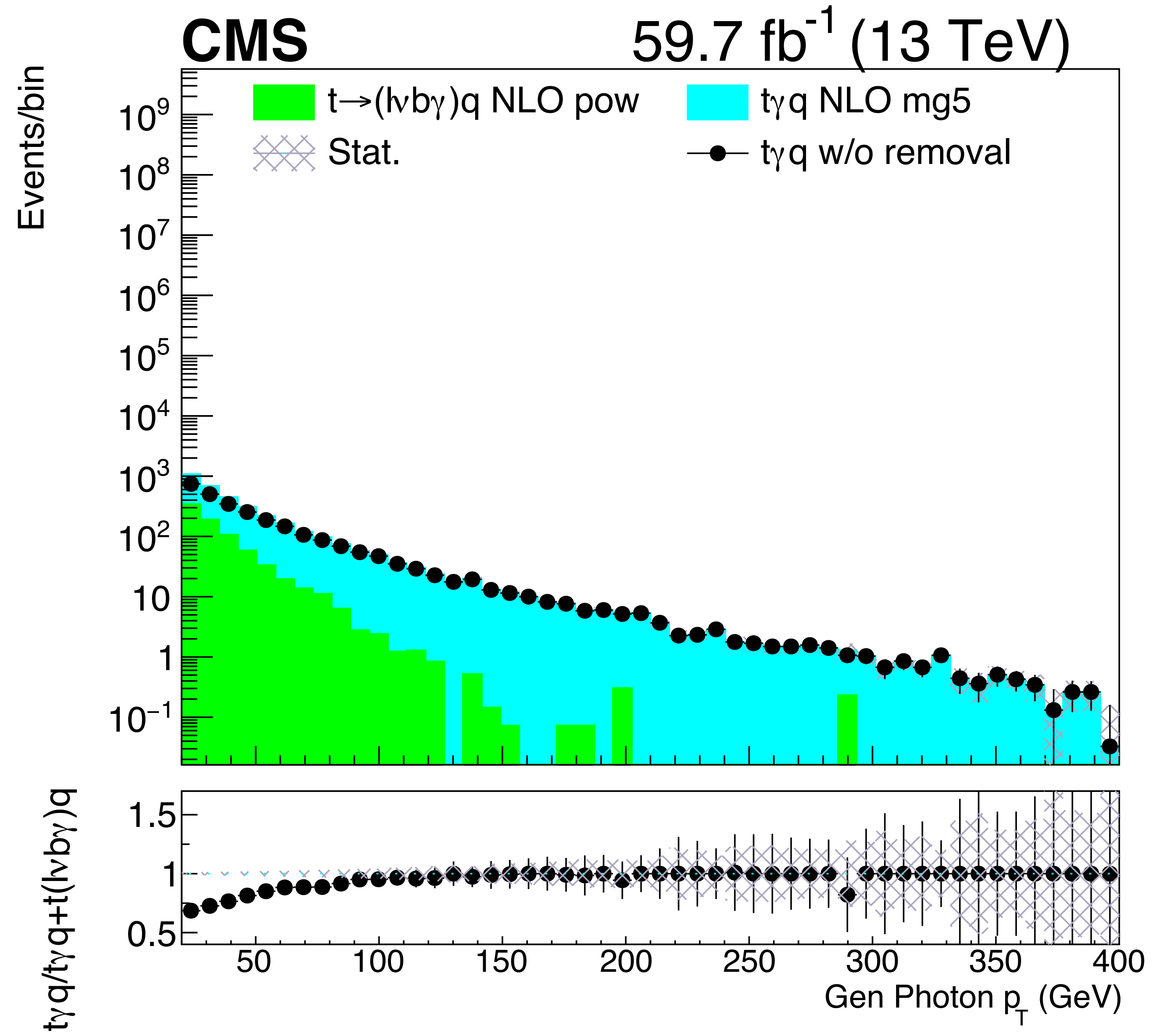
ptgmin = 10

R0gamma = 0.05

backup

- **Generator level** with requirements
- N_γ =1, N_ℓ=1, N_j≥2, N_b≥1 and pass removal requirement





Selection	gen-lepton	gen-photon	gen-Jet	gen-bJet
p_T/GeV	> 30	> 15	> 30	> 30
|η|	< 2.5	< 2.5	< 4.7	< 2.5
status	1	1	—	—
|pdgID|	13/11	22	—	—
Others	No meson mother	• No meson mother • Isolated • $\Delta R(\ell,\gamma) > 0.1$	• $\Delta R(\ell,j) > 0.4$ • $\Delta R(\ell,\gamma) > 0.1$	• |partonFlavour| = 5 • $\Delta R(\ell,j) > 0.4$ • $\Delta R(\ell,\gamma) > 0.1$

The loss due to incomplete simulation in tγq NLO sample is not large which accounts for ≈25%, mainly concentrating on low photon p_T and low $\Delta R(\ell,\gamma)$

(Signal) simulation — $t\bar{t}\gamma$

$t\bar{t}\gamma$ LO production with 5-flavour scheme:

Process card

generate p p > t t~ > l+ vl b ds uc~ b~ a
add process p p > t t~ > uc ds~ b l- vl~ b~ a

$t\bar{t}\gamma$ NLO production with 5-flavour scheme:

Process card

generate p p > t t~ a [QCD] @0
add process p p > t t~ a j [QCD] @1
output TTGJets_5f_NLO_FXFX -nojpeg

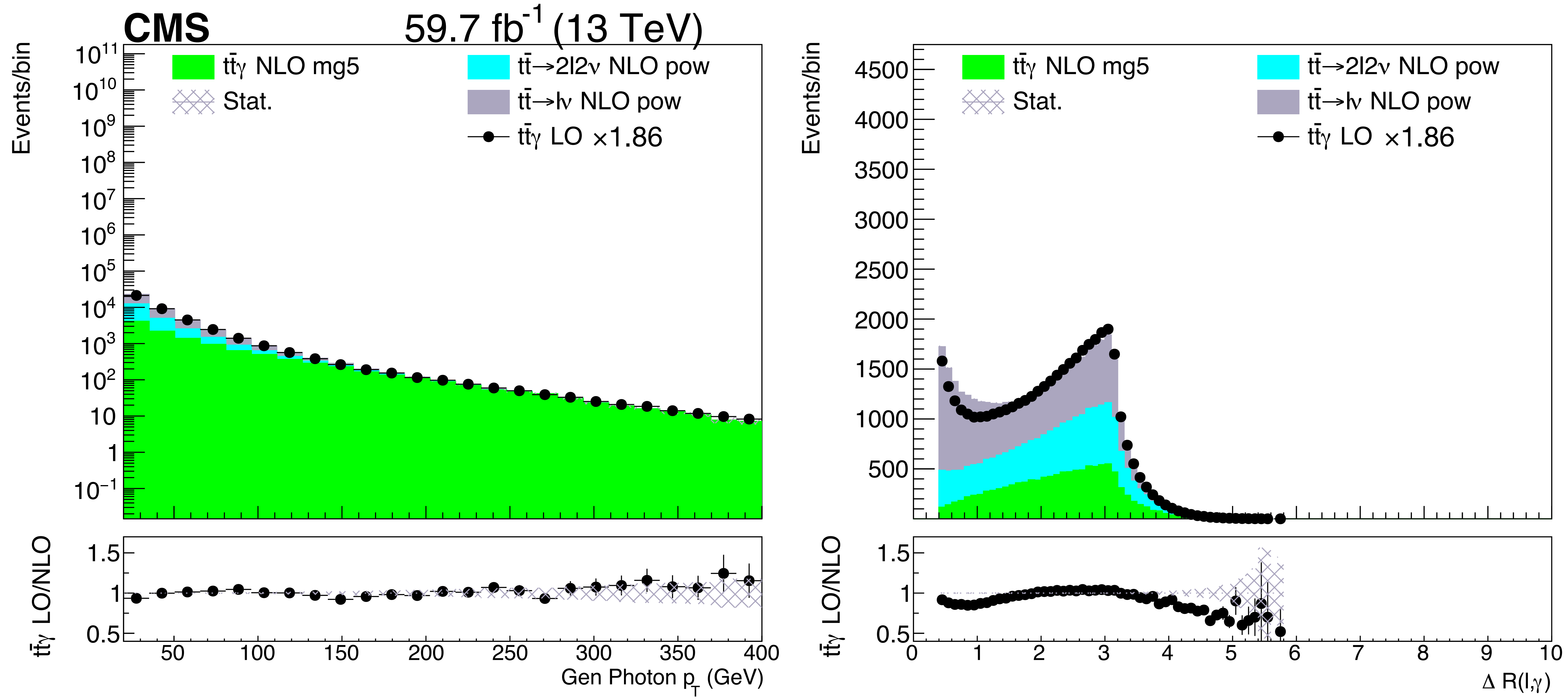
Madspind card

decay t > w+ b, w+ > all all
decay t~ > w- b~, w- > all all

k-factor for $t\bar{t}\gamma$ LO : 1.86

backup

- **Generator level** with requirements
- N_γ =1, N_ℓ=1, $N_j \geq 2$, $N_b \geq 1$ and pass removal requirement



For LO $t\bar{t}\gamma$:

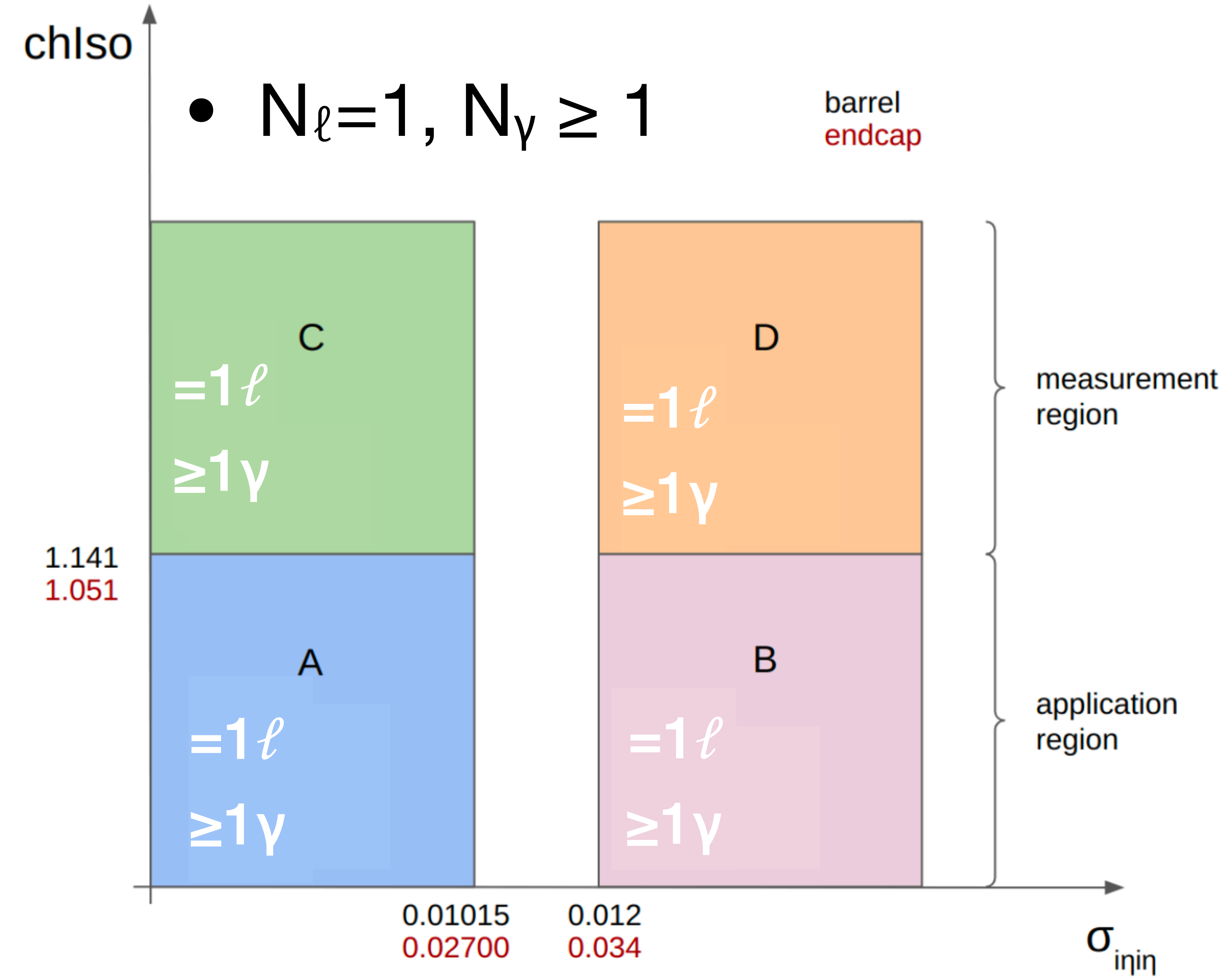
- A large value of the k-factor is needed → Important to do dependency studying in differential xs measurement

For NLO $t\bar{t}\gamma$:

- Loss due to the incomplete simulation is large. Need to consider treatment of the contribution from the $t\bar{t}$

Background estimation — Nonprompt γ

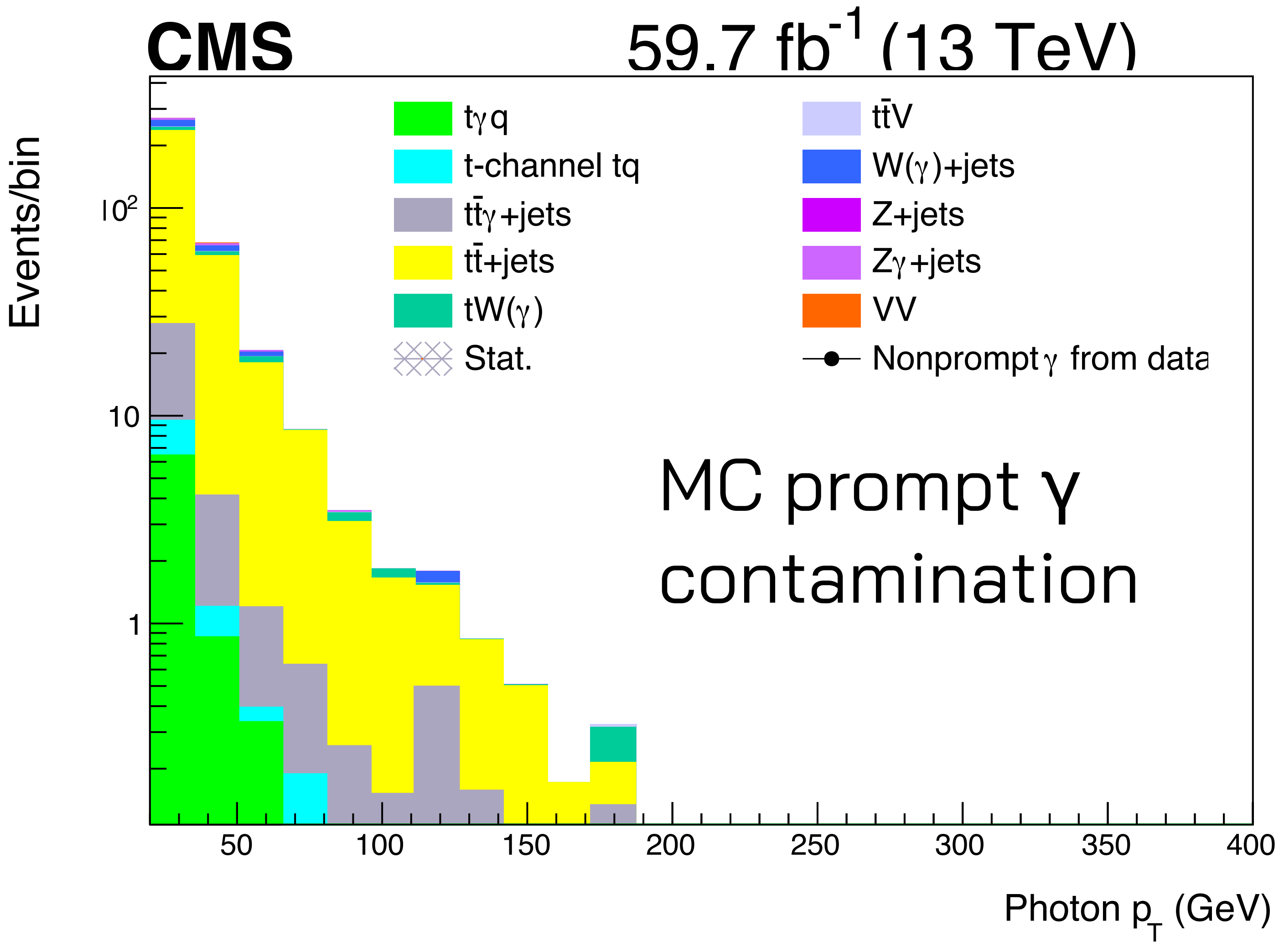
- $N_\ell = 1$, $N_\gamma \geq 1$

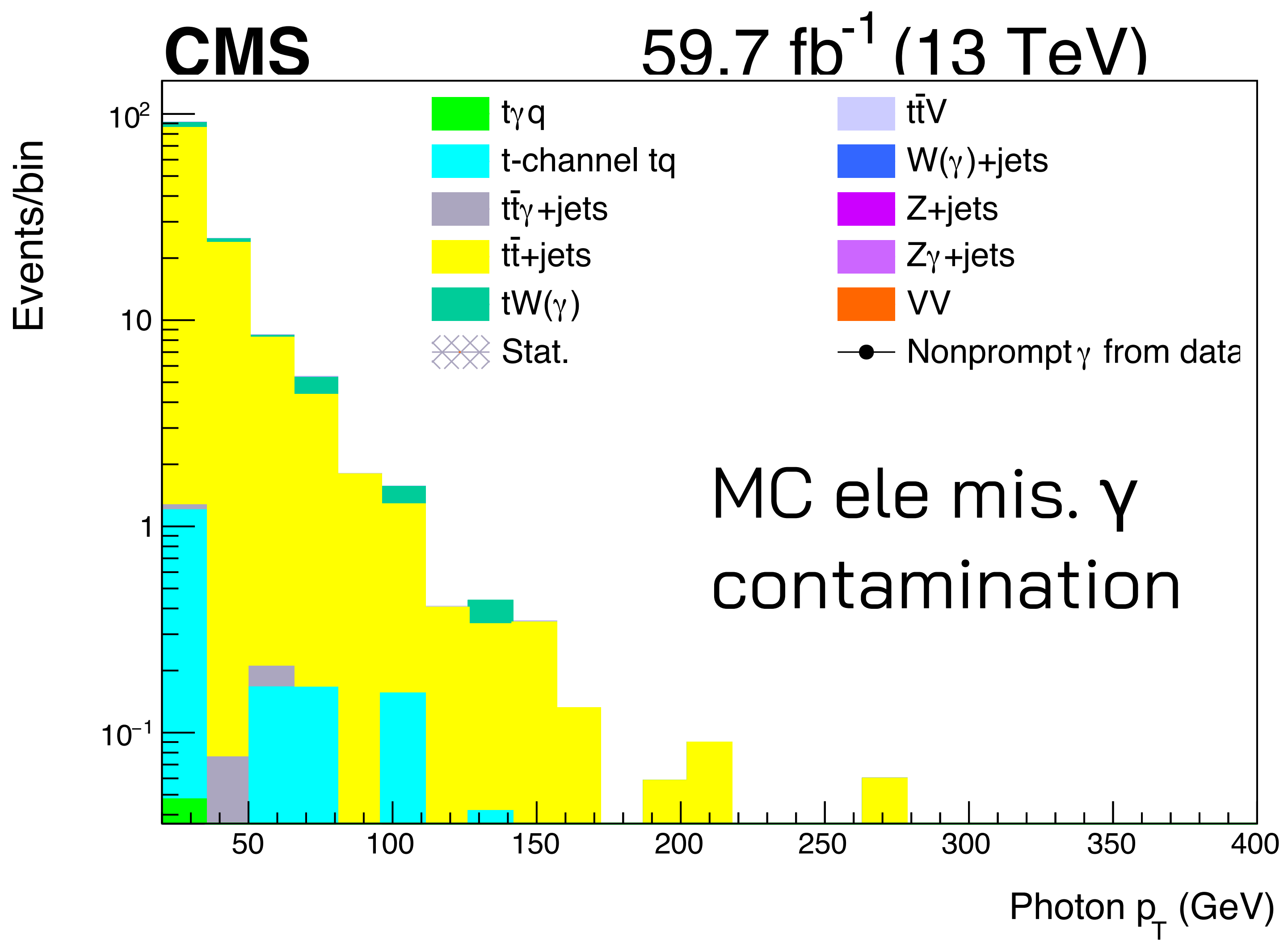


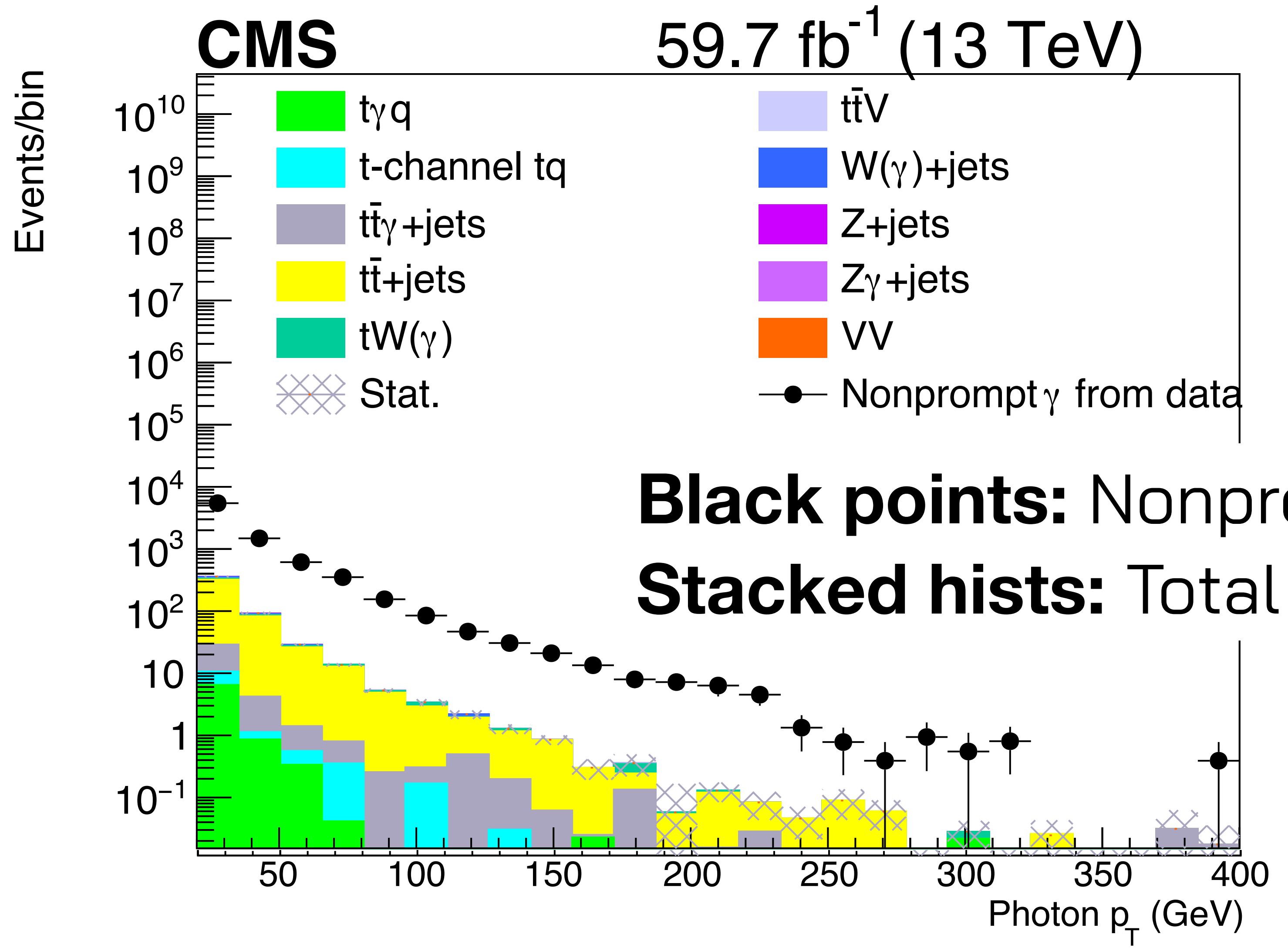
$$\text{fake rate}^{ij} = \frac{\text{Data}^{ij}_{C} - (\text{prompt} + \text{ele mis.})\,\text{MC}^{ij}_{C}}{\text{Data}^{ij}_{D} - (\text{prompt} + \text{ele mis.})\,\text{MC}^{ij}_{D}}$$

$$k^{ij}_{\text{MC}} = \frac{\text{nonprompt MC}^{ij}_{A}}{\text{nonprompt MC}^{ij}_{B}} \div \frac{\text{nonprompt MC}^{ij}_{C}}{\text{nonprompt MC}^{ij}_{D}}$$

$$\text{nonprompt contribution} = \sum_{ij} \left(\text{data}^{ij}_{B} \times \text{fake rate}^{ij} \times k^{ij}_{\text{MC}}\right) - \sum_{ij} \left((\text{prompt} + \text{ele mis.})\,\text{MC}^{ij}_{B} \times \text{fake rate}^{ij} \times k^{ij}_{\text{MC}}\right)$$

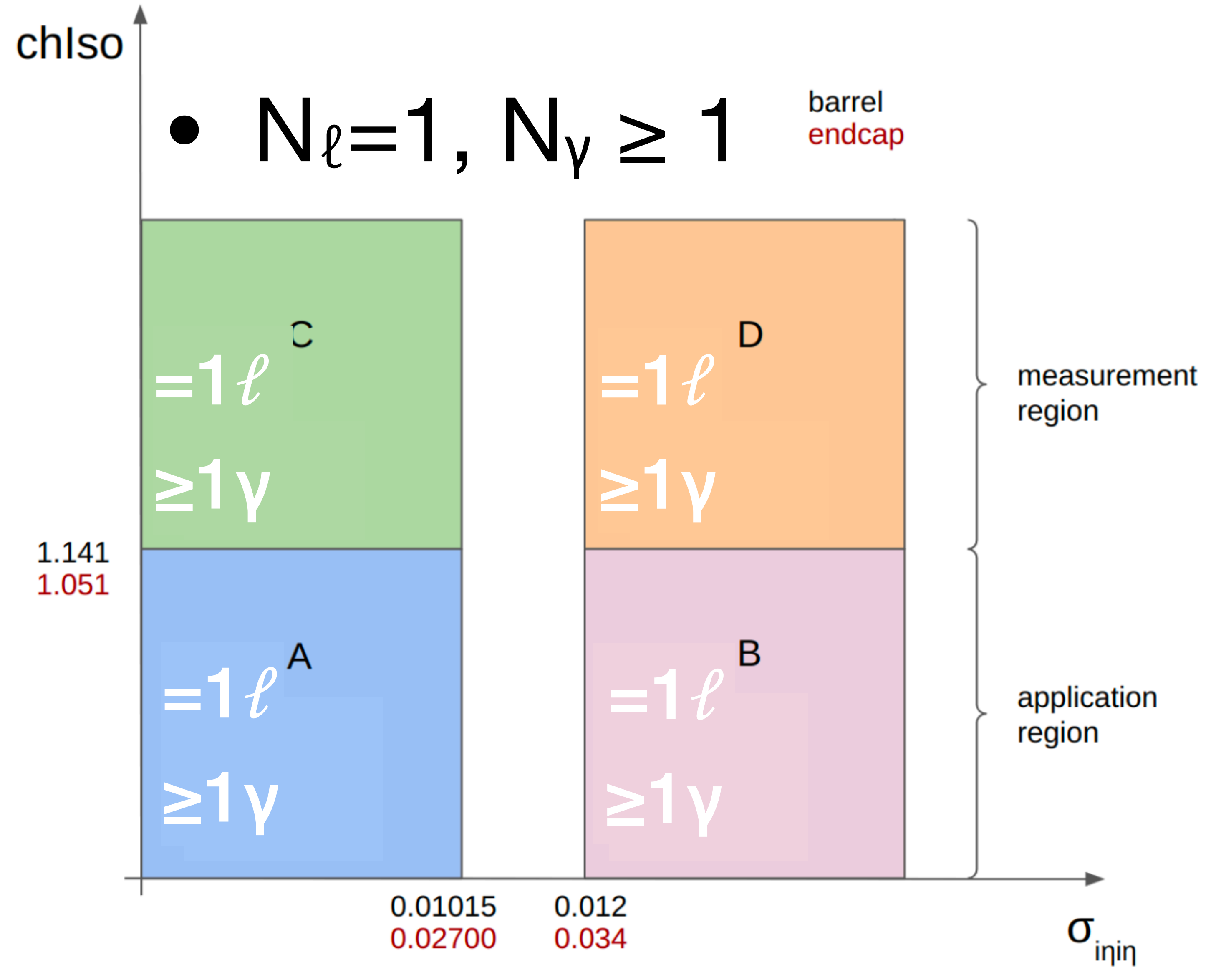






Black points: Nonprompt γ from data
Stacked hists: Total MC contamination

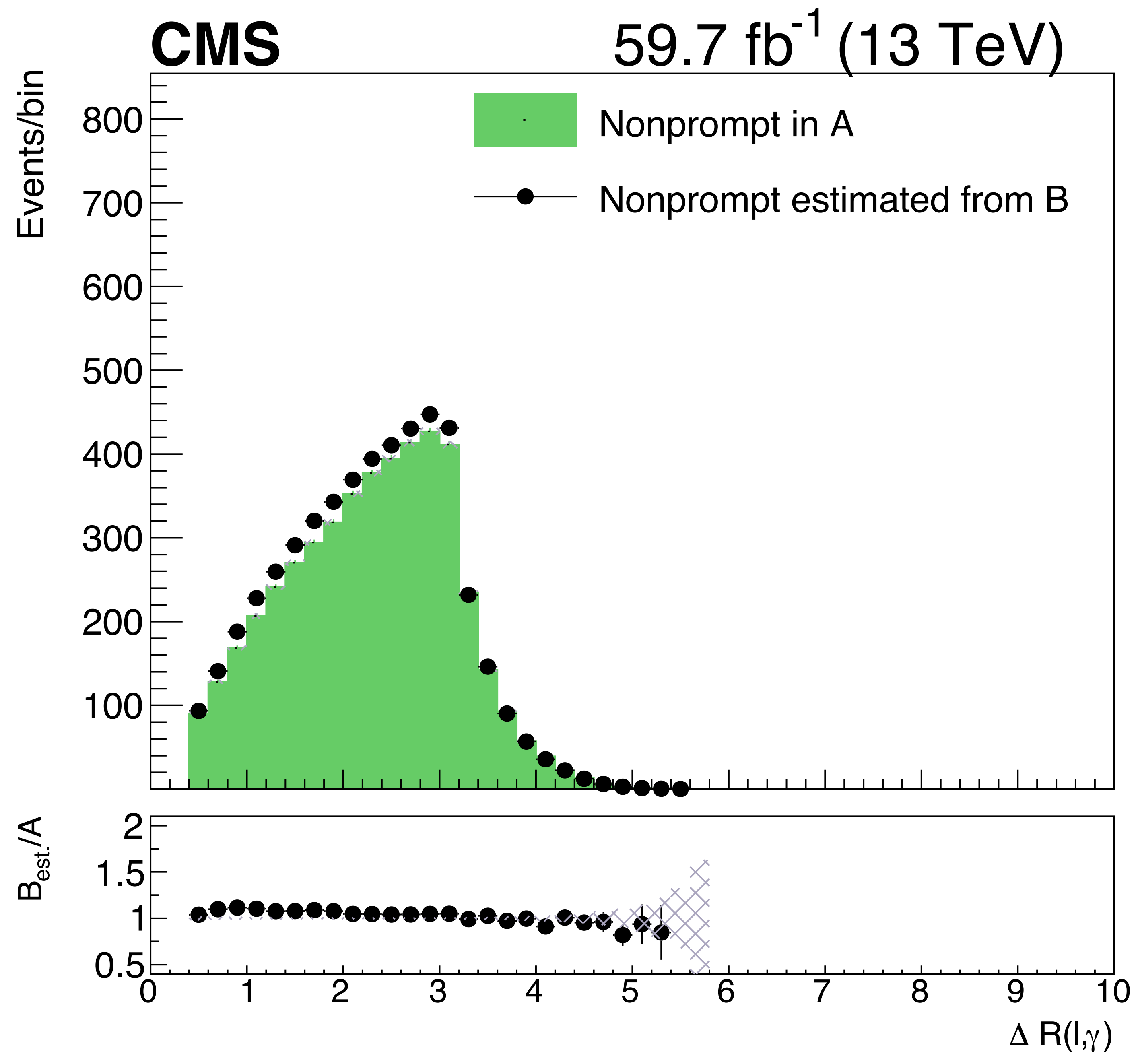
Nonprompt γ estimation — closure

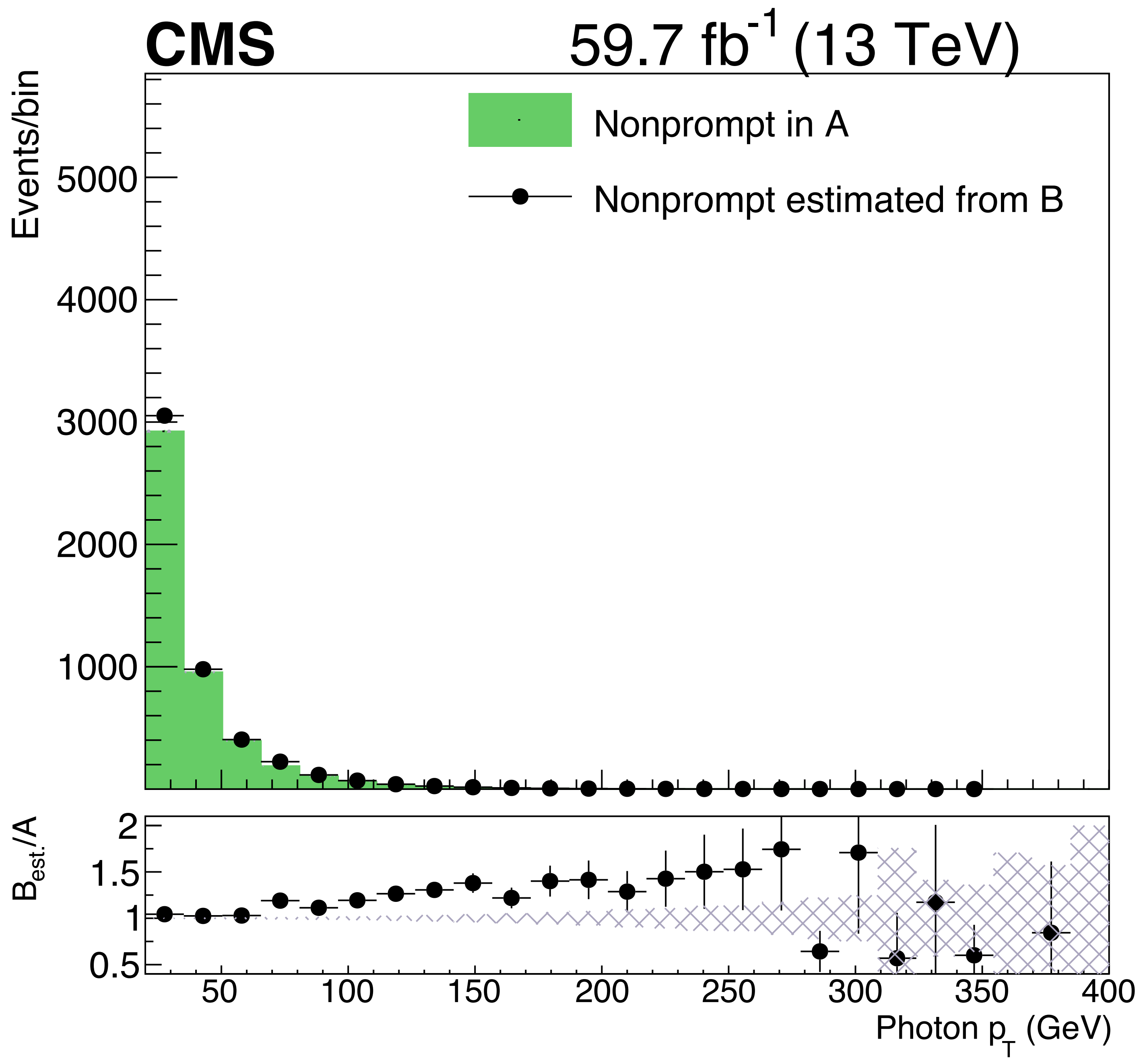


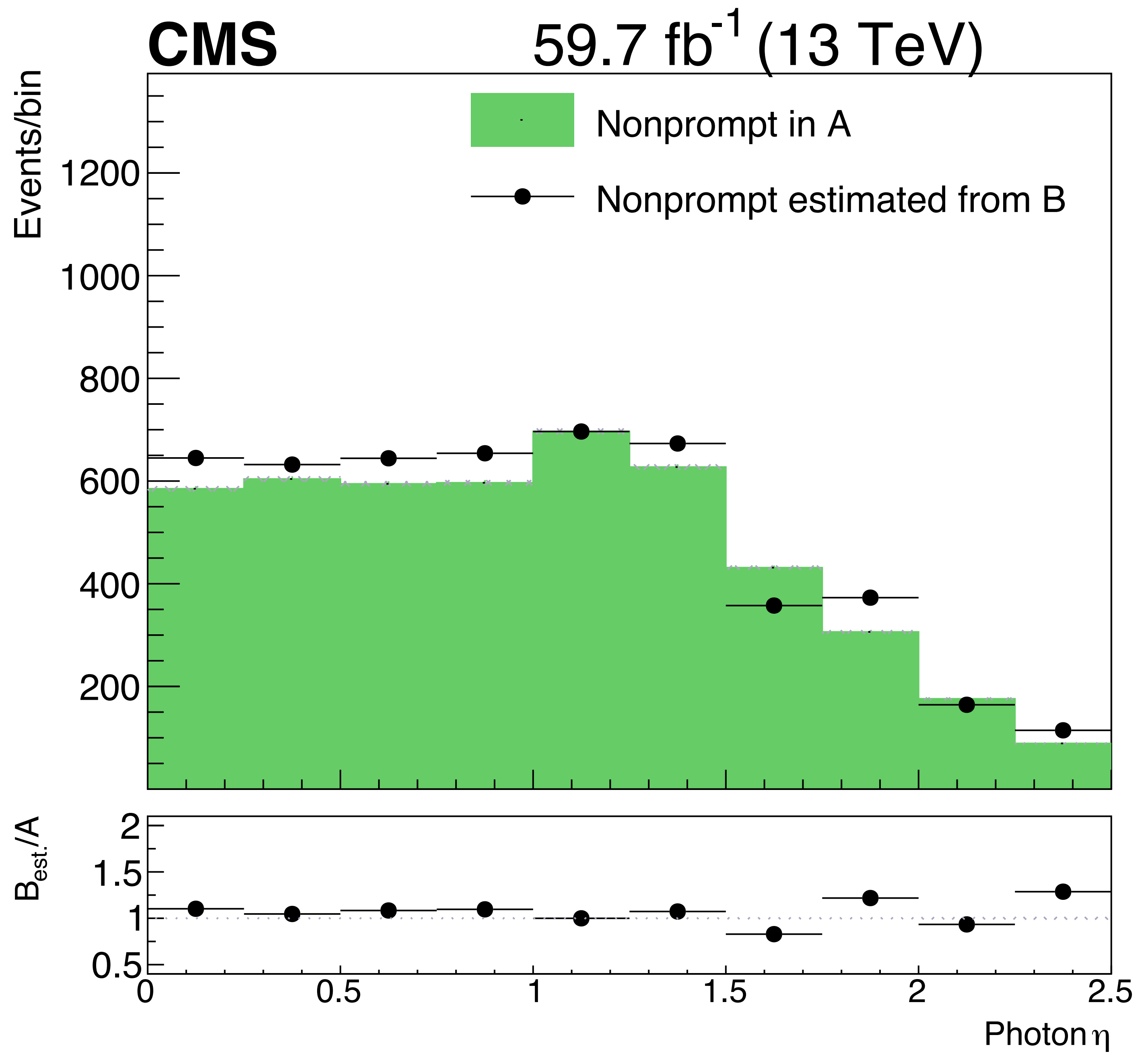
$$\text{fake rate}^{ij}_{t\bar{t}} = \frac{\text{nonprompt } t\bar{t}^{ij}_{C}}{\text{nonprompt } t\bar{t}^{ij}_{D}}$$

$$\text{nonprompt contribution}^{t\bar{t}}_{A} = \sum_{ij} (t\bar{t}^{ij}_{B} \times \text{fake rate}^{ij}_{t\bar{t}} \times k^{ij}_{t\bar{t}})$$

Within 20% uncertainty, agreement is acceptable



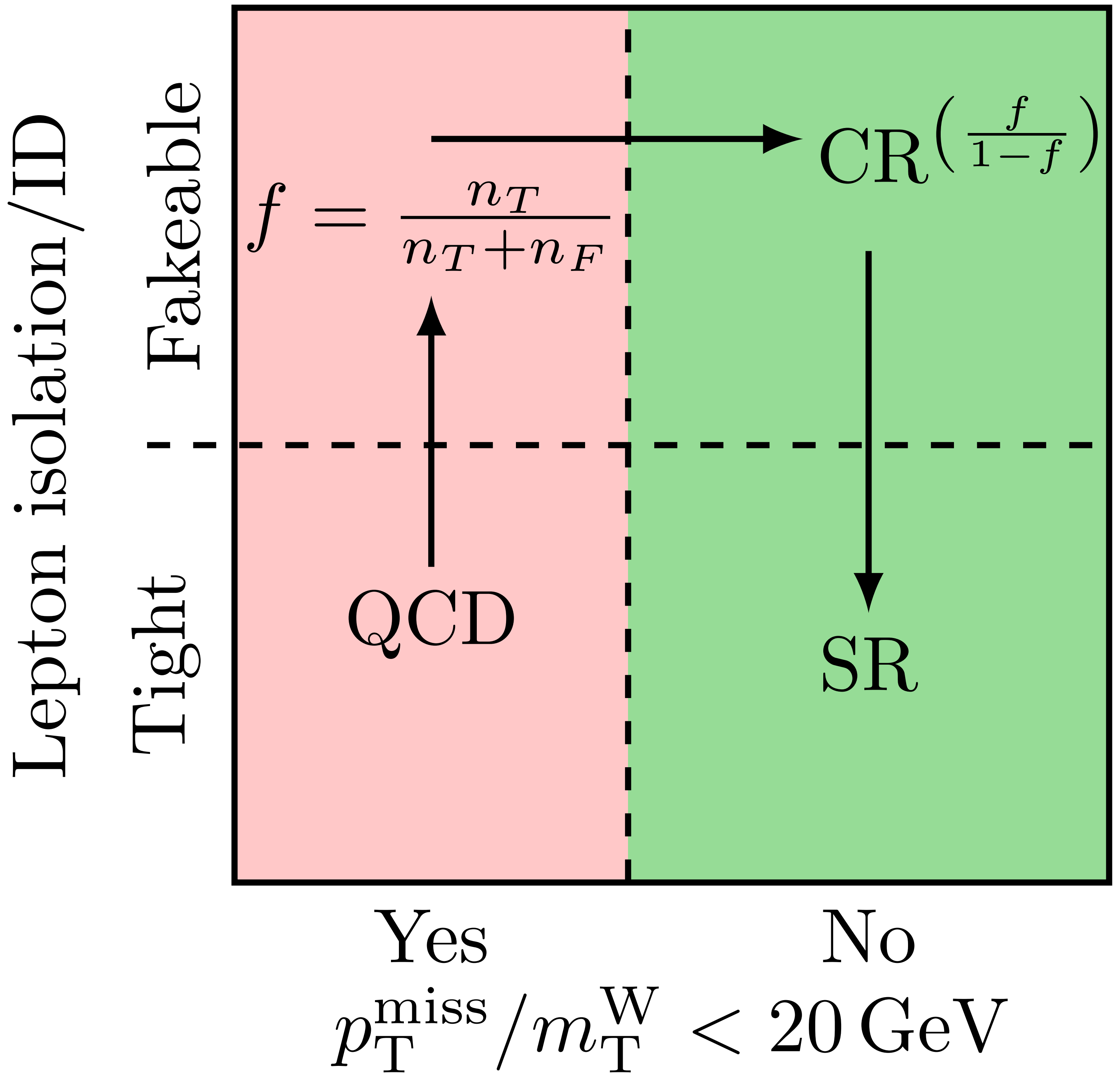


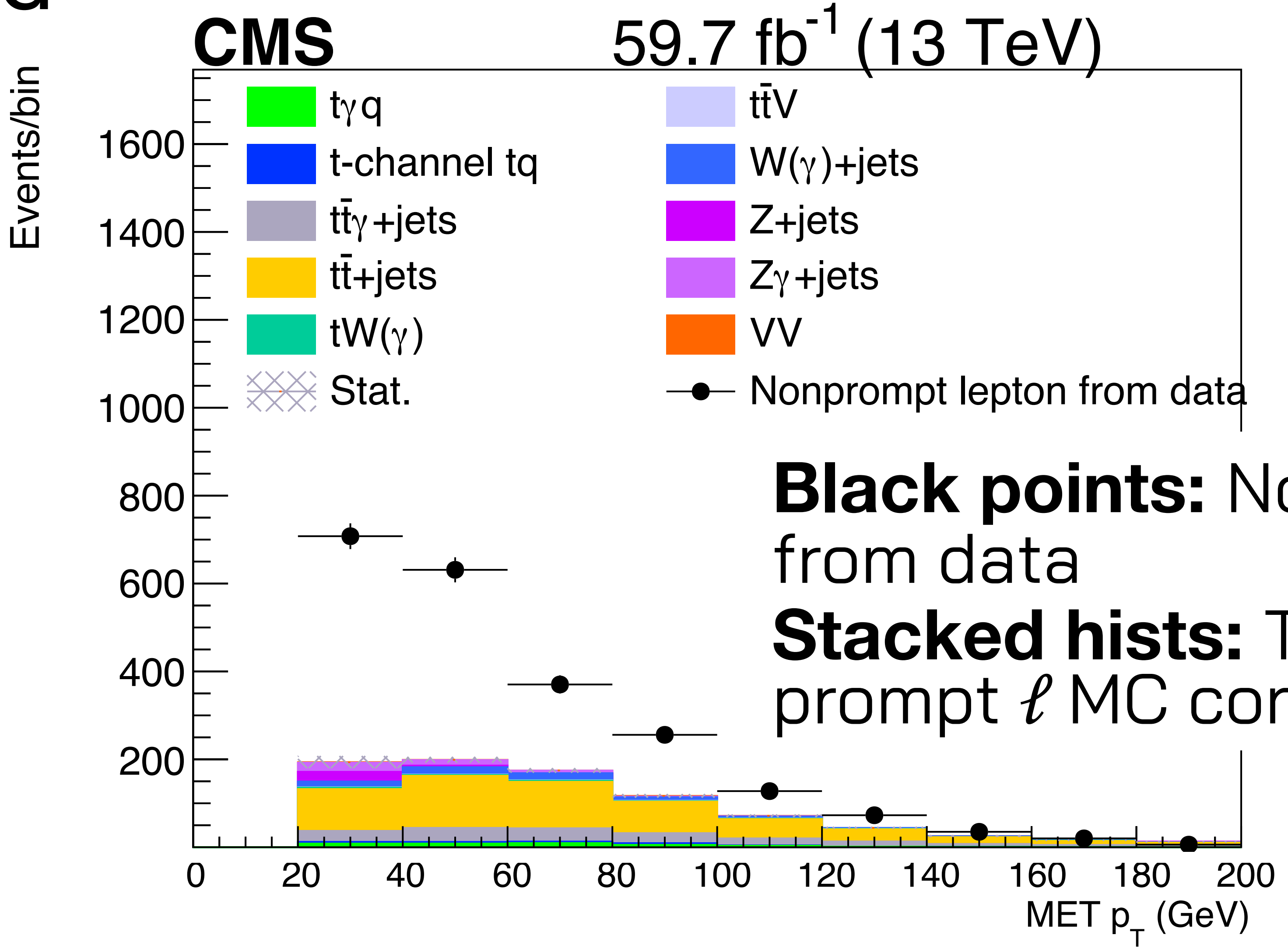


Background estimation — Nonprompt ℓ

1. Build QCD jet-enriched region with requirements of
   - Exactly one lepton
   - $p_T^{miss} < 20$ GeV and $m_T^W < 20$ GeV ($m_T^W = \sqrt{2p_T^{\ell} p_T^{miss}[1 - \cos\Delta\phi(\ell, \vec{p}_T^{miss})]}$)
   - At least one jet with $p_T > 30$ GeV and $\Delta R(\ell, j) > 0.4$
2. Measure the tight-to-loose rate $f = \frac{n_T}{n_T + n_F}$
   - n_T the number of leptons passing tight ℓ ID in QCD jet-enrich region
   - n_F the number of leptons passing fakeable ℓ ID in QCD jet-enrich region
3. Build nonprompt ℓ data-driven CR with fakeable ℓ ID and applied to SR with weights $f/(1-f)$

$$n^{SR}_{\text{nonprompt}\,\ell} = \sum_{ij} (\text{data}^{ij}_{CR} \times \frac{f^{ij}}{1-f^{ij}}) - \sum_{ij} (\text{prompt}\,\ell\,\text{MC}^{ij}_{CR} \times \frac{f^{ij}}{1-f^{ij}})$$

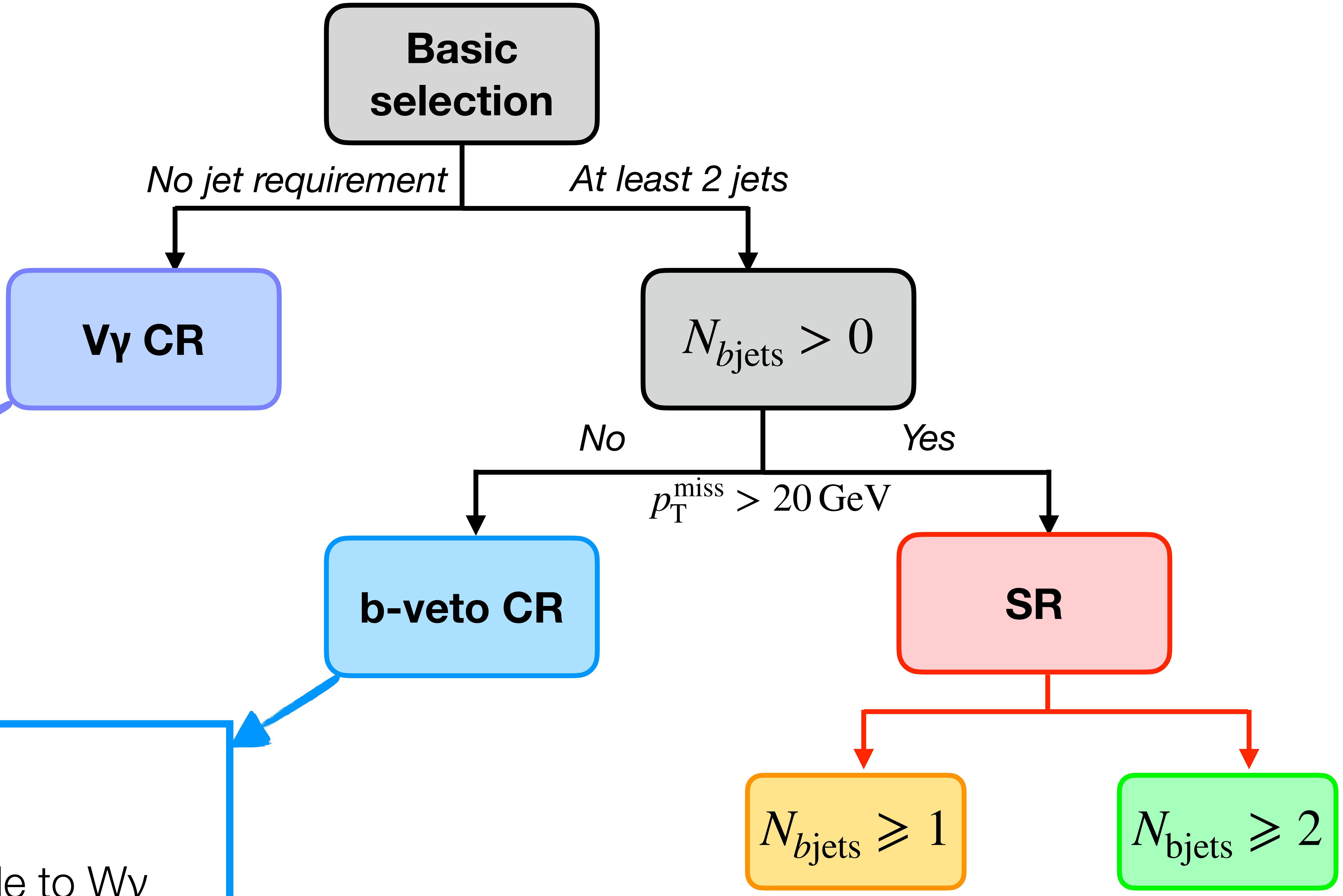




Black points: Nonprompt ℓ from data
Stacked hists: Total prompt ℓ MC contamination

Event categorisation

Baseline selection	A. Exactly one lepton with $p_T^{\mu/e}$> 30 (35) GeV, $\lvert\eta\rvert$ < 2.4 (2.5) B. Pass single e(μ) HLT paths C. At least one photon with p_T^{γ}> 20 GeV in the barrel or endcap



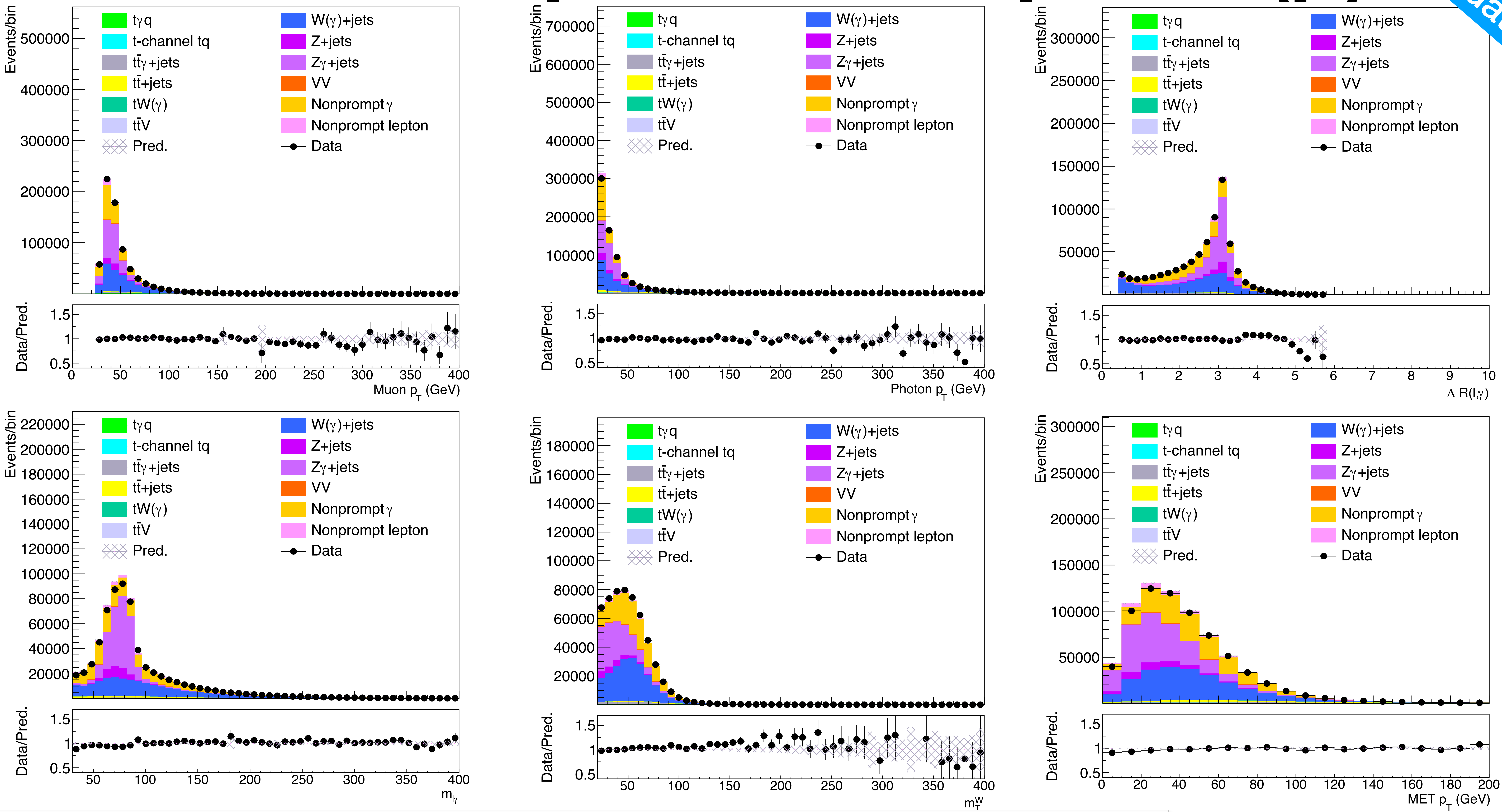
Vγ enriched region
- Nonprompt γ contribution is comparable
- Very large statistics is helpful to constrain corresponding background normalisation

Wγ enriched region
- Nonprompt ℓ contribution is visible
- Nonprompt γ contribution is comparable to Wγ
- Constrain the corresponding background normalisation by correlating the same systematic uncertainty in the SR

Events are further separated to **e channel** and **μ channel**

Control plots — Vγ CR (μ)

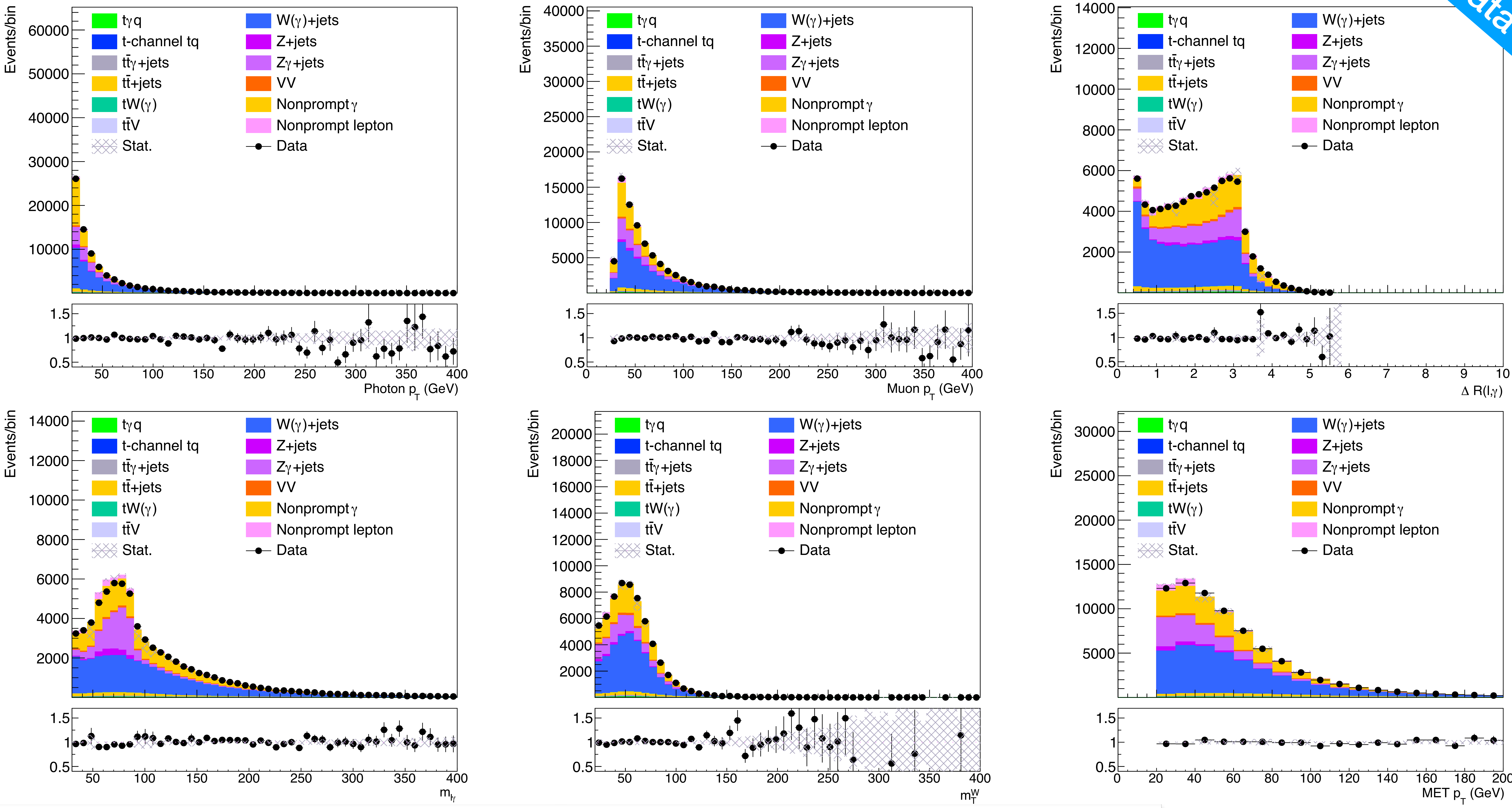
2018 data and MC



Use tt̄γ NLO. The tt̄γ LO also gives good agreement, since this is a Vγ dominant region

Control plots — b-veto CR (μ)

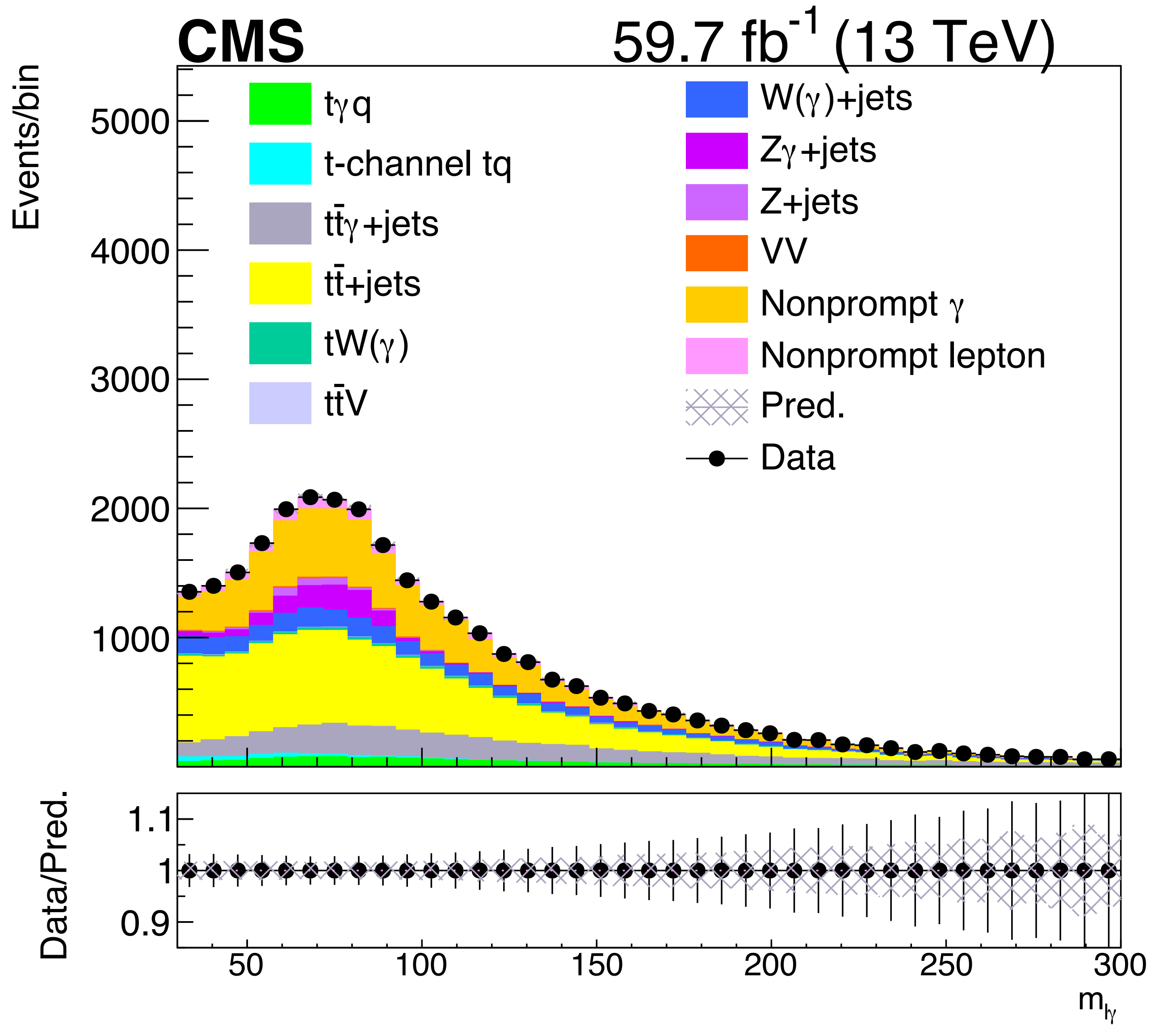
2018 data and MC

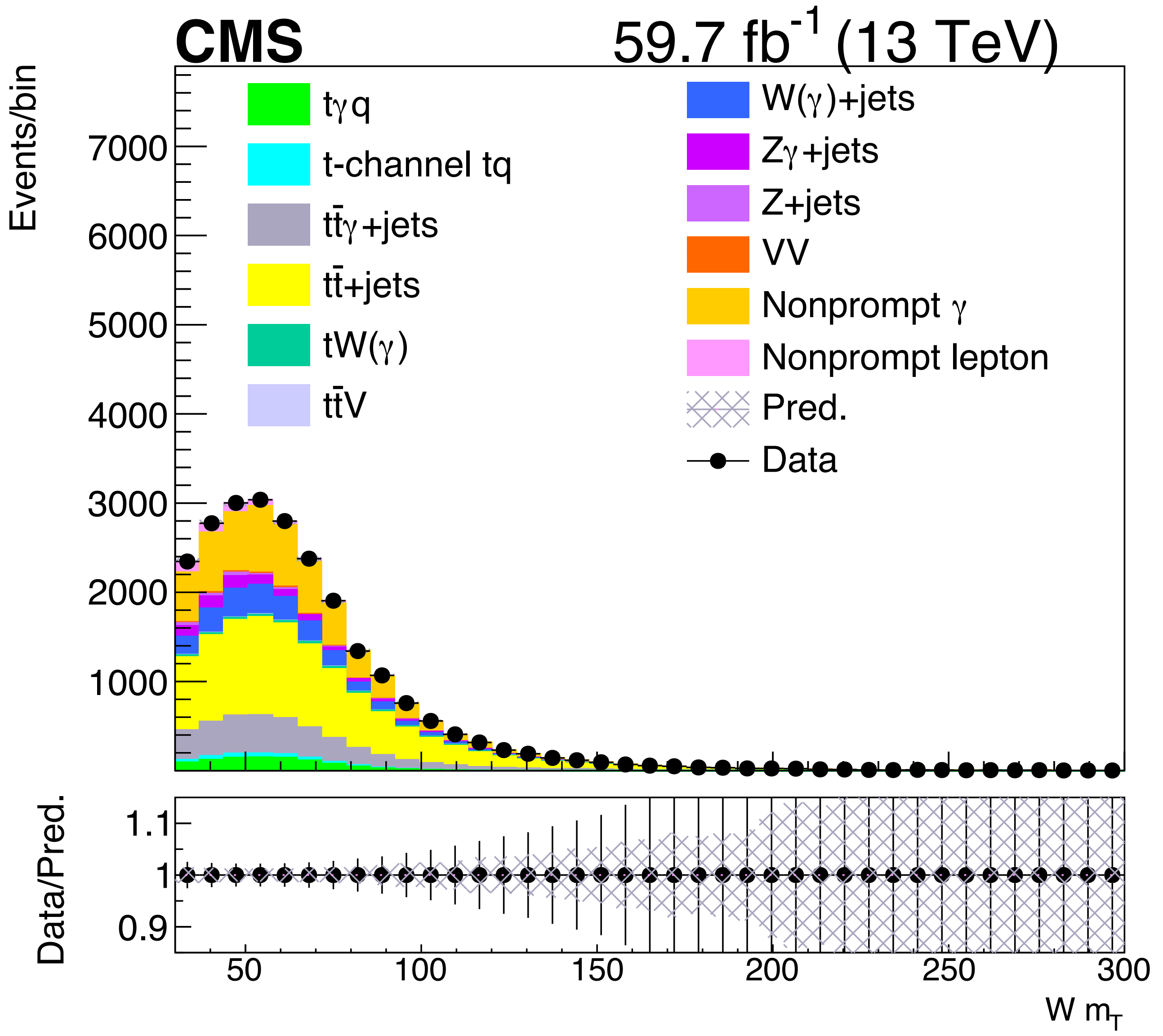


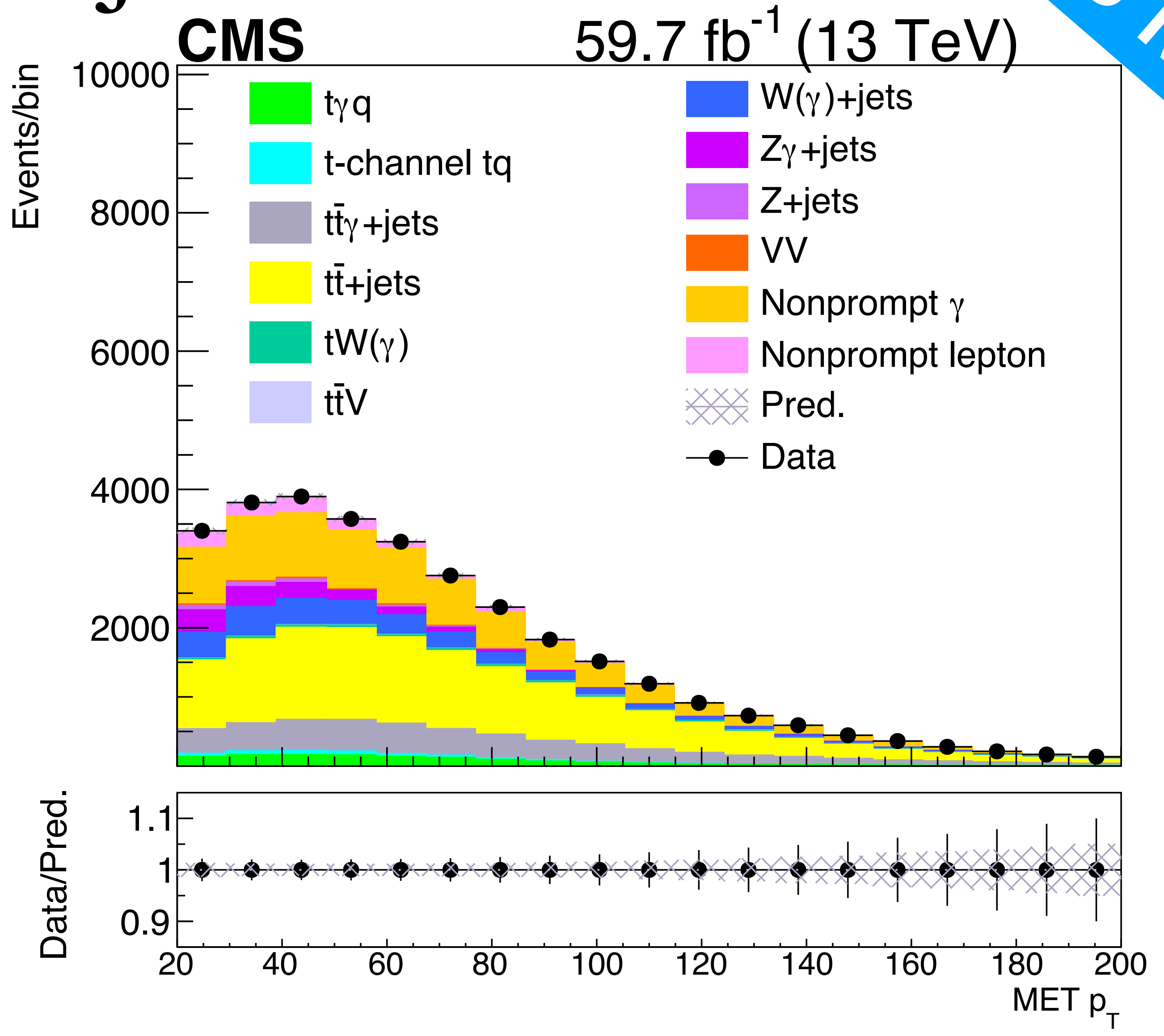
Use tt̄γ NLO. The tt̄γ LO also gives good agreement, since this is a Vγ dominant region

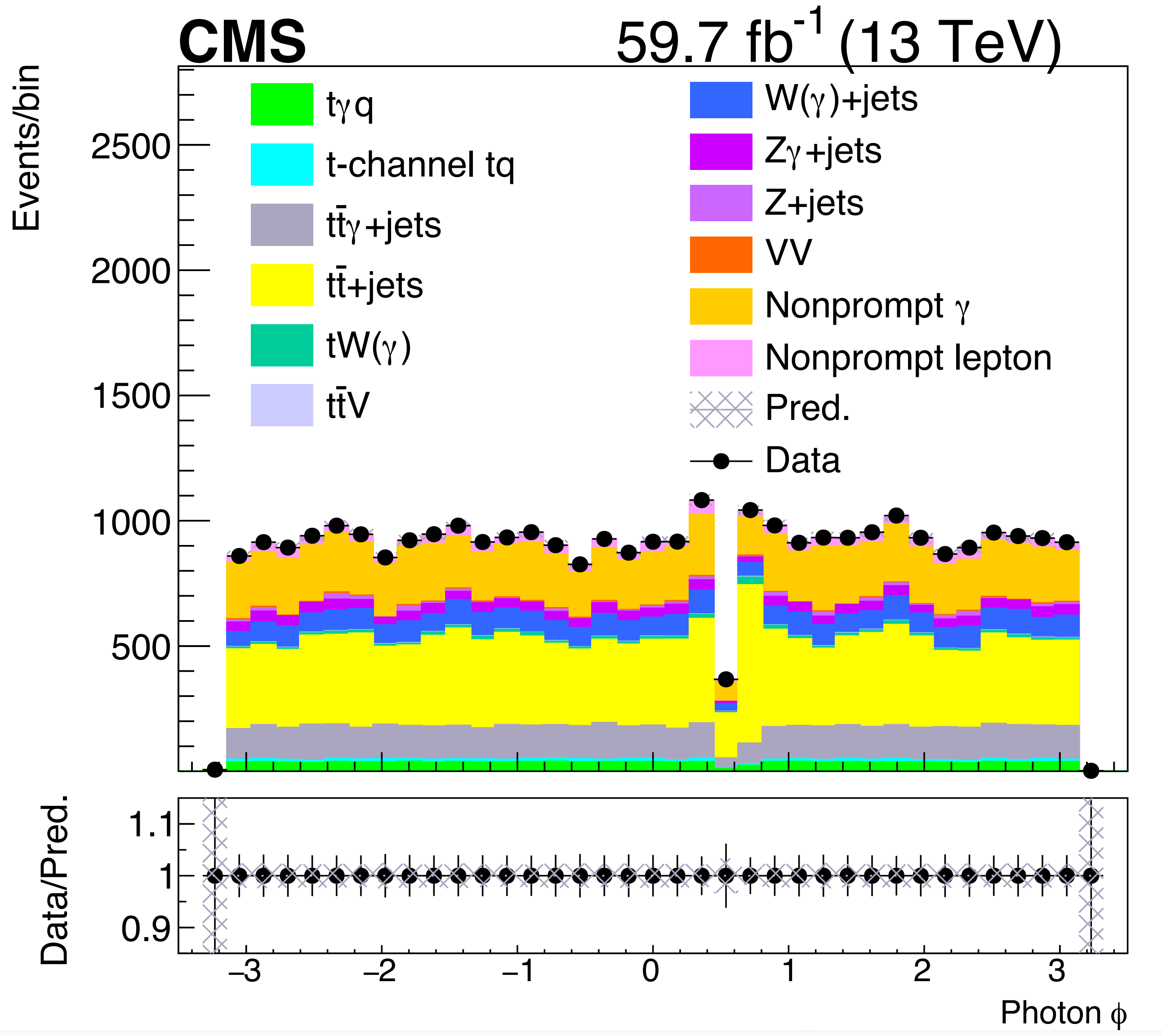
SR plots — $N_{\mathrm{j}} \geq 2$ $N_{\mathrm{b-jets}} \geq 1$

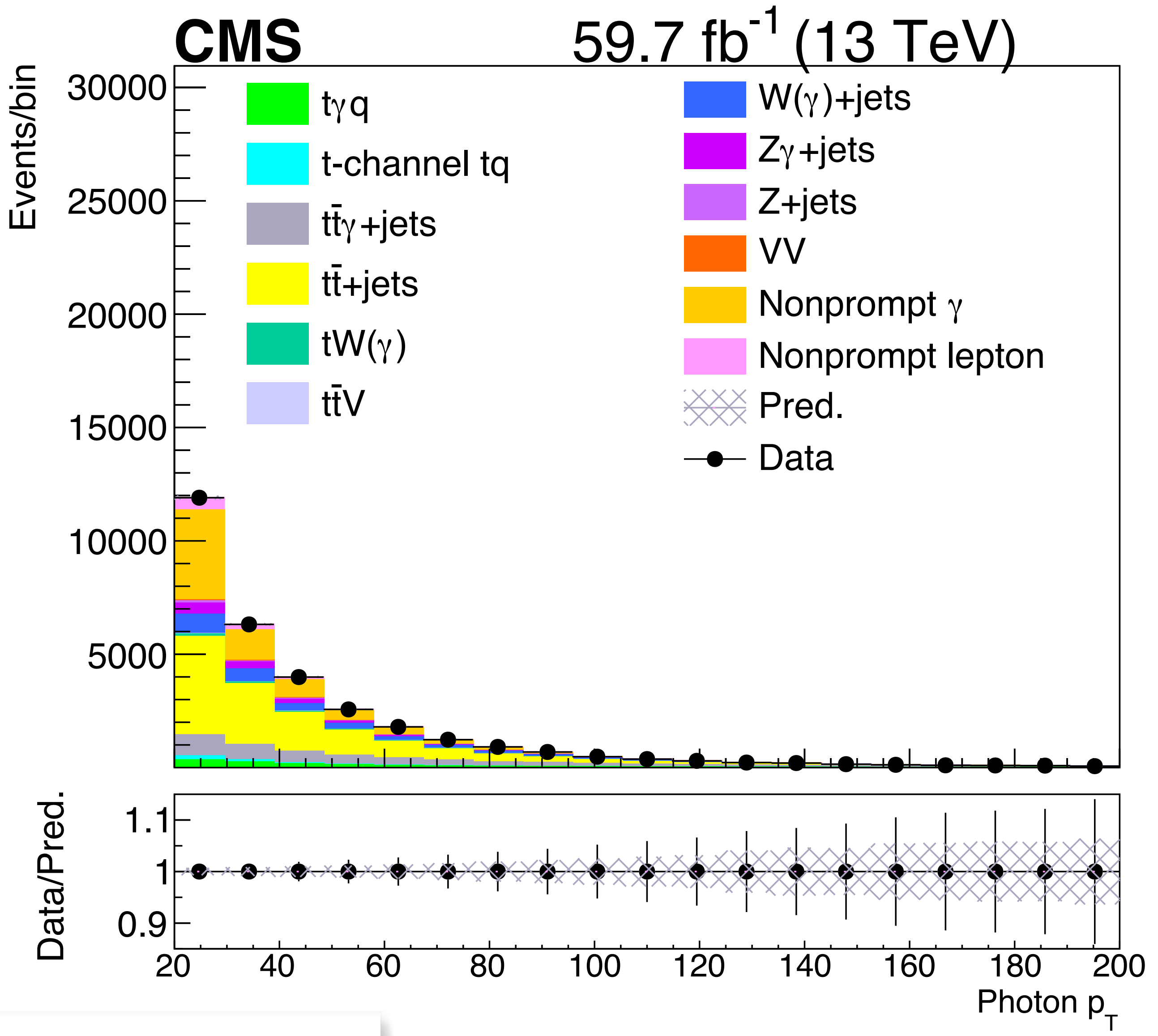
2018 MC only

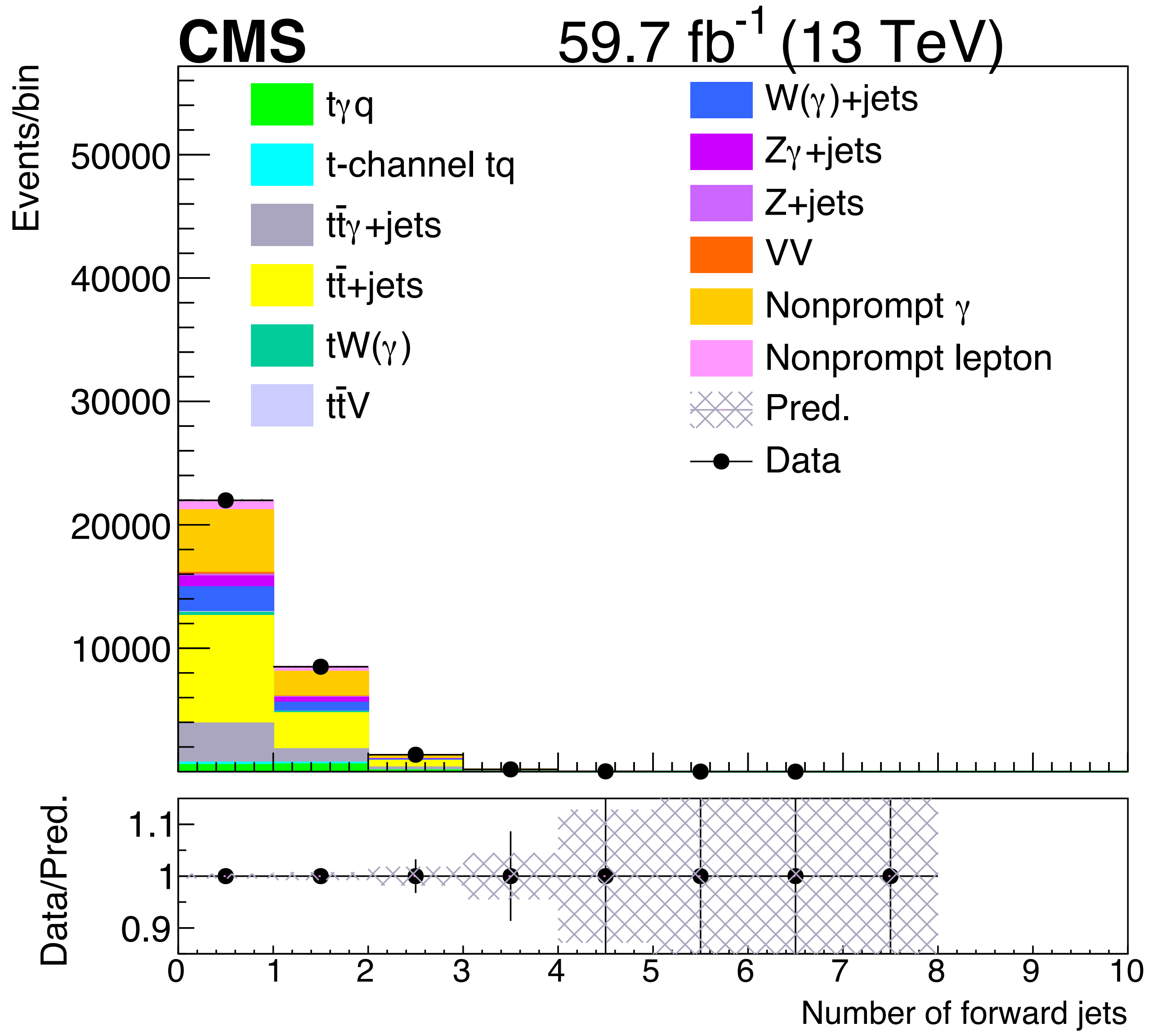










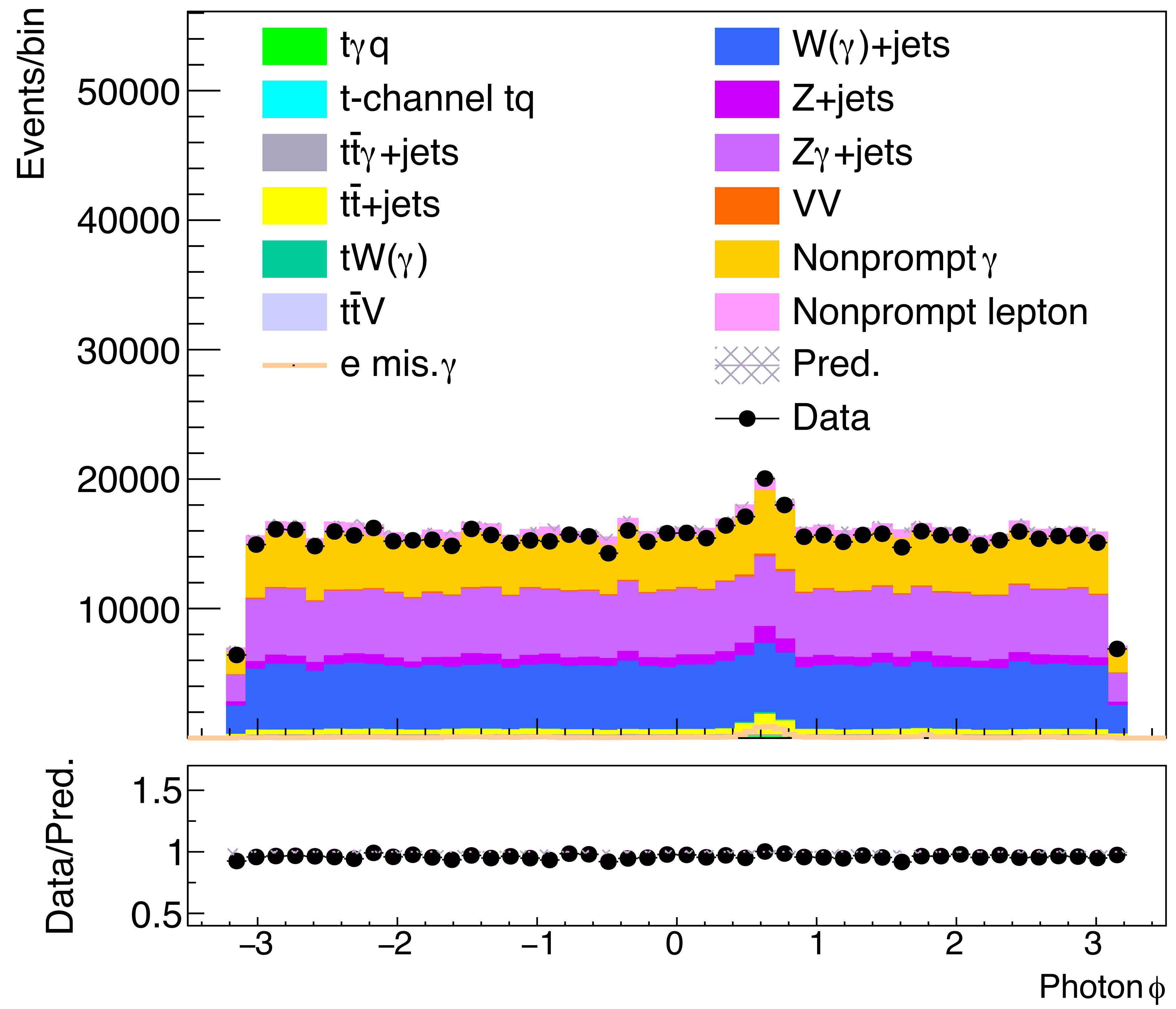


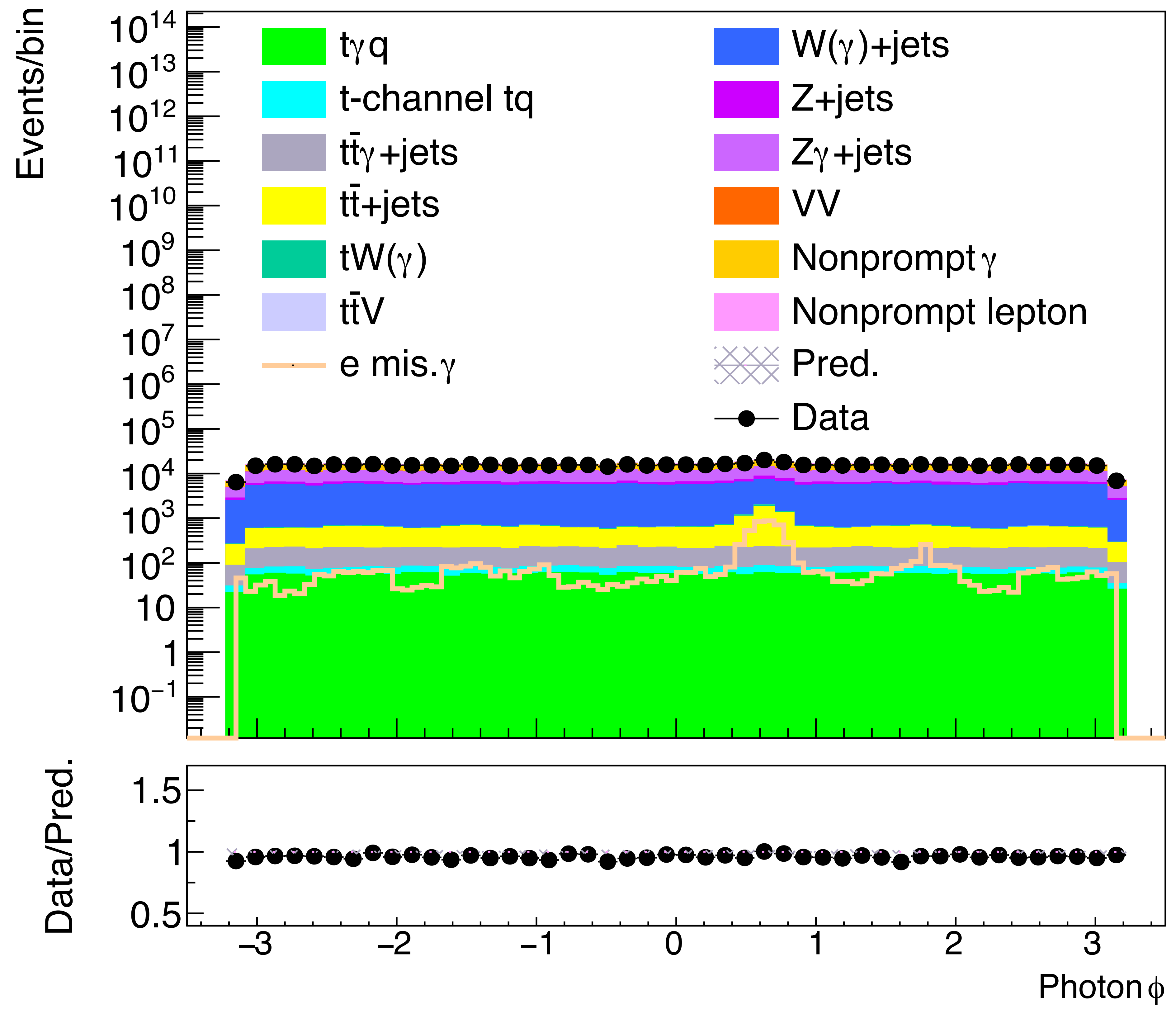
Use tt̄γ NLO, so the tt̄ (yellow) has large contribution

backup

Photon ϕ spike in muon

- N_ℓ=1, $N_\gamma \geq 1$ (Only lepton and photon selection)





The case in muon is not extreme as the electron channel. In order not to introduce any bias, ***events in region of photon ϕ [0.5, 0.7] are removed***

Systematic uncertainties

Theoretical uncertainties:

- Renormalisation and factorisation → envelope after exclusion of two extreme cases
- PDF
- Parton shower (FSR, ISR)

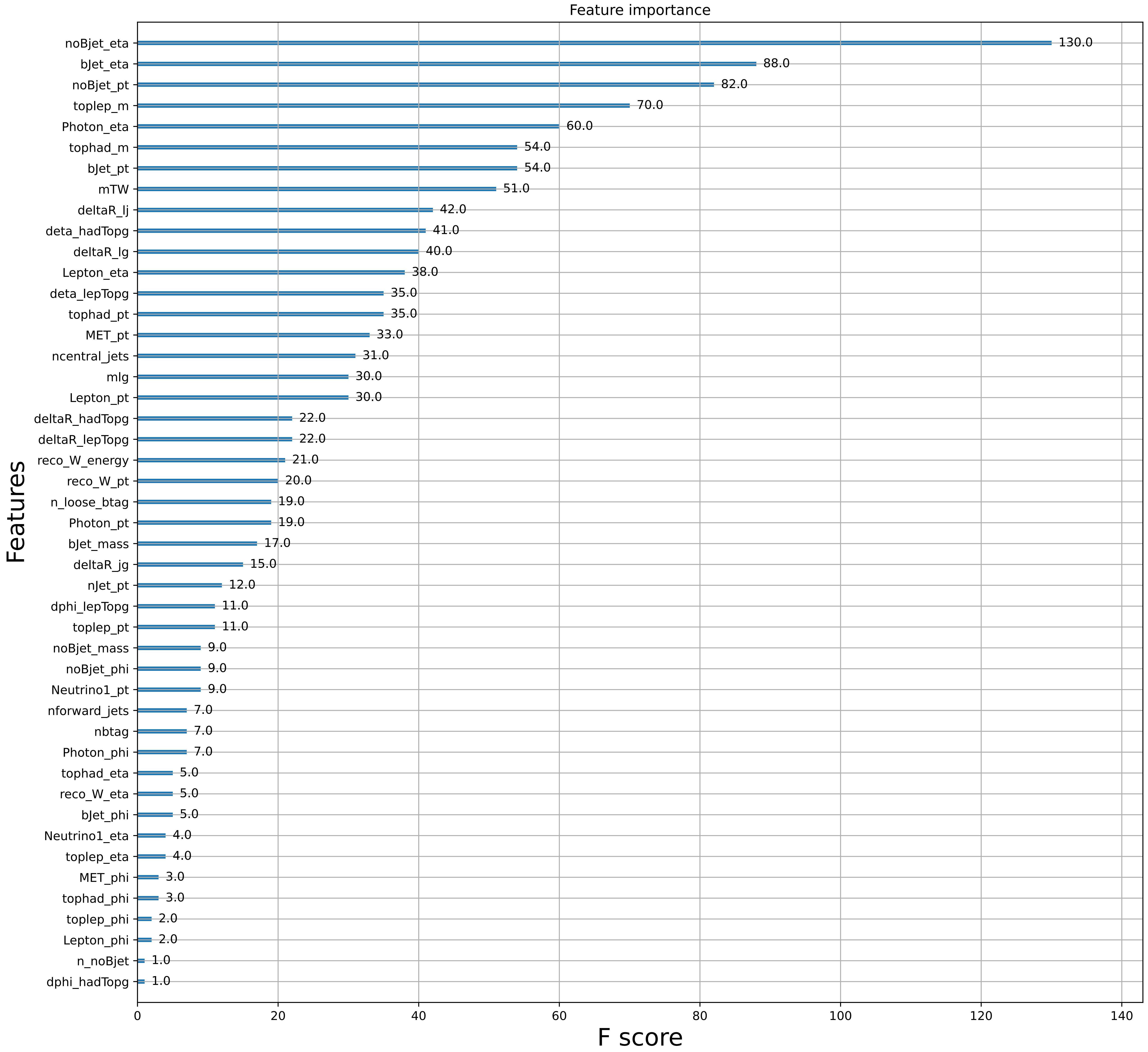
Experimental uncertainties:

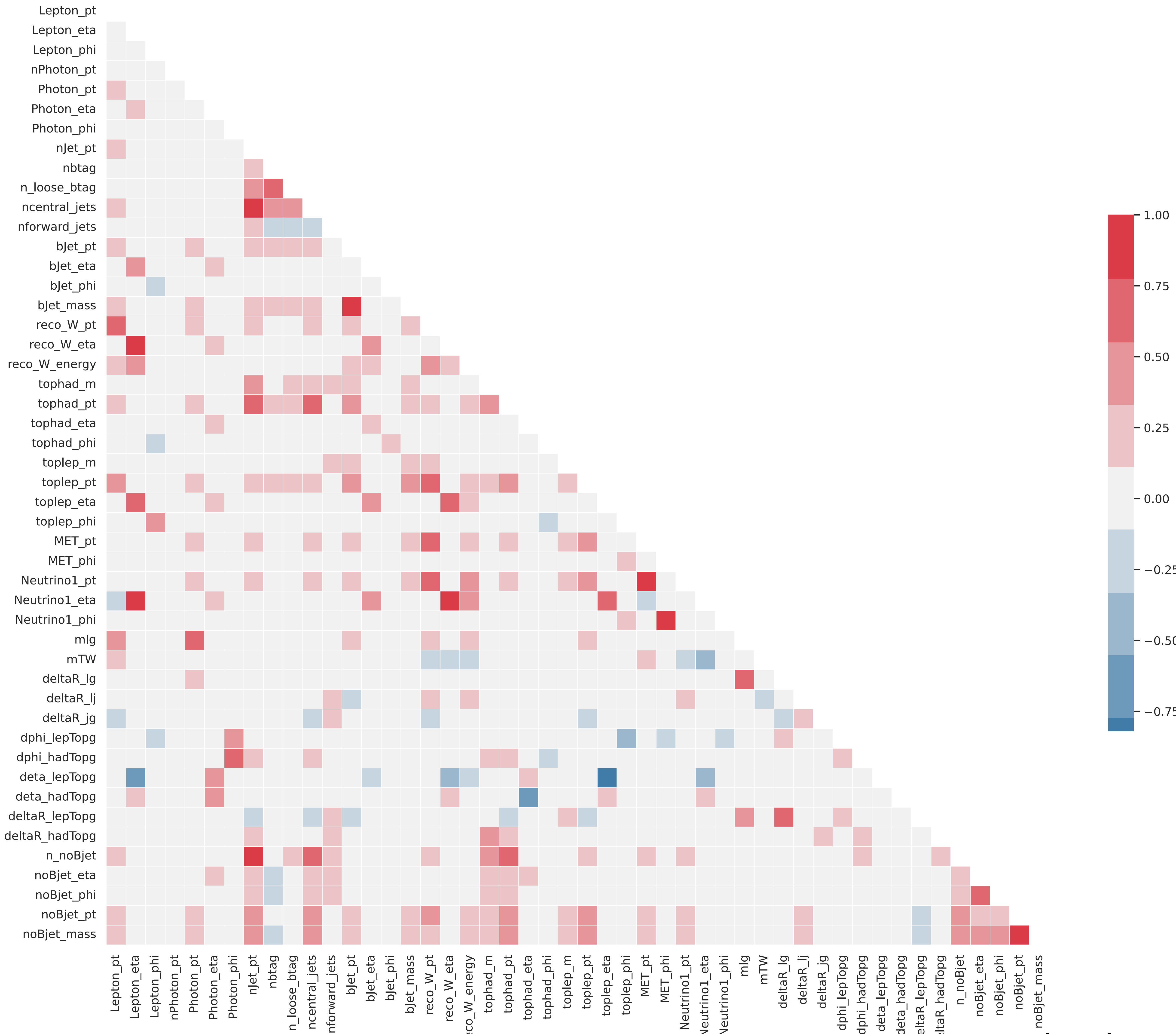
- Luminosity, PU, L1 pre-firing (2016 and 2017)
- Lepton ID/ISO/reco/HLT
- Photon ID/veto scale factors
- Nonprompt photon/lepton estimation → 30% per bin
- Jet energy scale and resolution → not split
- Statistical uncertainties from MC and data-driven

Shape style in data card:
Both shape and normalisation effected are considered

ML training

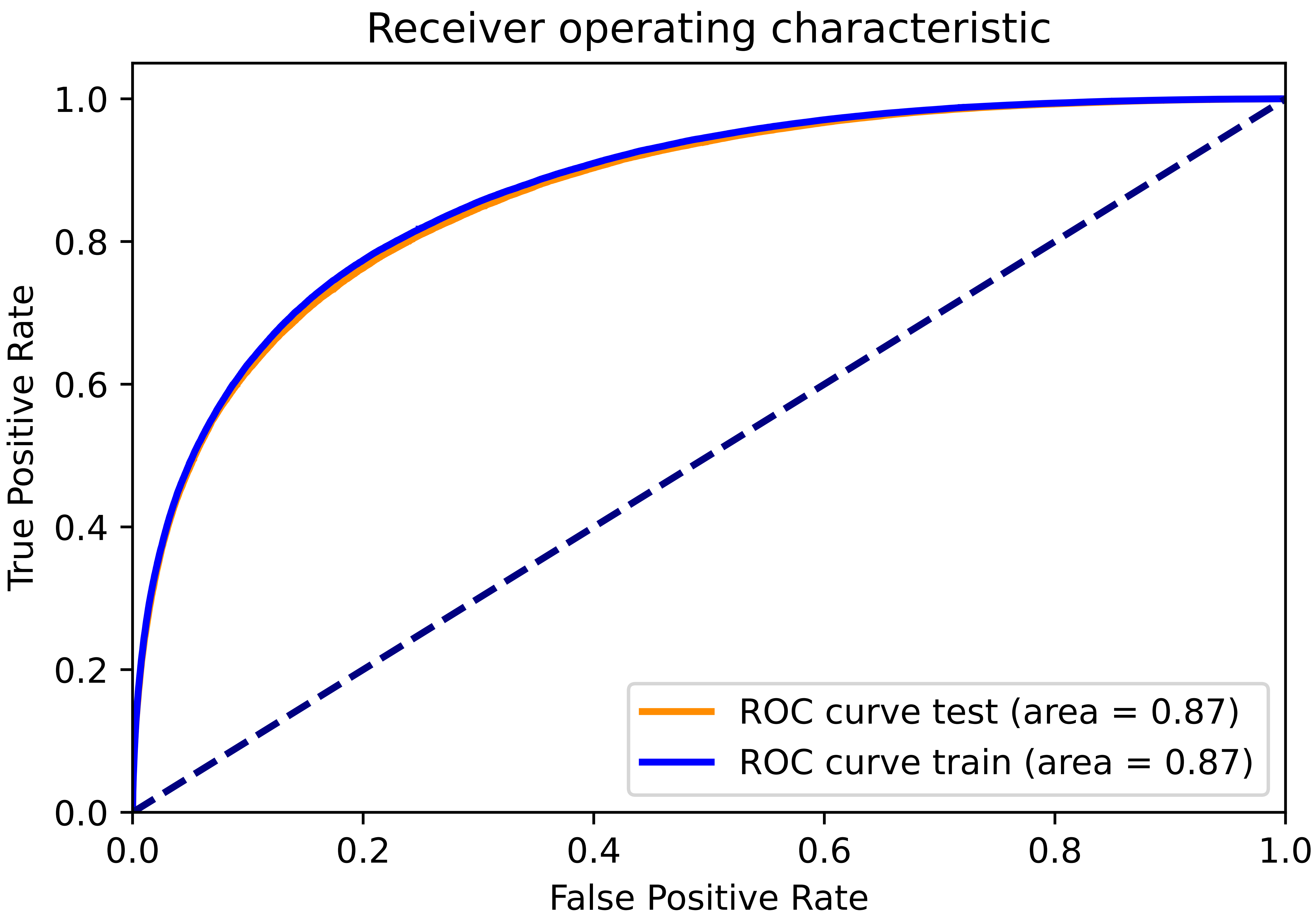
A BDT is trained in the signal region $N_\ell=1$, $N_\gamma \geq 1$, $N_j \geq 2$, $N_b \geq 1$



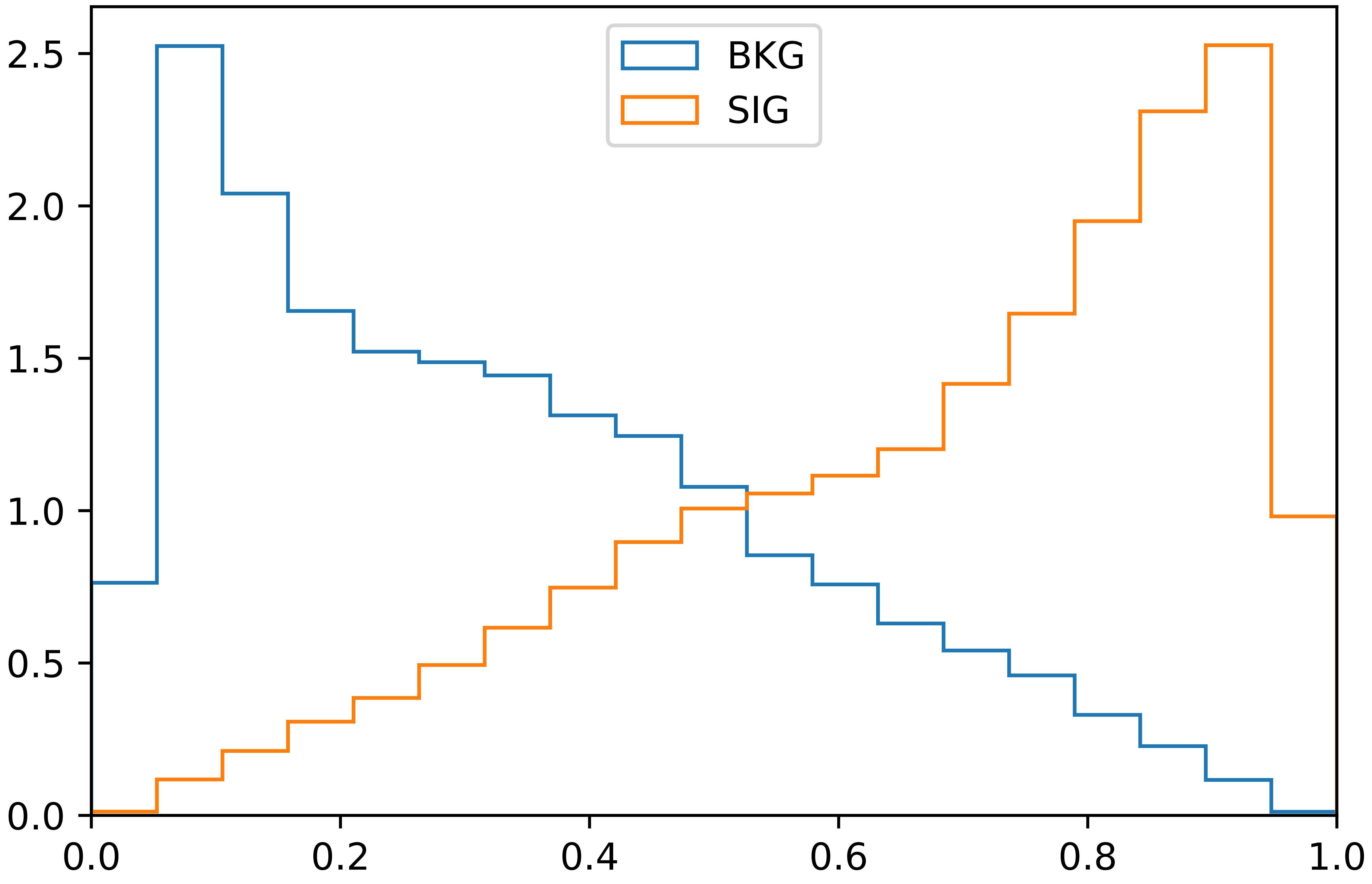


backup

ML training



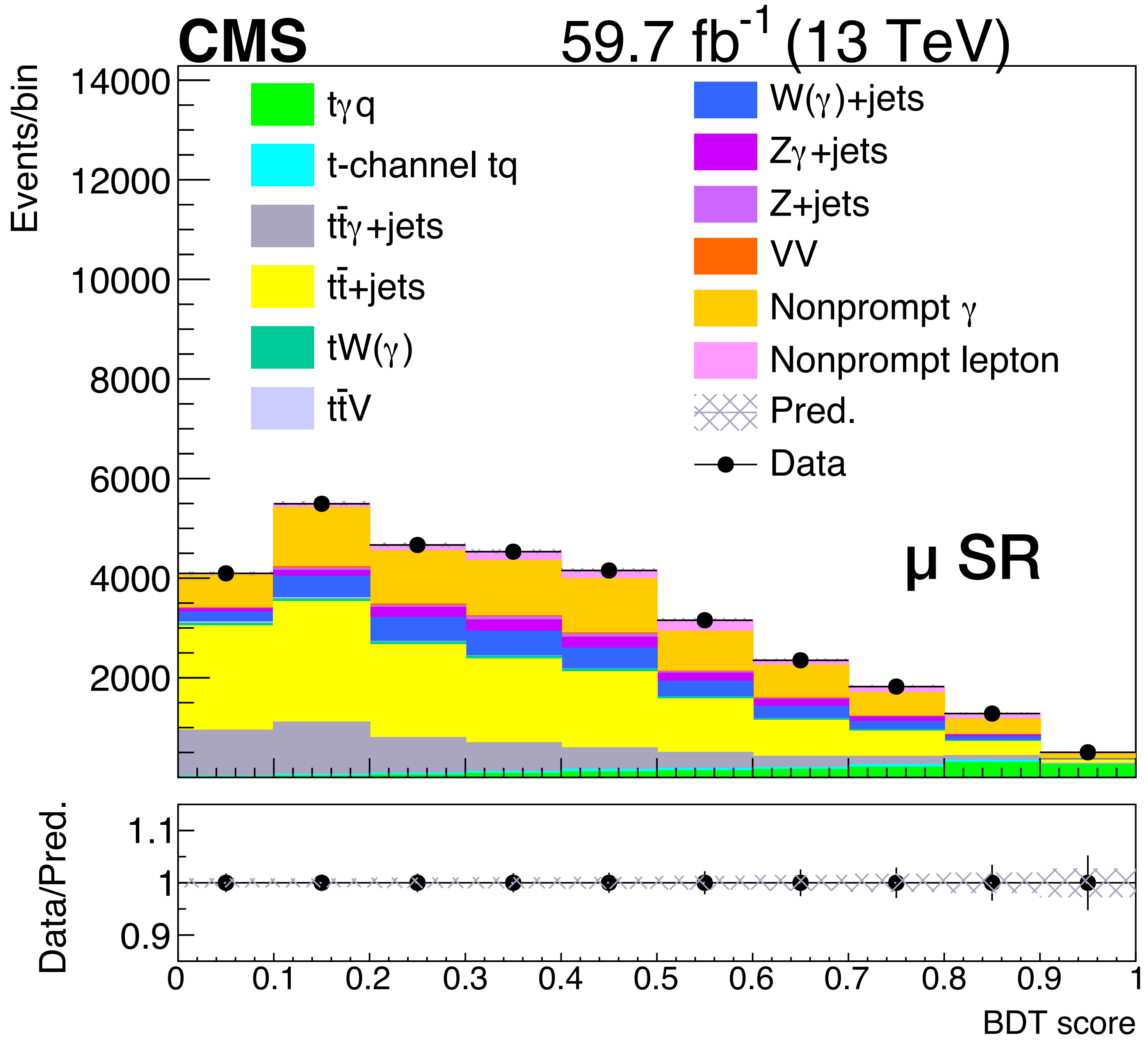
Receiver operating characteristic
True Positive Rate
False Positive Rate
ROC curve test (area = 0.87)
ROC curve train (area = 0.87)
0.0
0.2
0.4
0.6
0.8
1.0

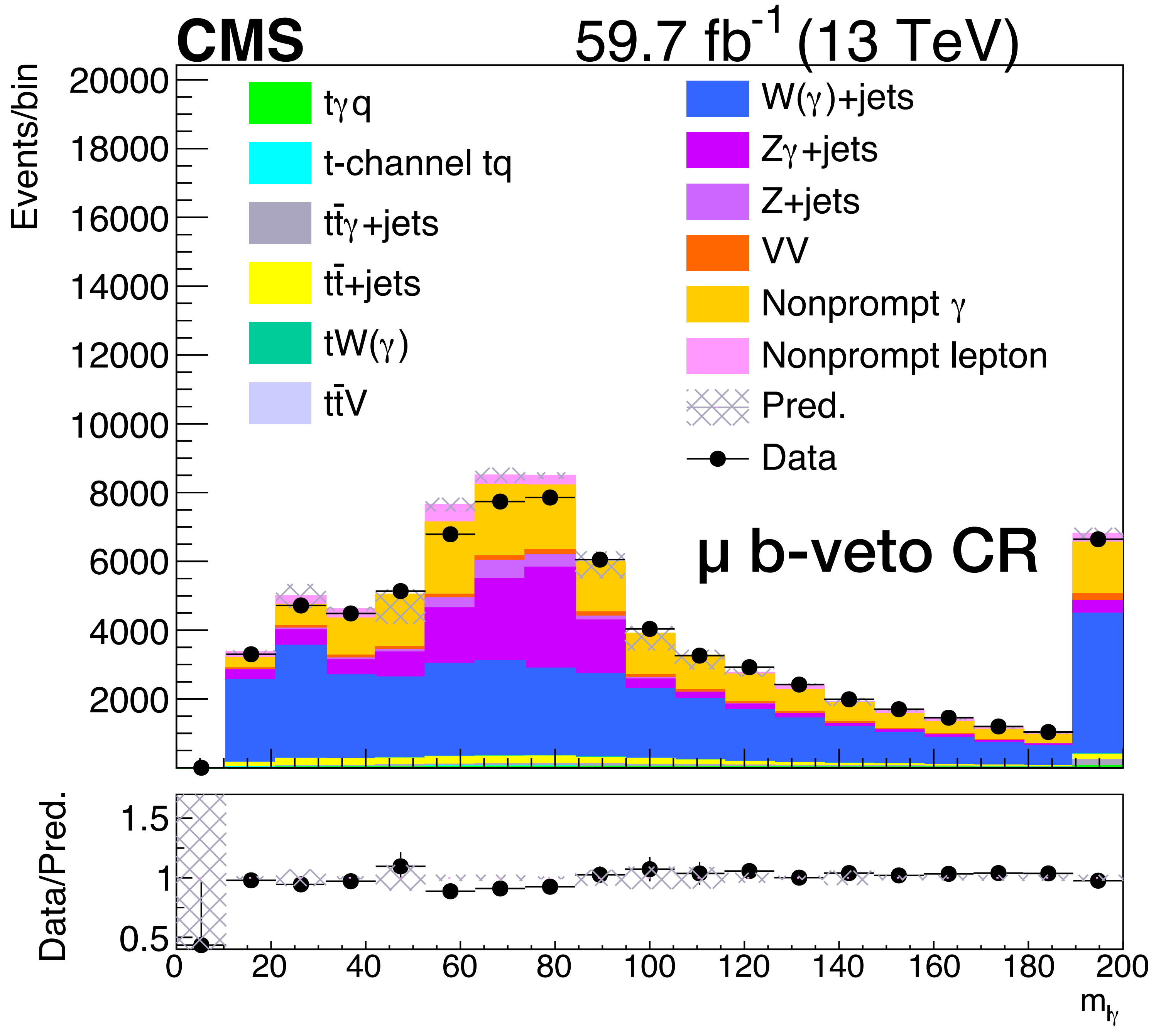


BKG
SIG
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0

Fit strategy

- The signal significance is calculated by performing a simultaneous fit for events in the **signal** (N_ℓ=1, $N_\gamma \geq 1$, $N_j \geq 2$, $N_b \geq 1$) and **b-veto control regions**
- The signal region uses the BDT distribution → 10 bins from 0 to 1 uniformly
- The control region uses the $m_{\ell\gamma}$ distribution → 10 bins from 0 to 200 uniformly
- Some uncertainties are correlated between two regions



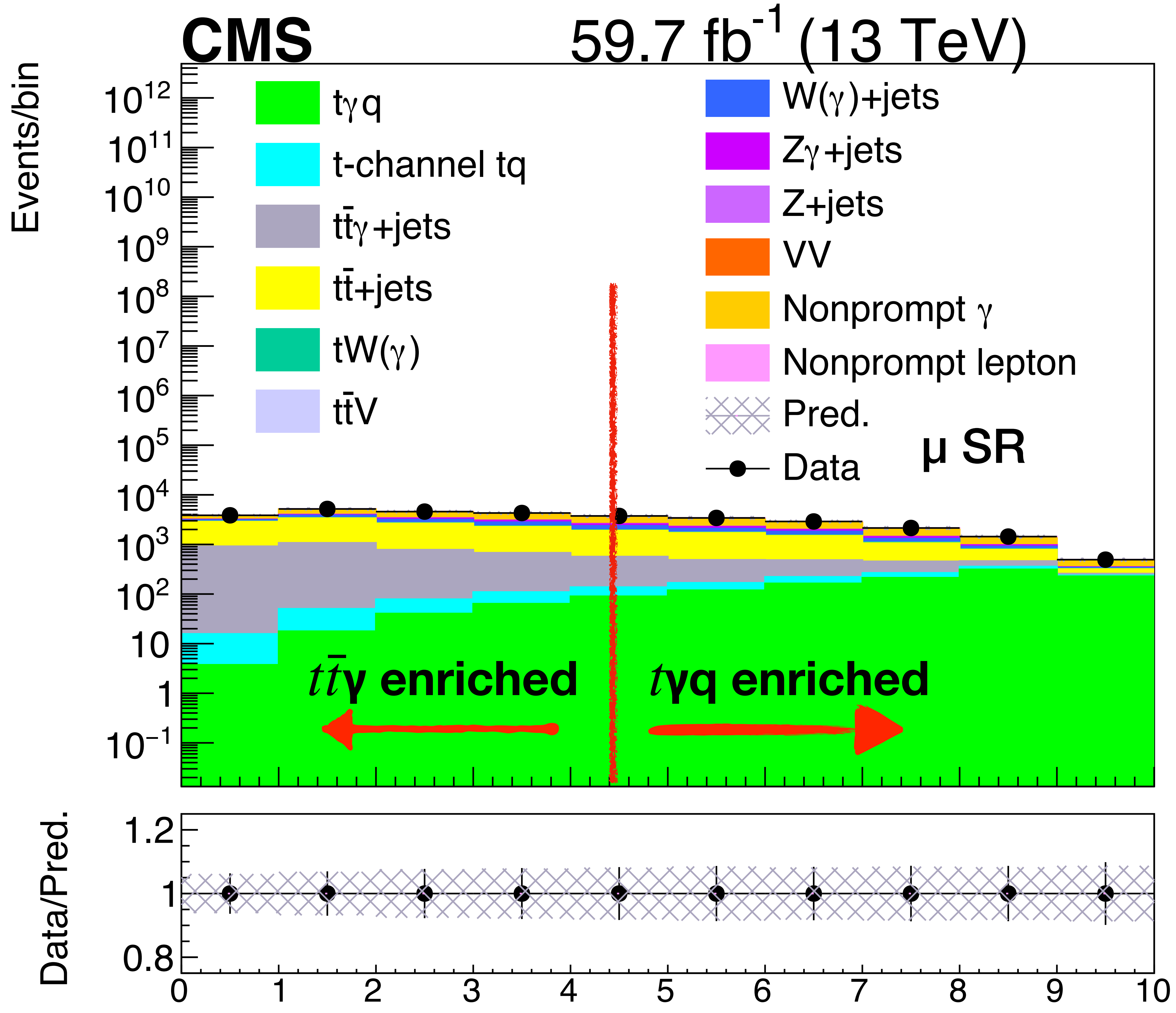


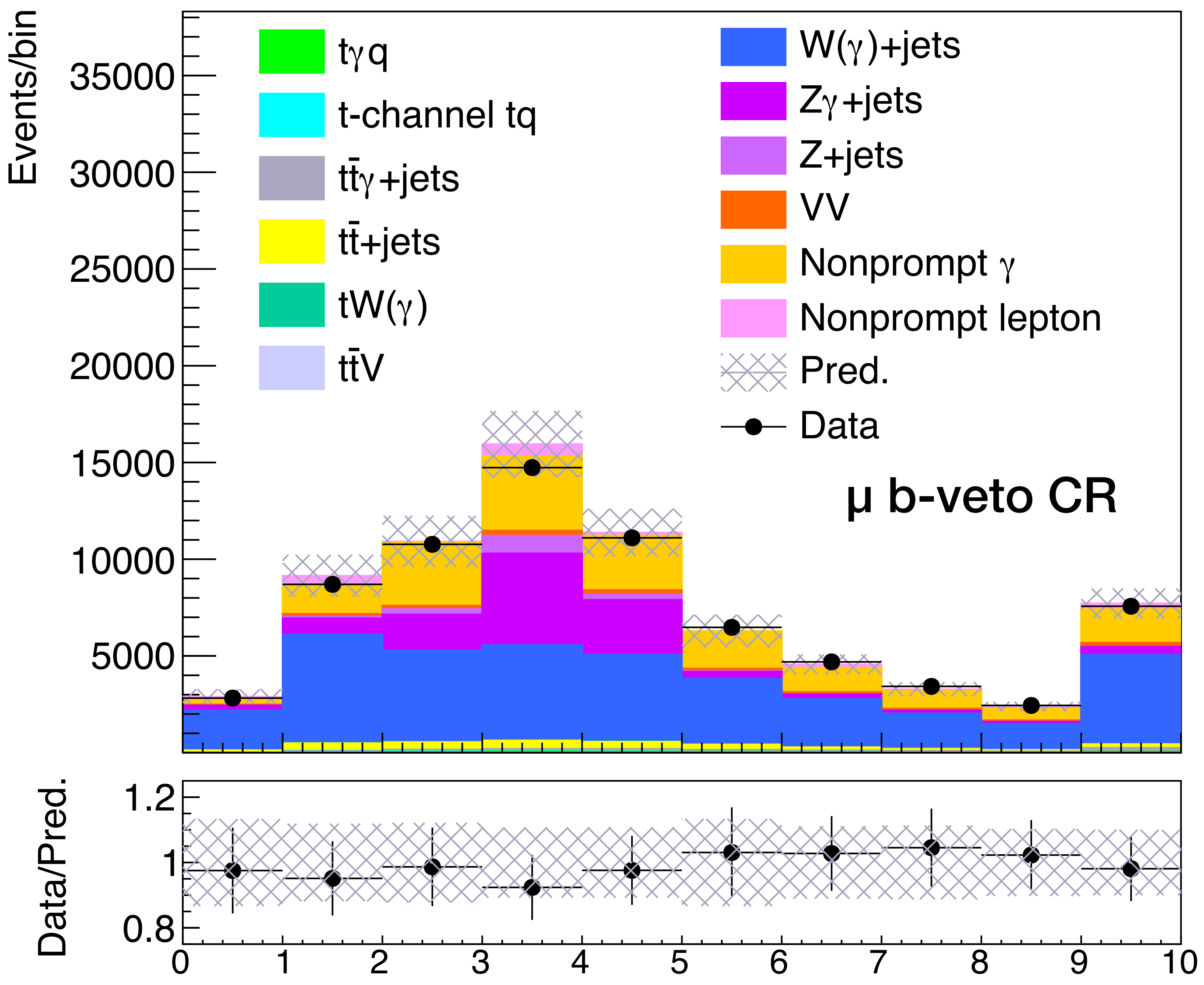
Only statistical uncertainties are added

Inclusive fit result of $t\gamma q$

The expected significance: 7.3 σ

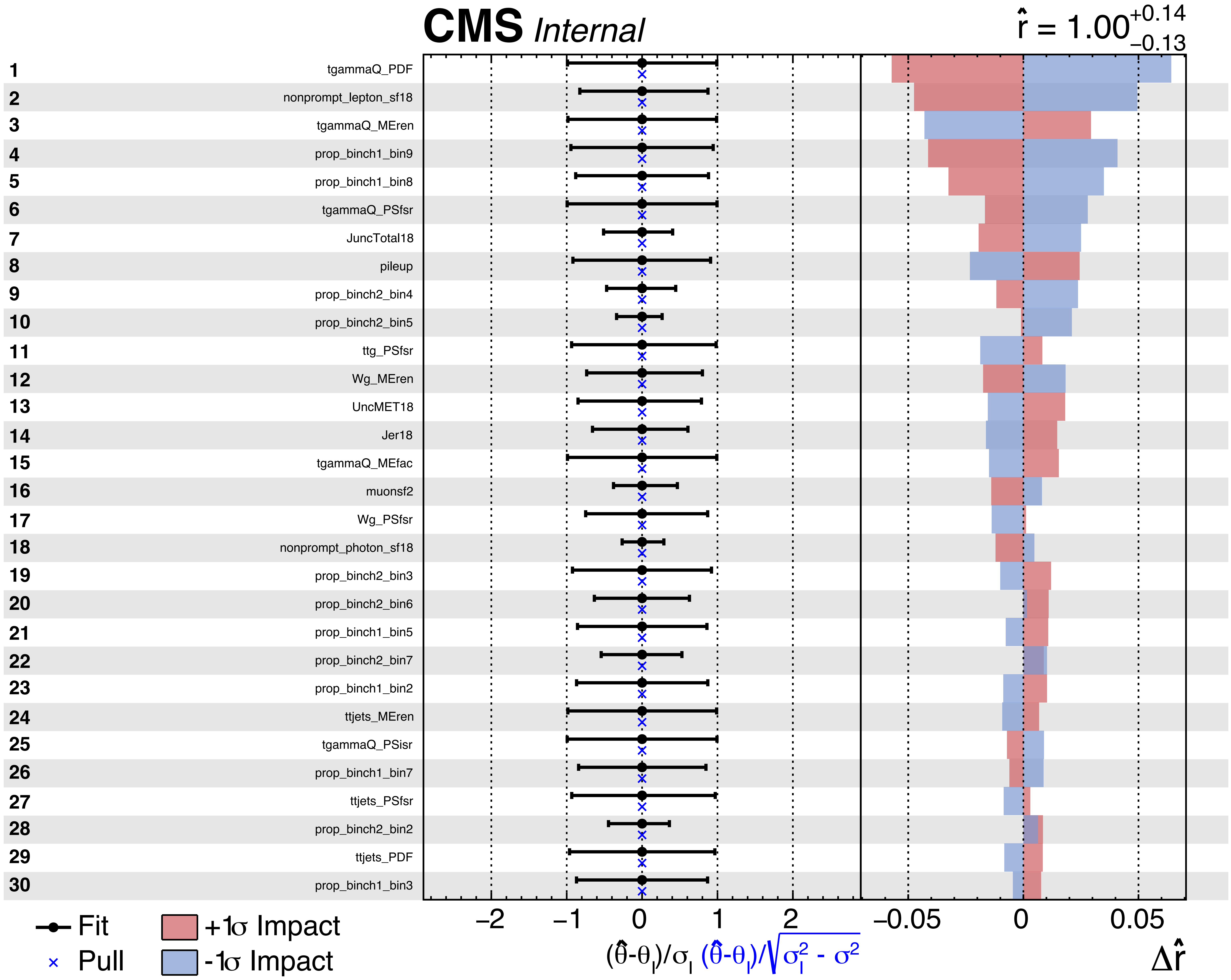
The expected signal strength (μ): $1.00^{+0.14}_{-0.13}$





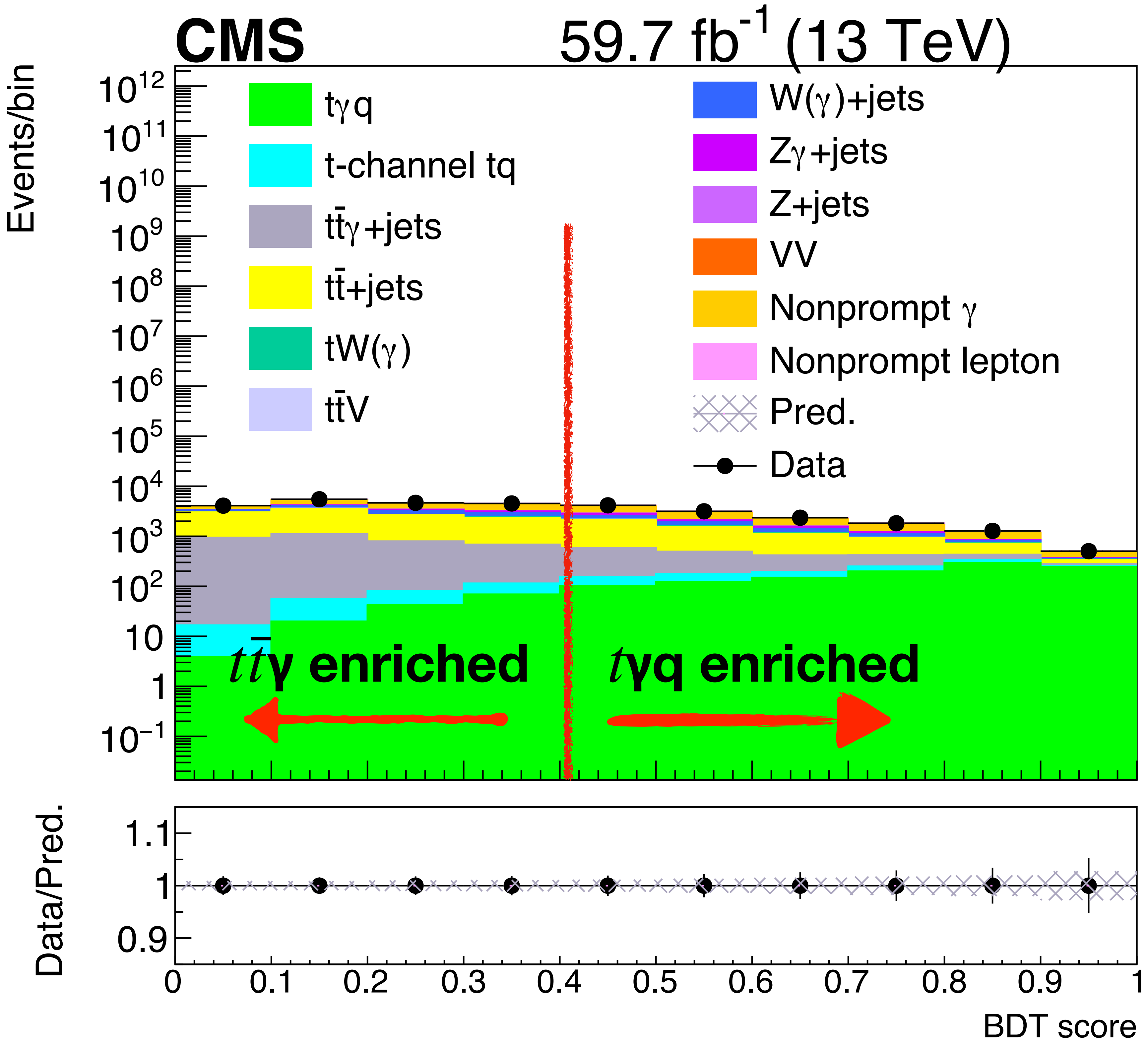
All uncertainties are added

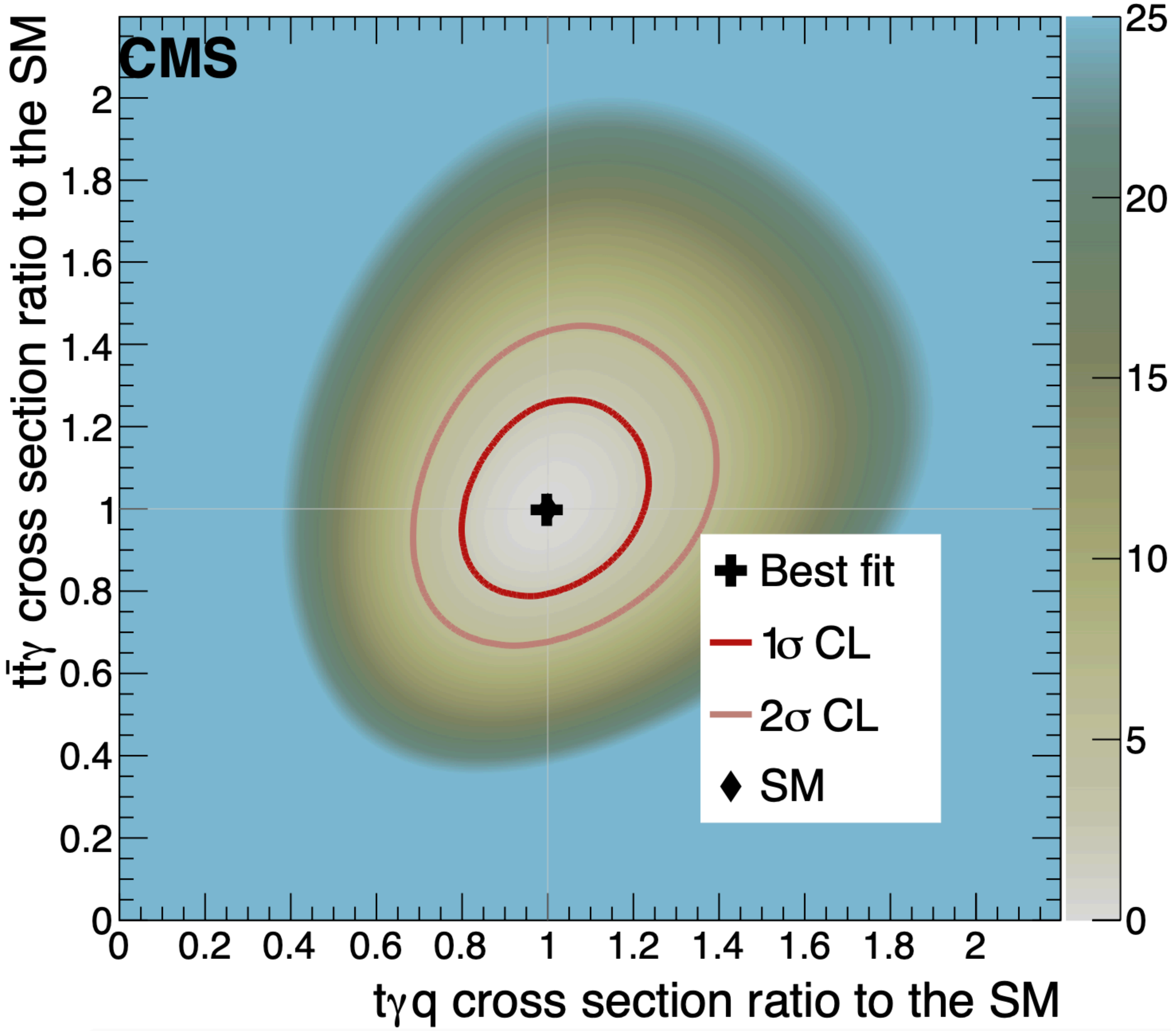
Inclusive fit result — impact



Simultaneous fit of $t\gamma q+t\bar{t}\gamma$

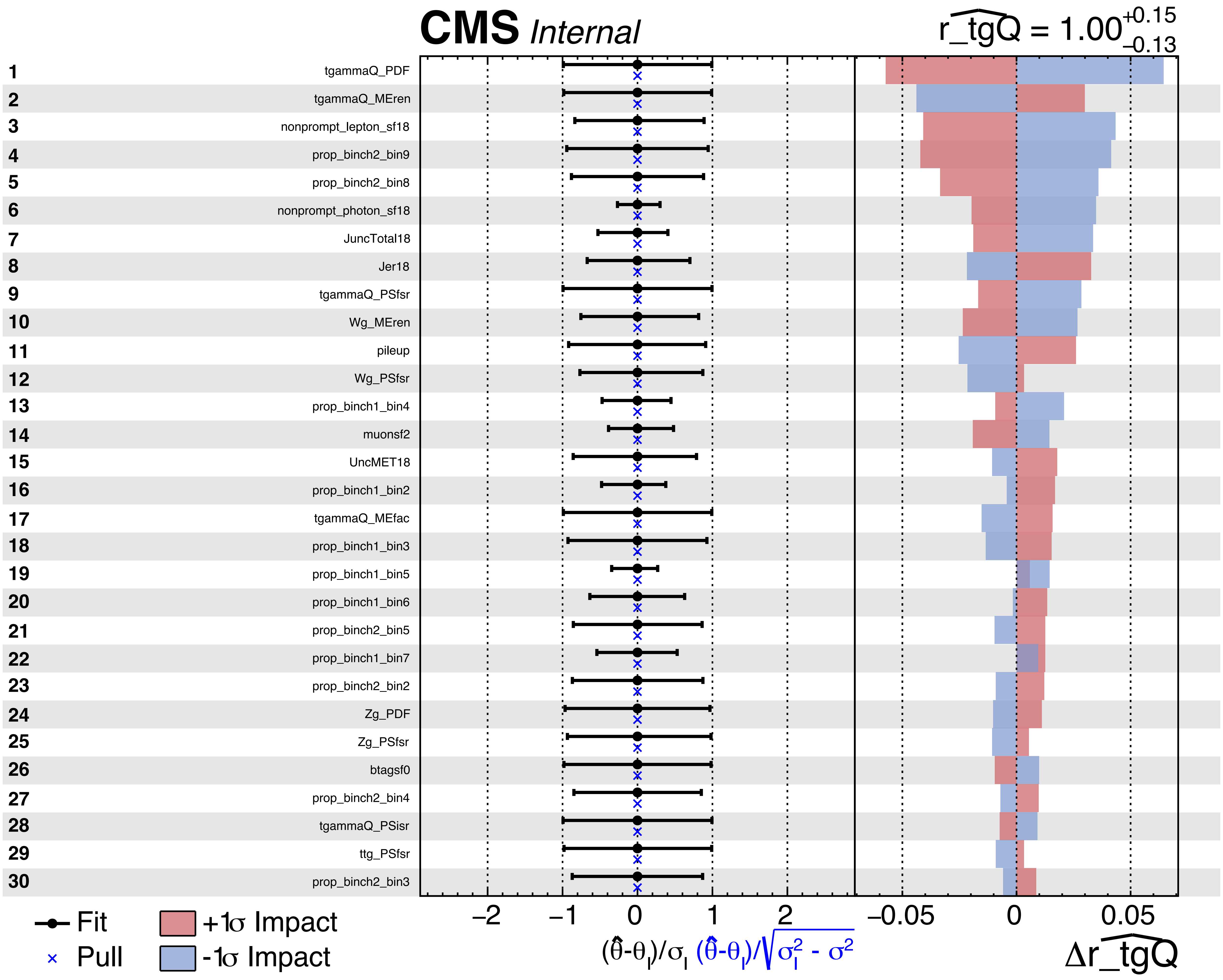
- The signal significance is calculated by performing a simultaneous fit for events in the **signal** (N_ℓ=1, $N_\gamma \geq 1$, $N_j \geq 2$, $N_b \geq 1$) and **b-veto control regions**
- Both $t\gamma q$ and $t\bar{t}\gamma$ are regarded as signal

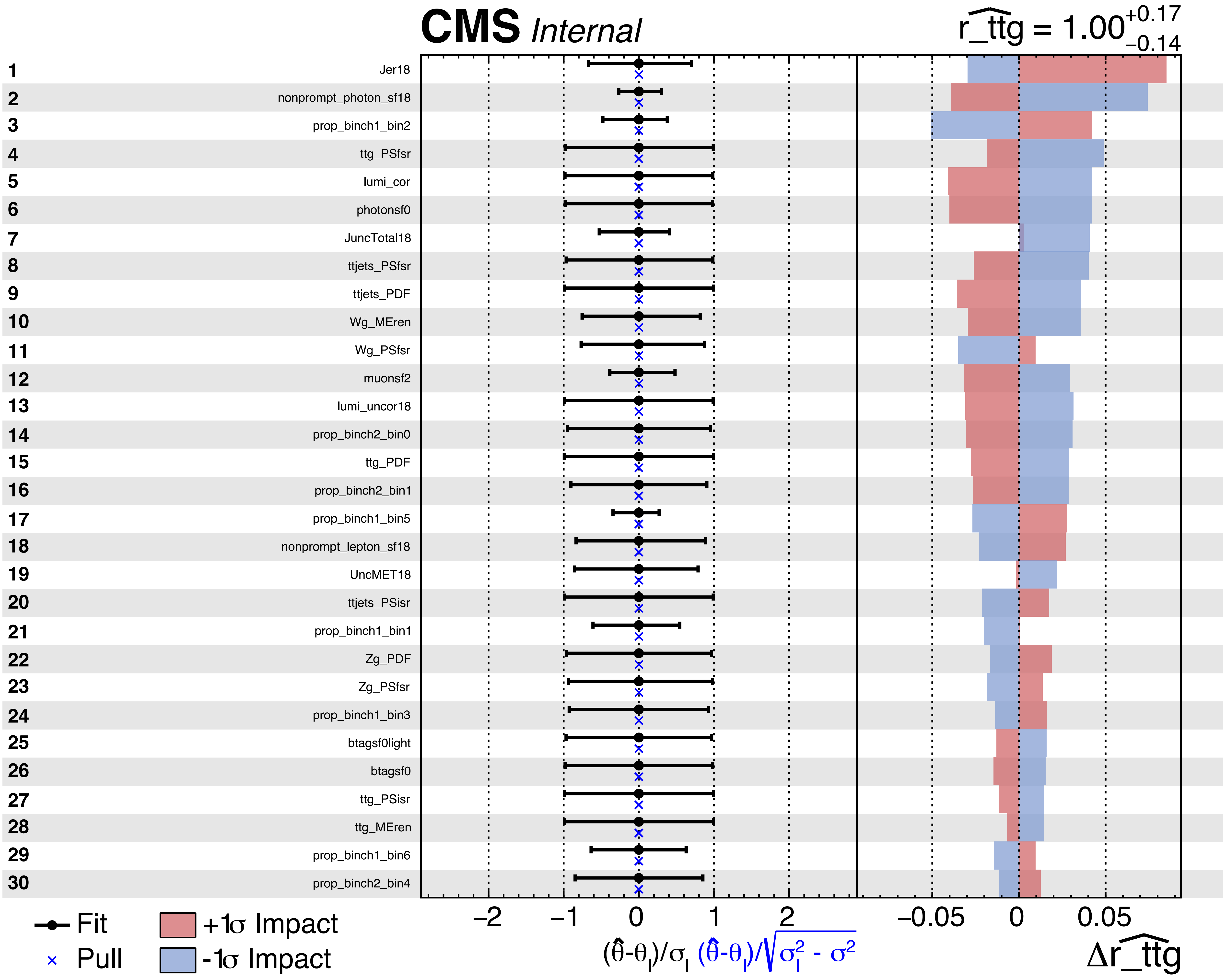




Expected Best fit results: $r_{t\gamma q} = 1.0^{+0.15}_{-0.13}$ $r_{t\bar{t}\gamma q} = 1.0^{+0.17}_{-0.14}$

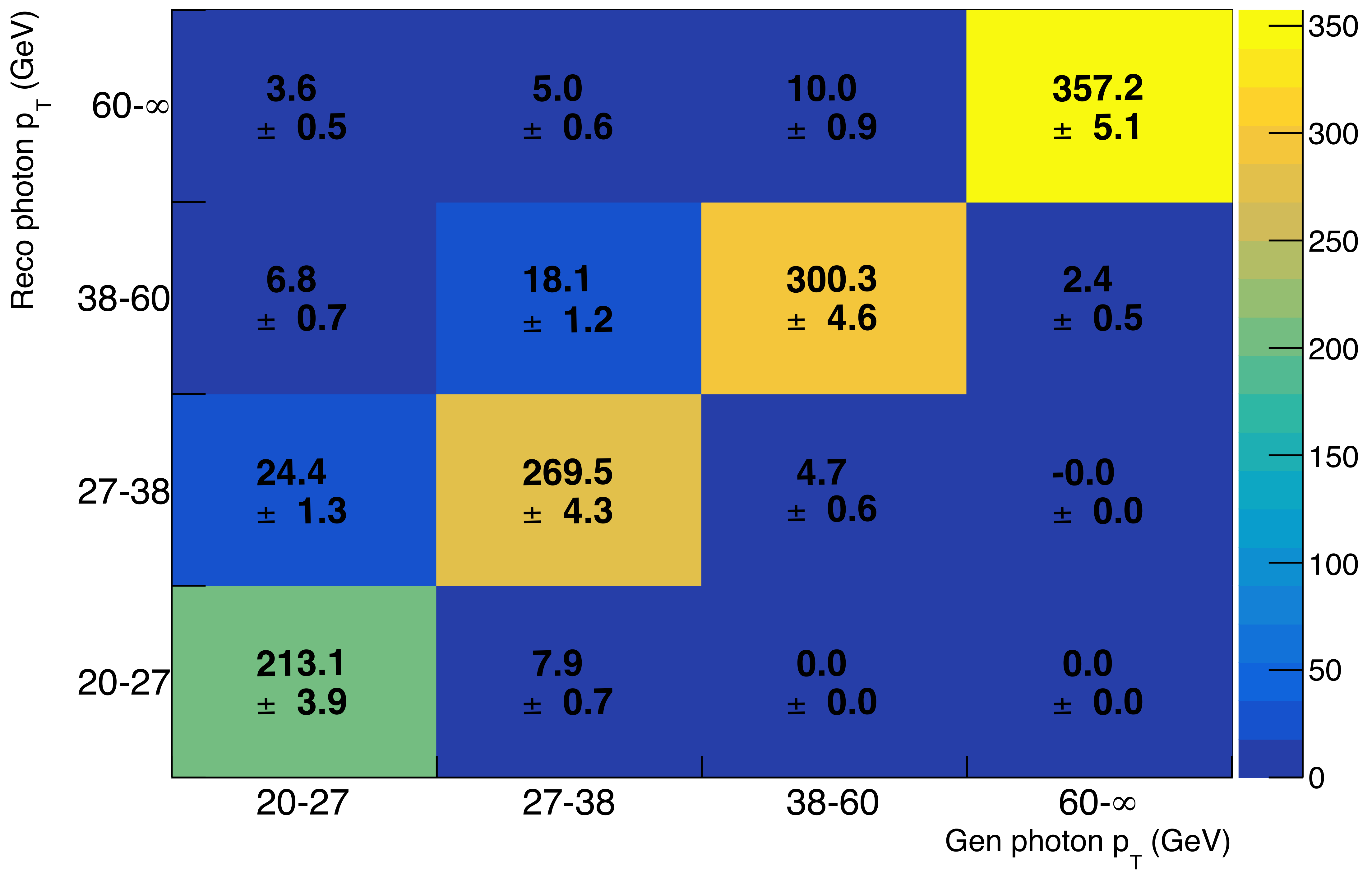
Simultaneous fit of $t\gamma q + t\bar{t}\gamma$ — impact

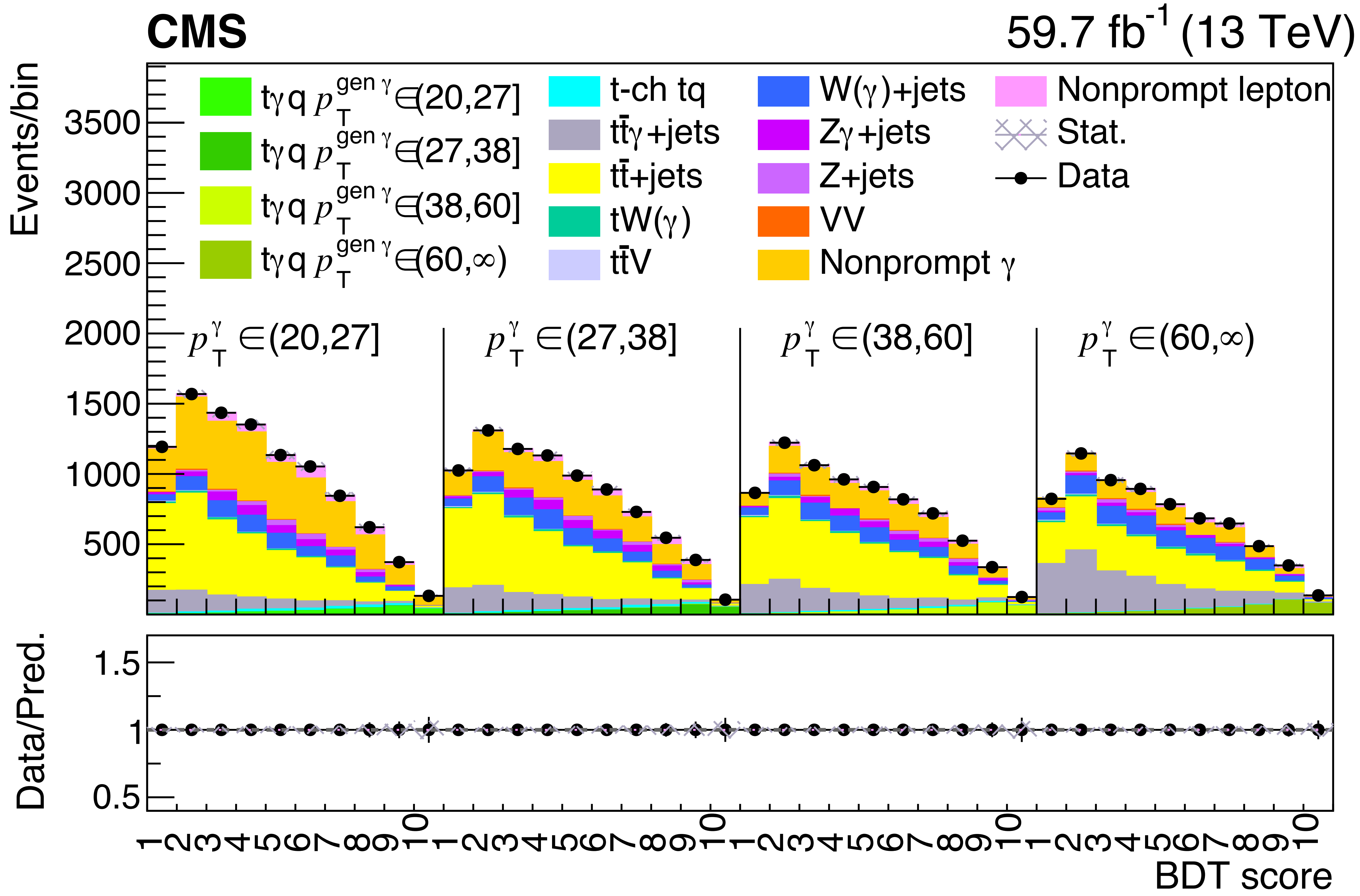




Differential fit of $t\gamma q$ — p_T^{γ}

- Check response matrix→ 2-dimensional distribution of p_T in reco- and gen-level
  - Quite diagonal response matrices: condition number < 10, regularization not needed
- Split signal into different bins according to the variable measured in generator-level
- Categorise events to different bins according to the variable measured in reco-level
  - BDT distribution used in every bin category

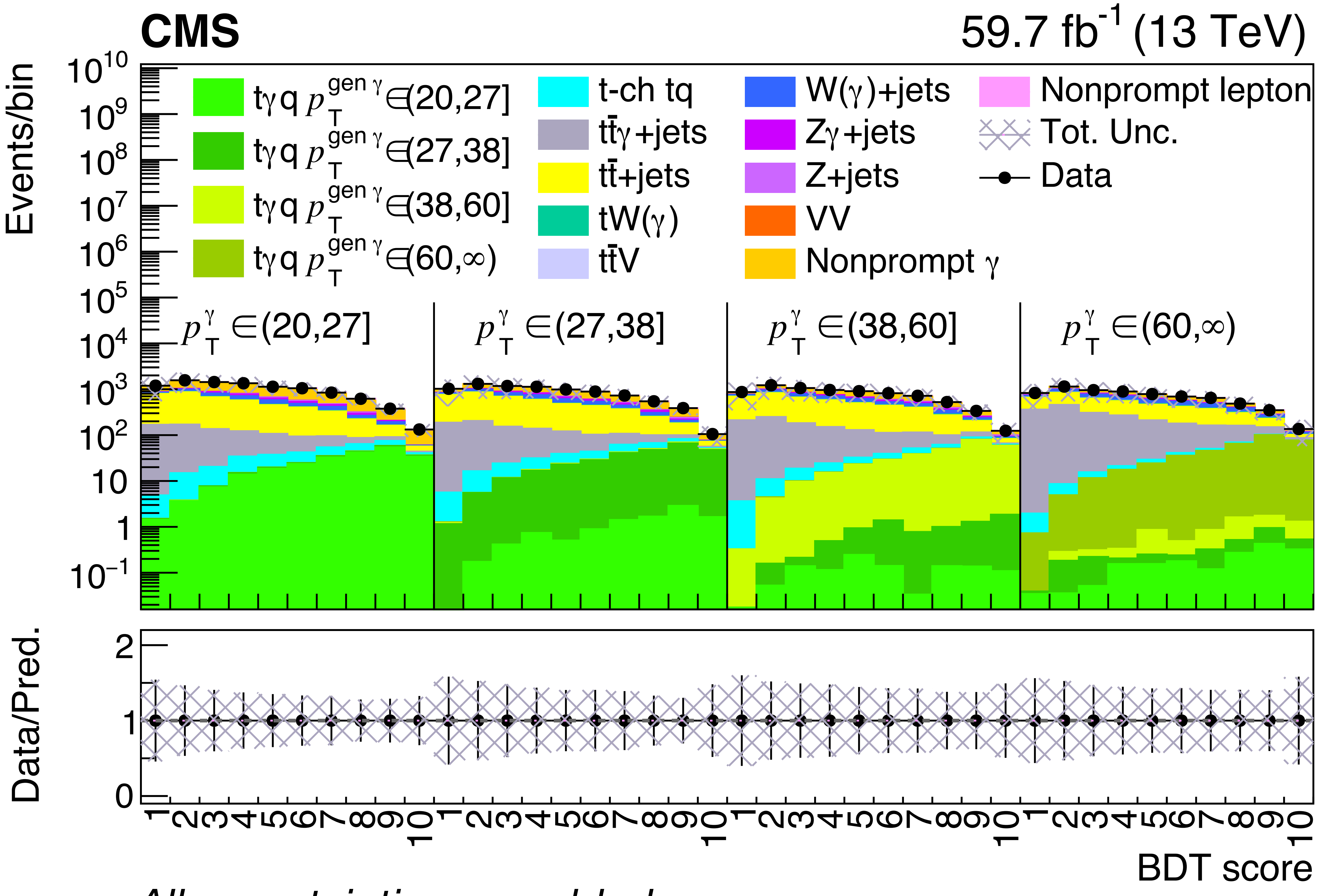




[backup](#)

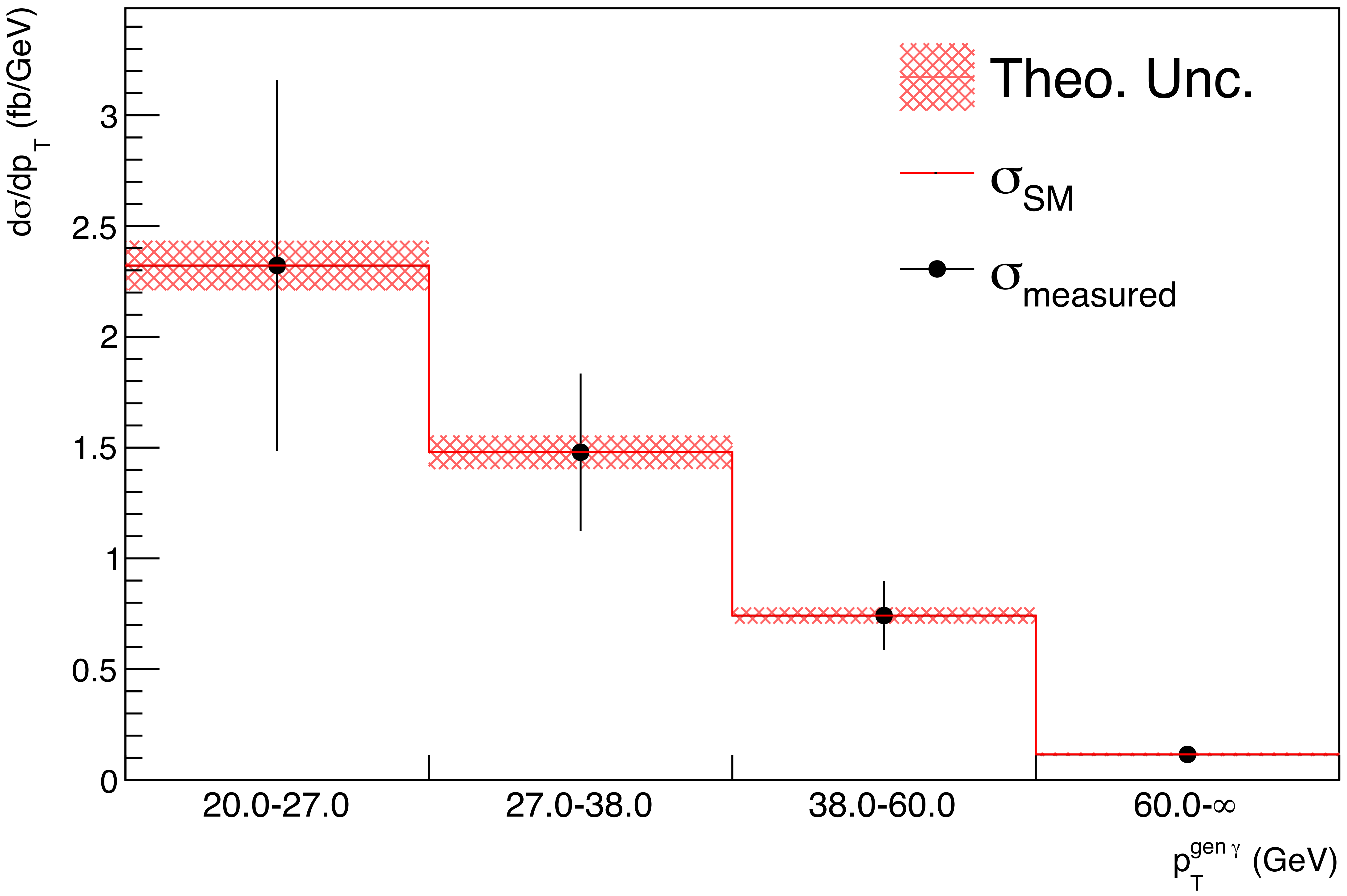
Differential fit of $t\gamma q$ — p_T^{γ}

- POIs: signal events in four $p_T^{\mathrm{gen}\,\gamma}$ bins
- Fit signal region ($N_j \geq 2$, $N_b \geq 1$) (only SR now) → will include control regions



All uncertainties are added

Variable	r_1	r_2	r_3	r_4
p_T^{γ}	1.0 ± 0.36	1.0 ± 0.24	1.0 ± 0.21	1.0 ± 0.17



Next to do

- Build DNN with three output nodes → tγq, ttγ, and others
  - DNN tuning → hyper-parameters, input features, structures
- Photon ϕ spike solution → get confirmation with EGamma group (really large effect in electron channel)
- e mis. γ data-driven background → get electron channel results
- Get Full Run2 results
- Start the EFT interpretation
- Prepare documentation (AN note and etc.)

Backup

Rough estimation for $t\gamma q$ full Run2

CMS $t\gamma q$ with 2016 data had significance 4.4 (3) σ in μ channel and only barrel photon

Simple calculation for sensitivity with full Run 2 data:

- It could be 8.5 (6) σ in μ channel
- If add e channel, it could be 10 (8) σ with precision 0.1
- If add endcap photon events, could increase a little bit more

CMS $t\gamma q$ post-fit yields table

Process	Event yield
$t\bar{t}+\gamma$	1401 ± 131
$W\gamma$+jets	329 ± 78
$Z\gamma$+jets	232 ± 55
Misidentified photon	374 ± 74
$t\gamma$ (s- and tW-channel)	57 ± 8
$VV\gamma$	8 ± 3
Total background	2401 ± 178
Expected signal	154 ± 24
Total SM prediction	2555 ± 180
Data	2535

ATLAS $t\gamma q$ post-fit yields table Particle level measurement

	$\geq$ 1fj SR	0fj SR	$t\bar{t}\gamma$ CR	$W\gamma$ CR
$tq\gamma$	2360 ± 250	2450 ± 310	880 ± 120	1260 ± 140
$t(\rightarrow l\nu b\gamma)q$	500 ± 170	660 ± 210	180 ± 60	330 ± 120
$t\bar{t}\gamma$ (production)	3100 ± 400	4700 ± 700	4300 ± 600	2700 ± 400
$t\bar{t}\gamma$ (radiative decay)	3800 ± 600	9200 ± 1400	5600 ± 600	4200 ± 900
$W\gamma$+jets	2500 ± 400	9200 ± 1400	1170 ± 320	31 700 ± 3000
$Z\gamma$+jets	970 ± 310	2700 ± 800	430 ± 150	7700 ± 2400
$e \rightarrow \gamma$ fake photons	5100 ± 500	10 400 ± 800	4900 ± 400	5500 ± 500
$h \rightarrow \gamma$ fake photons	1100 ± 400	2700 ± 900	1300 ± 500	2600 ± 800
Other prompt γ	1340 ± 350	2600 ± 900	1400 ± 400	4000 ± 600
Fake leptons	390 ± 190	1000 ± 500	110 ± 50	3600 ± 1700
Total	21 250 ± 150	45 720 ± 240	20 180 ± 150	63 590 ± 280
Data	21 227	45 723	20 194	63 592

back

Datasets

- TGJets_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8: **0.995 pb**
- ST_tchannel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8: **136.02 pb**
- ST_tchannel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8: **80.95 pb**
- TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8: **3.697 pb**
- TTGamma_SingleLept_TuneCP5_13TeV-madgraph-pythia8: **5.056 pb**
- TTGamma_Dilept_TuneCP5_13TeV-madgraph-pythia8: **1.495 pb**
- TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8: **833.9 pb×0.44≈367 pb**
- TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8: **833.9 pb×0.107≈89.2 pb**
- DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8: **6077.22 pb**
- ZGToLLG_01J_5f_lowMLL_lowGPt_TuneCP5_13TeV-amcatnloFXFX-pythia8: **174.1 pb**
- WJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-pythia8: **61526.7 pb**
- WGToLNuG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8: **191 pb**
- TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8": **0.5104 pb**
- TTZToNuNu_TuneCP5_13TeV-amcatnlo-pythia8”: **0.1476 pb**
- TTZToLL_TuneCP5_13TeV_amcatnlo-pythia8”: **0.050 pb**
- TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8": **0.243 pb**
- TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8": **0.502 pb**
- ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8": **34.9 pb**
- ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8": **34.0 pb**
- ST_tWAto2L2Nu_5f_TuneCP5_13TeV-madgraph-pythia8": 191.1, **0.191 pb**

- WW_TuneCP5_13TeV-pythia8": **118.7 pb**
- WZ_TuneCP5_13TeV-pythia8": **47.13 pb**
- ZZ_TuneCP5_13TeV-pythia8": **16.523 pb**

back

Object selection

Electron	Good	Veto	Fakeable[1]
p_T/GeV	> 35	> 15	> 35
$\|\eta\|$	< 2.5 not in ECAL gap	< 2.5 not in ECAL gap	< 2.5 not in ECAL gap
ID	cut-based medium ID	cut-based veto ID	pass Veto but fail Good
Others	Impact (d_{xy},d_z)	—	Impact (d_{xy},d_z)

*[1] **Fakeable lepton** is used for estimating nonprompt ℓ*

Muon	Good	Veto	Fakeable[1]
p_T/GeV	> 30	> 15	> 30
$\|\eta\|$	< 2.4	< 2.4	< 2.4
ID	cut-based tight ID	cut-based loose ID	cut-based tight ID
Iso	Tight Iso (<0.15)	Very loose Iso (<0.4)	[0.15, 0.4]

Photon	Good	Fakeable
p_T/GeV	> 20	>20
$\|\eta\|$	< 2.5 not in ECAL gap	< 2.5 not in ECAL gap
ID	cut-based medium ID	part of cut-based medium ID
Electron-veto	pixel seed veto	pixel seed veto
Others	—	fail medium $\sigma_{i\eta i\eta}$

JetMET	Jet	b-jet	MET
p_T/GeV	> 30	> 30	> 20
$\|\eta\|$	< 4.7	< 2.5	—
Type	AK4CHS	AK4CHS	PFMET
ID	tight jet ID	medium deepjet ID	—

back

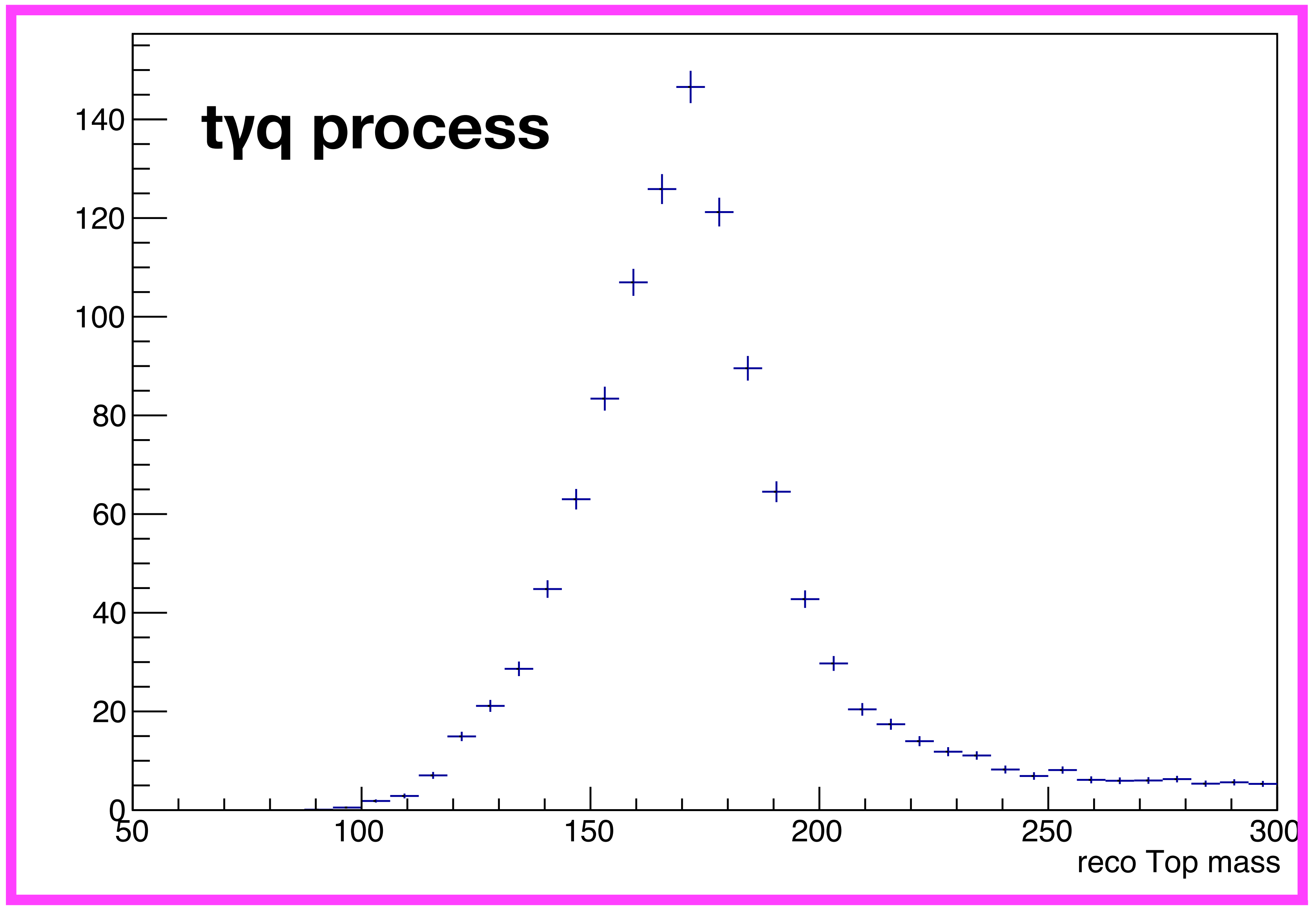
Top reconstruction

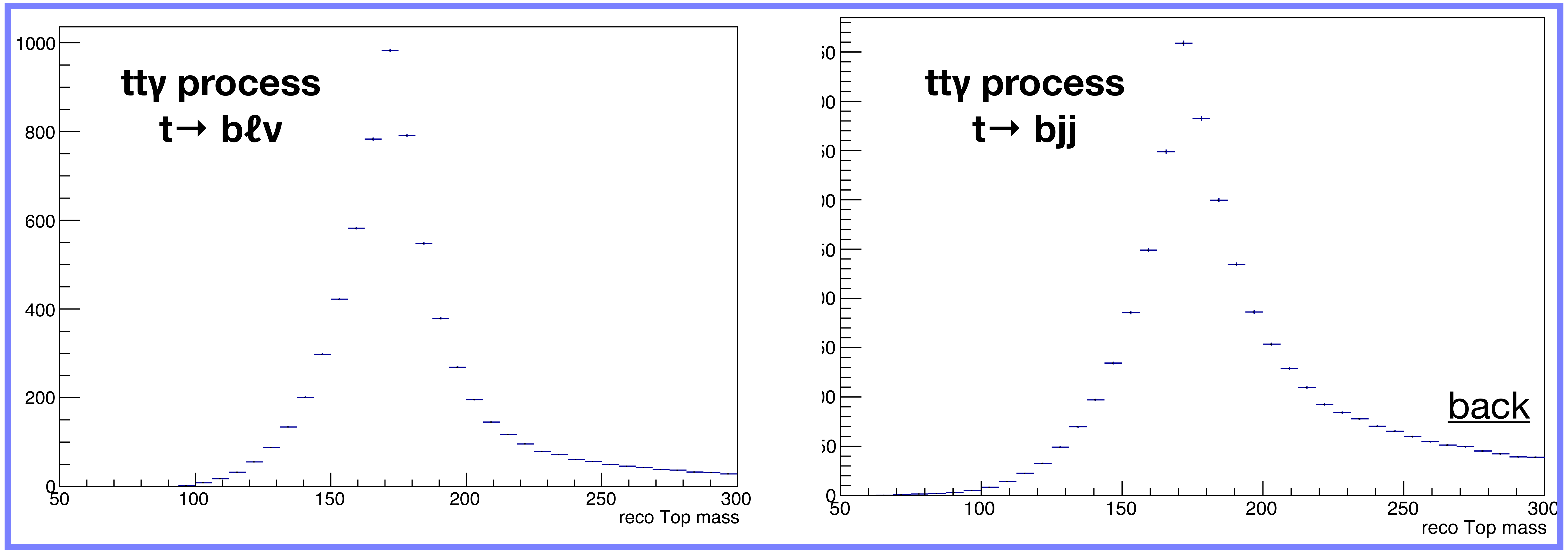
- Chi-square minimisation is performed
- Three different cases are considered:
  - 2 jets with 1 b-jets: only leptonic top is constructed
  - 3 jets with ≥ 1 b-jets: both hadronic and leptonic top are constructed with their lowest χ^2
  - ≥ 4 jets with ≥ 1 b-jets: both hadronic and leptonic top are constructed
- If the reconstruction is not possible, give a default value -10

$$\chi^2_{\mathrm{t,lep}} = \left(\frac{m_{\ell\nu \mathrm{b}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,lep}}}\right)^2$$

$$\chi^2_{\mathrm{t,had}} = \left(\frac{m_{\mathrm{bjj}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,had}}}\right)^2$$

$$\chi^2_{\mathrm{t}} = \left(\frac{m_{\ell\nu \mathrm{b}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,lep}}}\right)^2 + \left(\frac{m_{\mathrm{bjj}} - m_{\mathrm{t}}}{\sigma_{\mathrm{t,had}}}\right)^2$$





Top dead cone

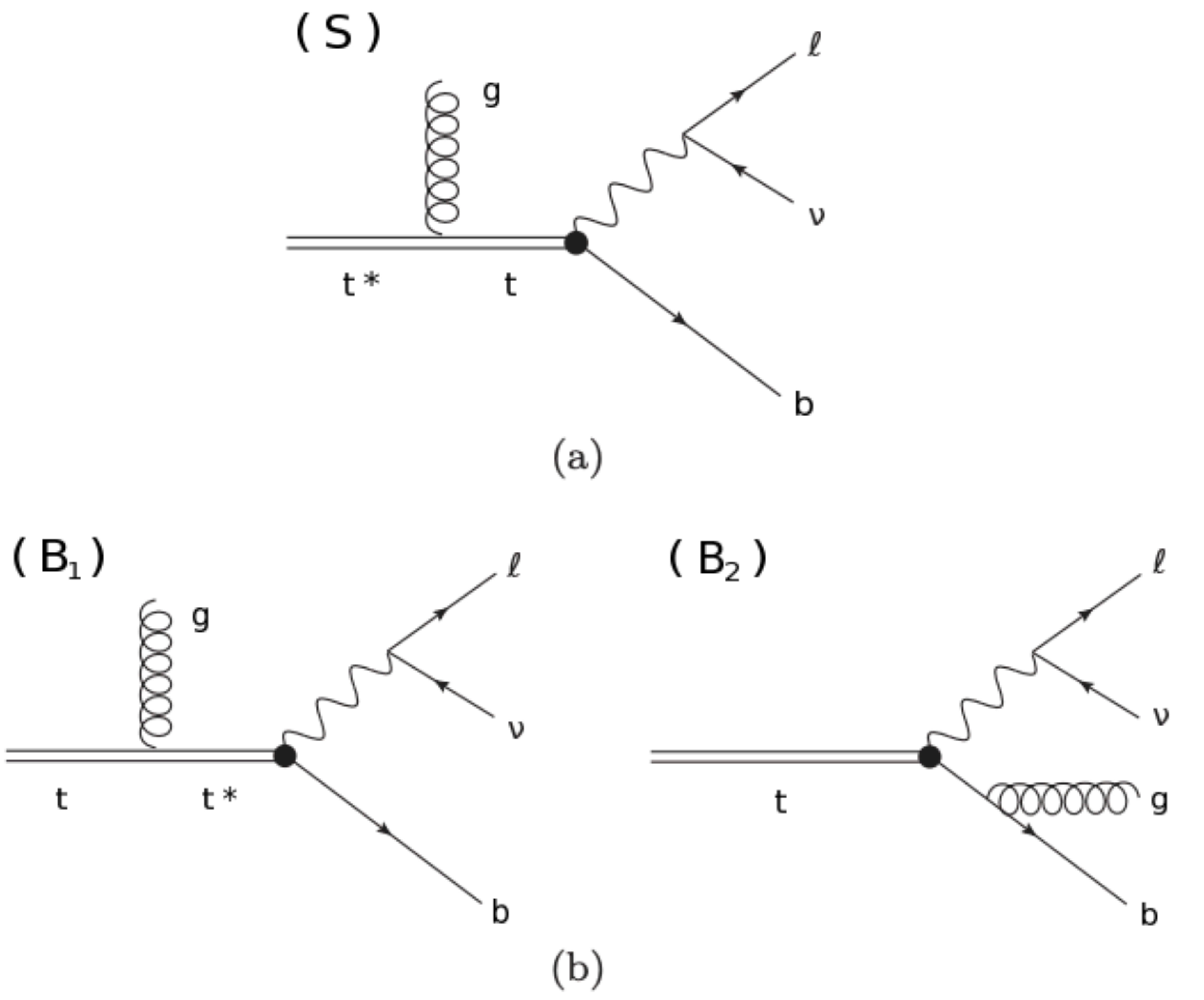
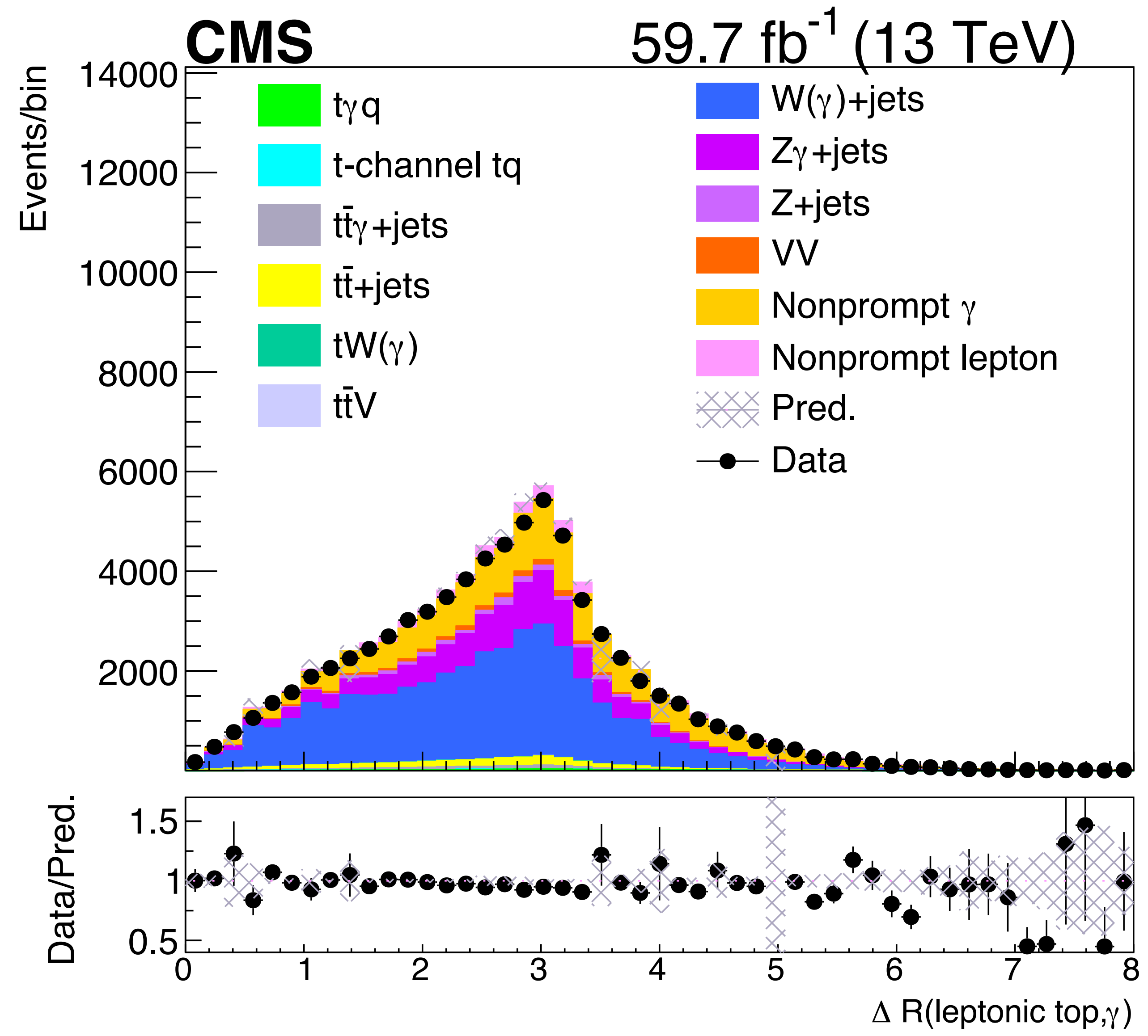
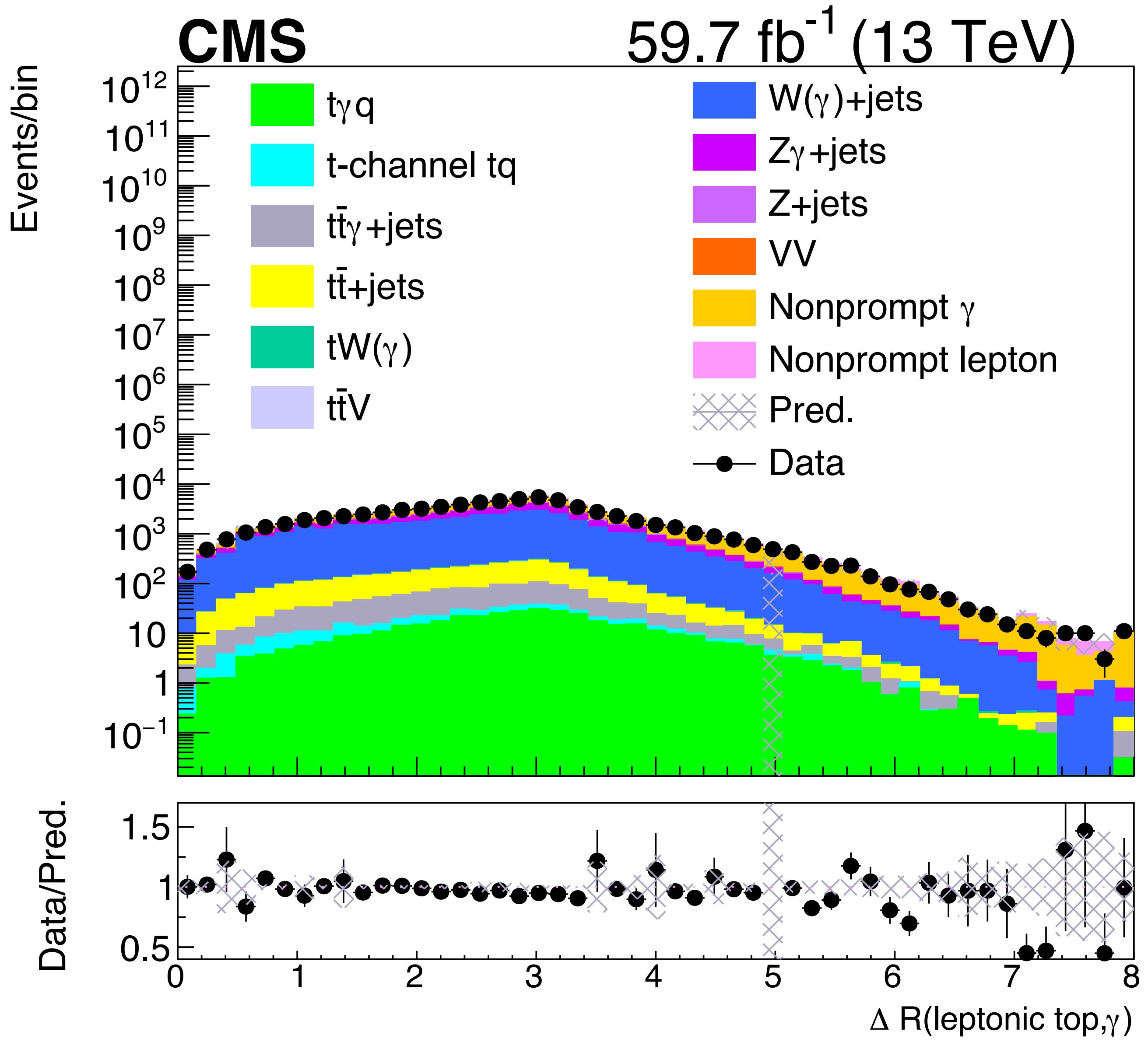


FIG. 4. Feynman diagrams for gluon radiation in (a) the signal process of top FSR $t^* \to tg$ and (b) the background process of top decay $t \to bWg$.

https://arxiv.org/pdf/1606.03449.pdf

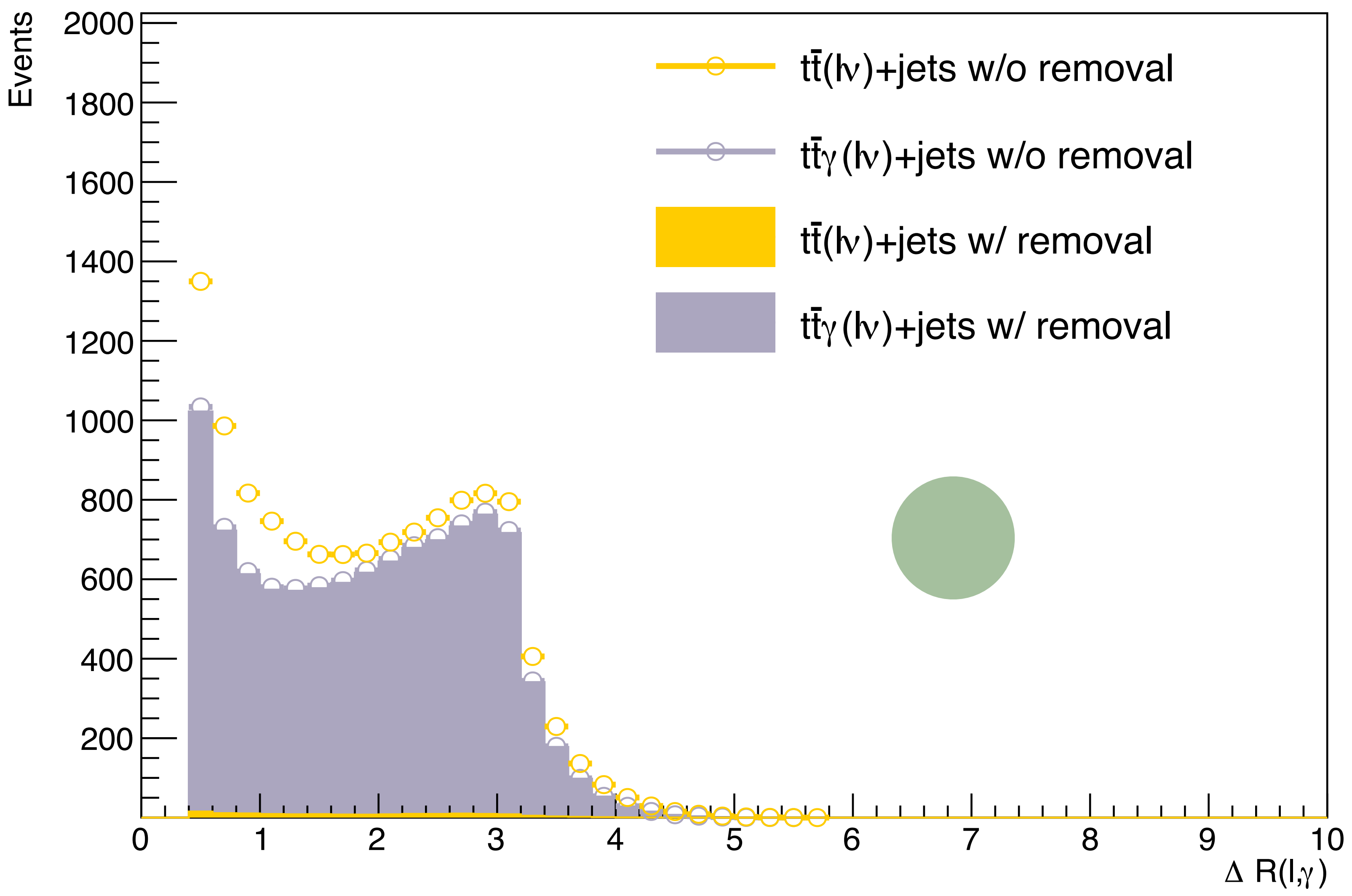


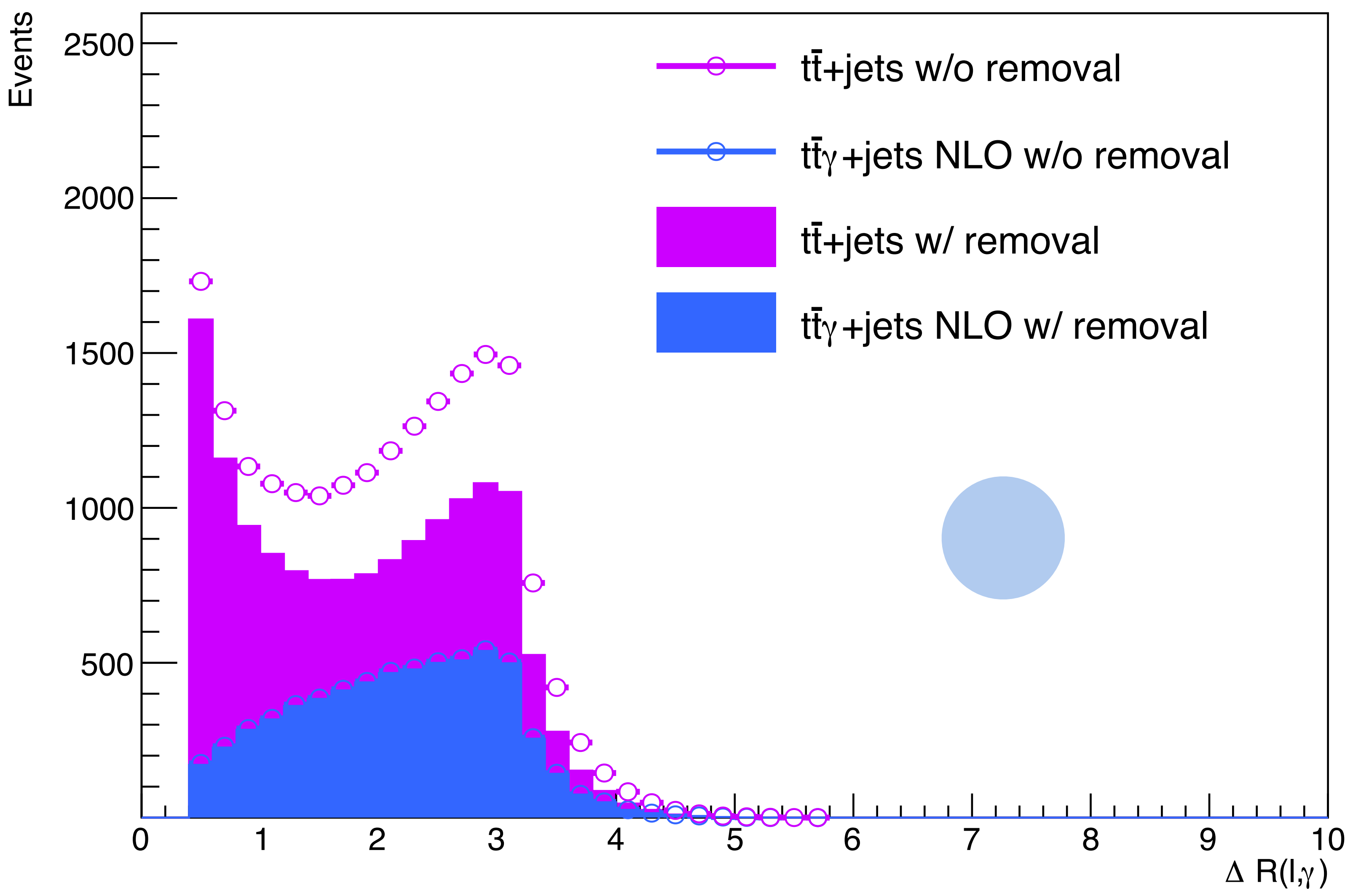


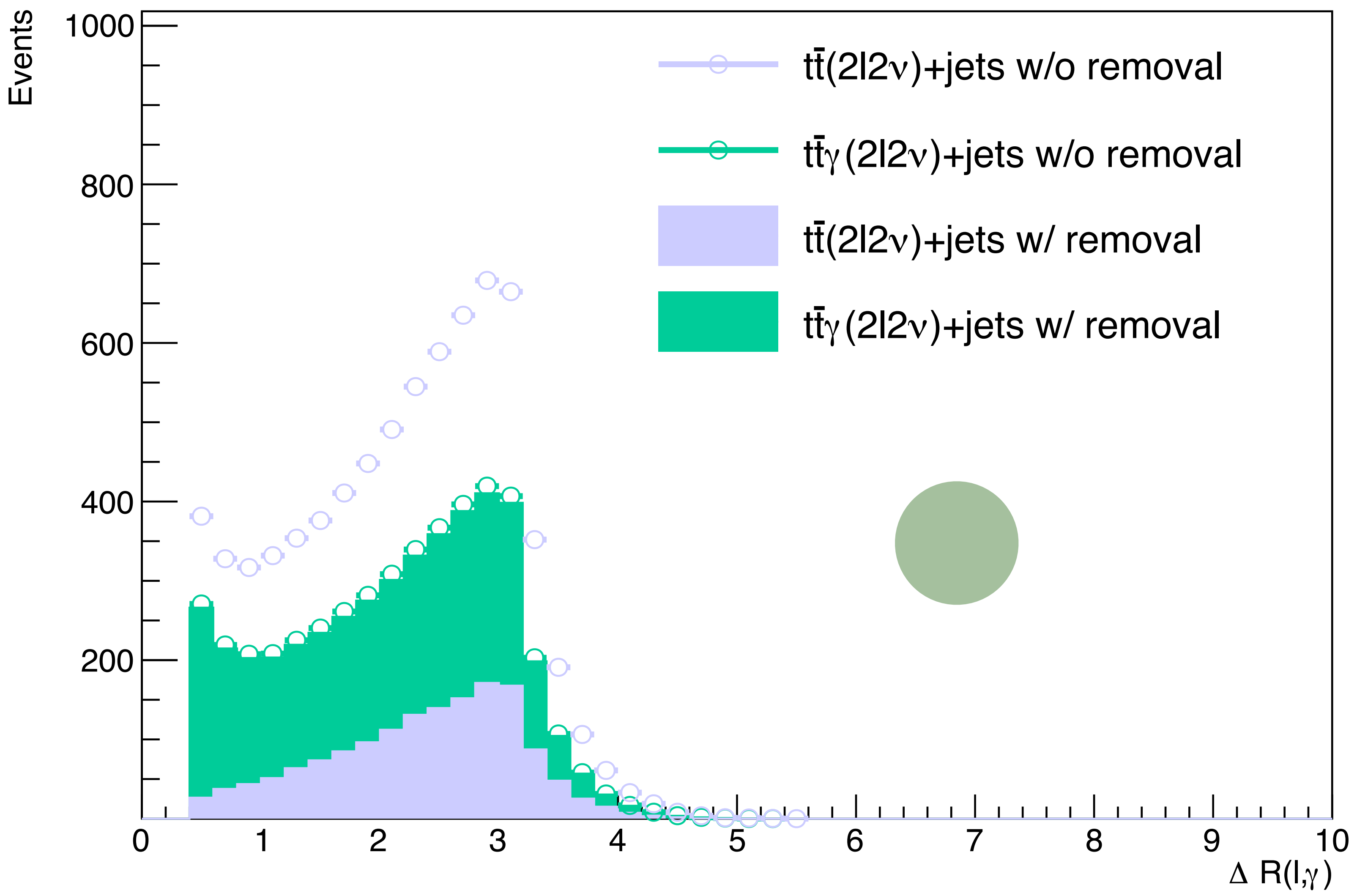
back

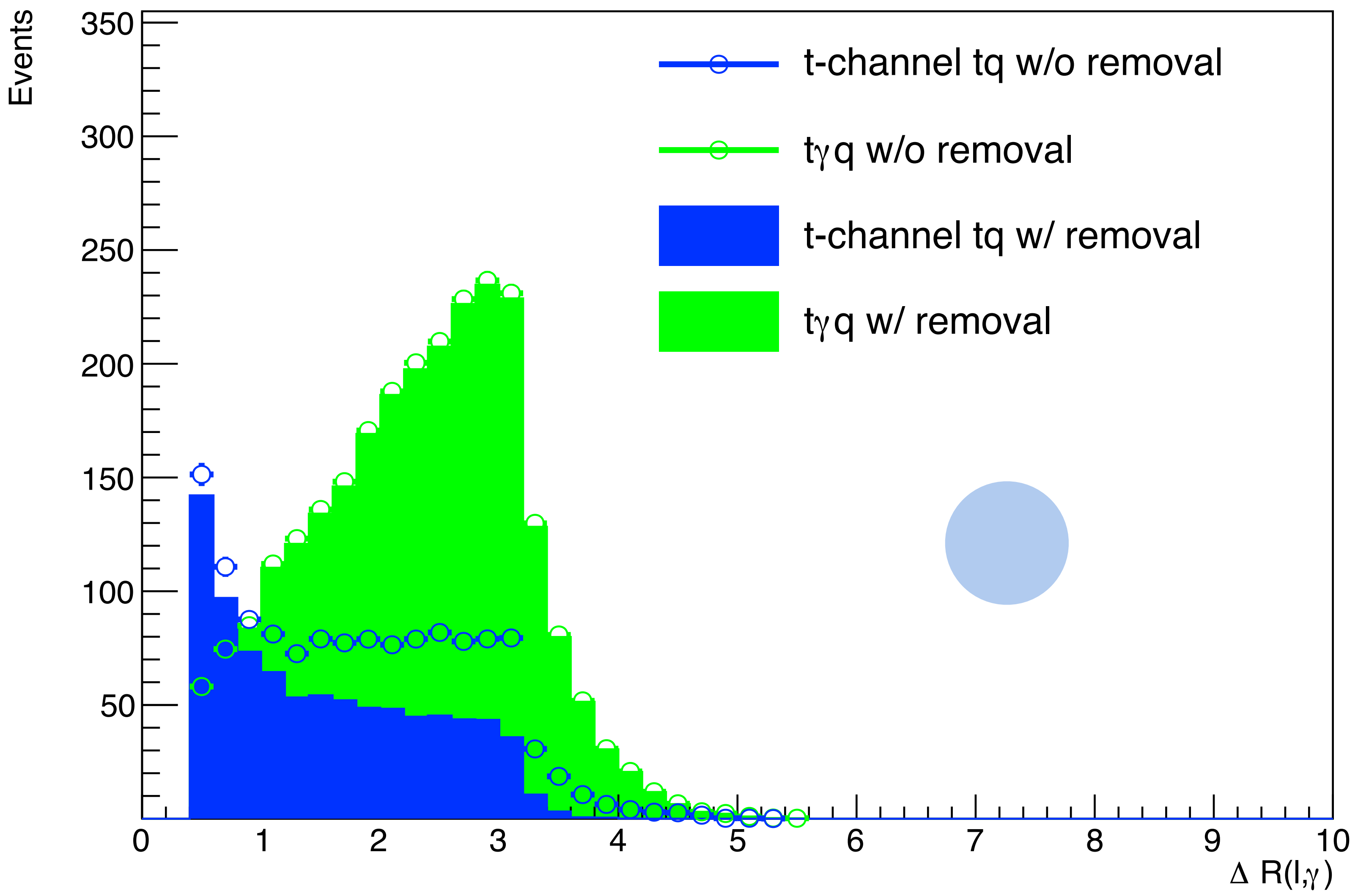
Distributions in X(γ)+jets — w/ and w/o removal

- **Reconstruction level** with only lepton and photon selection
- $N_{\ell}=1$, $N_{\gamma} \geq 1$, pass removal requirement









Removal procedure applied to

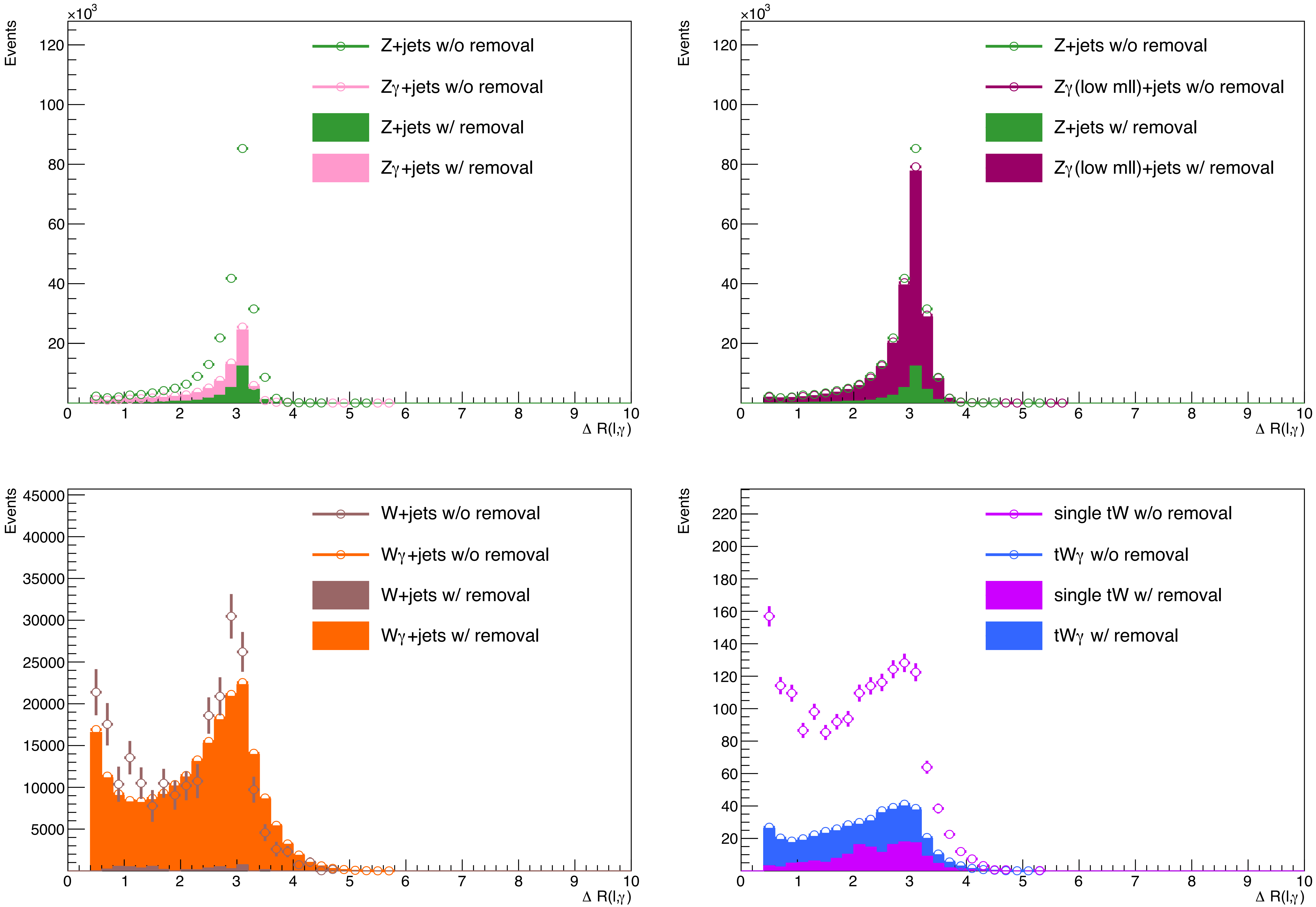
- NLO $t\bar{t}\gamma$ and $t\bar{t}$+jets($\ell\nu$+$2\ell2\nu$)
- LO $t\bar{t}\gamma$ ($\ell\nu$) and $t\bar{t}$+jets($\ell\nu$)
- LO $t\bar{t}\gamma$ ($2\ell2\nu$) and $t\bar{t}$+jets($2\ell2\nu$)
- Signal tγq + single top t-channel

After overlap removal,

1. If the Xγ+jets sample is simulated in a such completed phase space, the X+jets remaining few
2. If the Xγ+jets sample lacks dedicated production mode, the X+jets remaining some

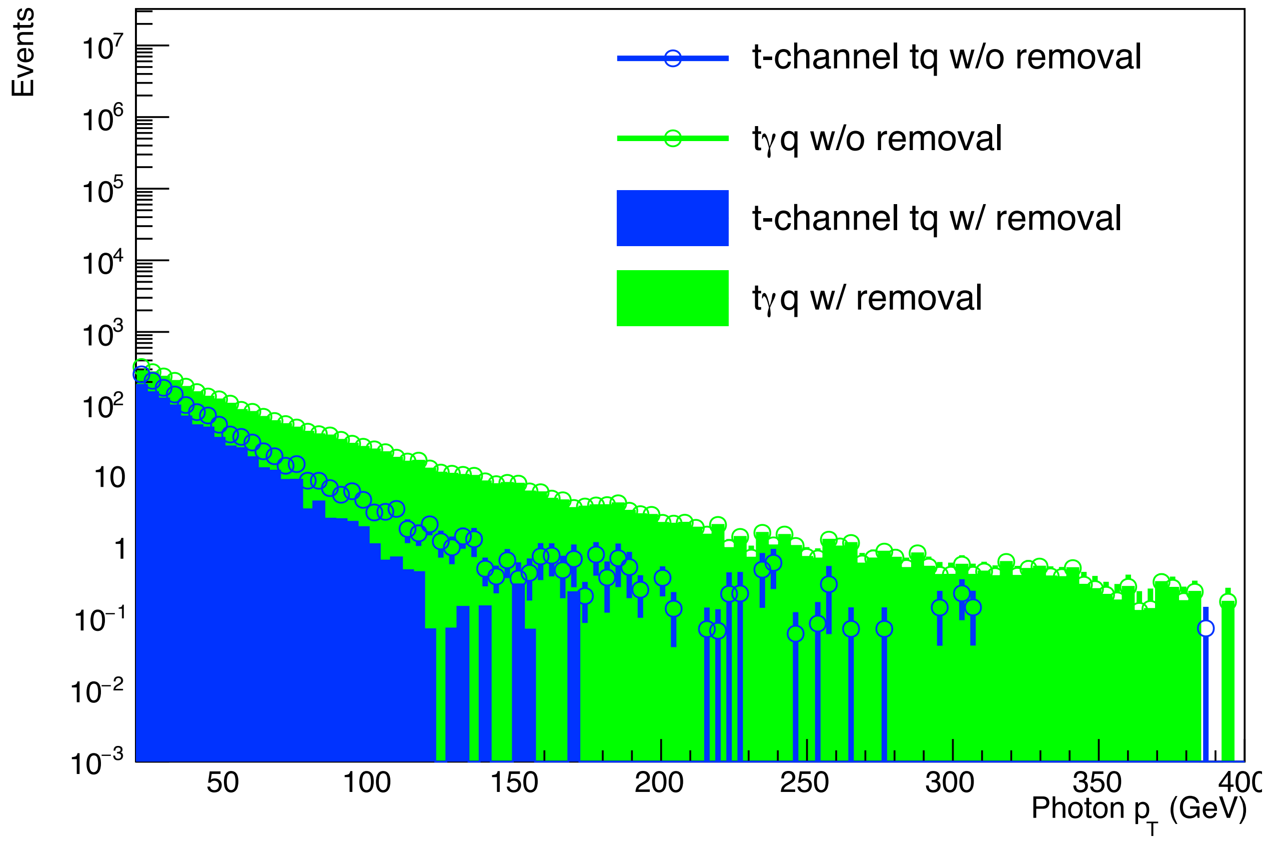
back More plots

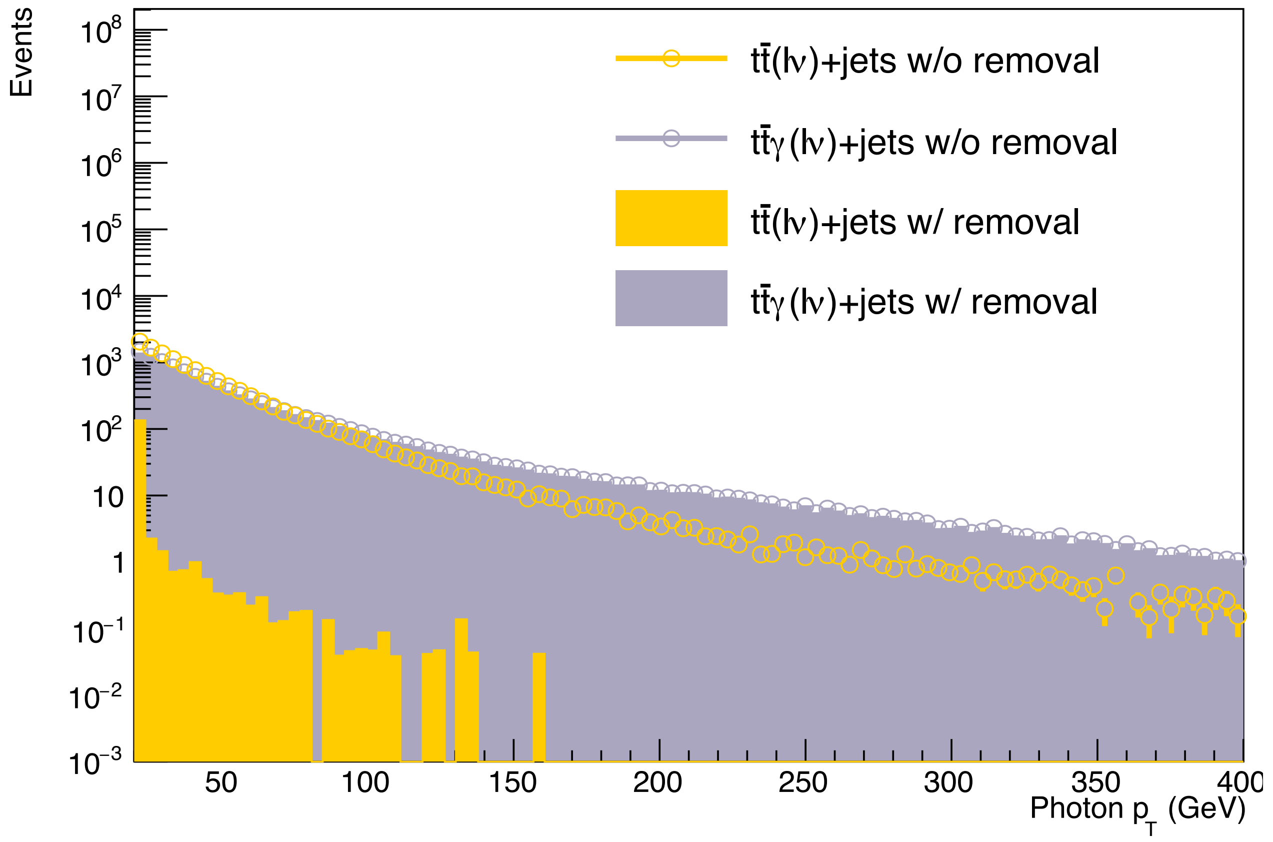
Distributions after removal

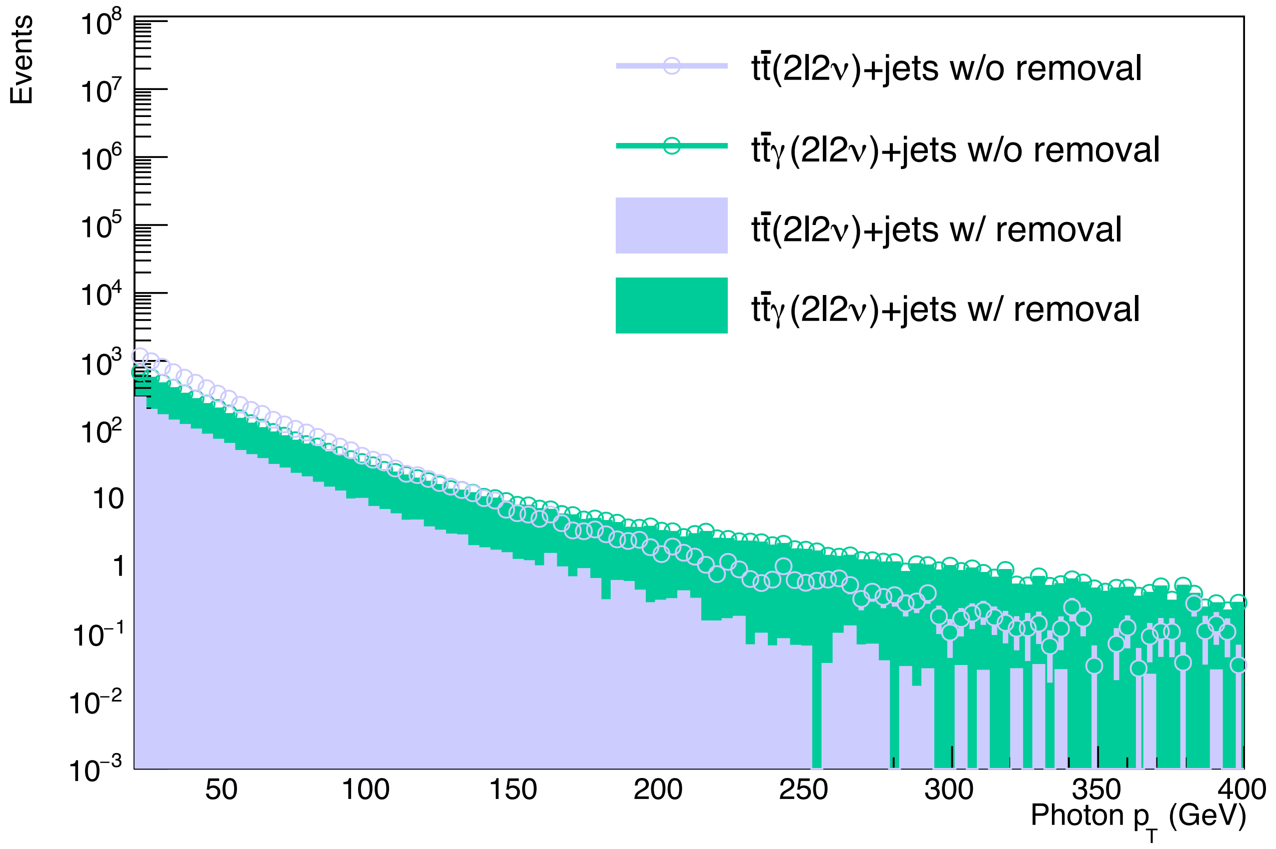


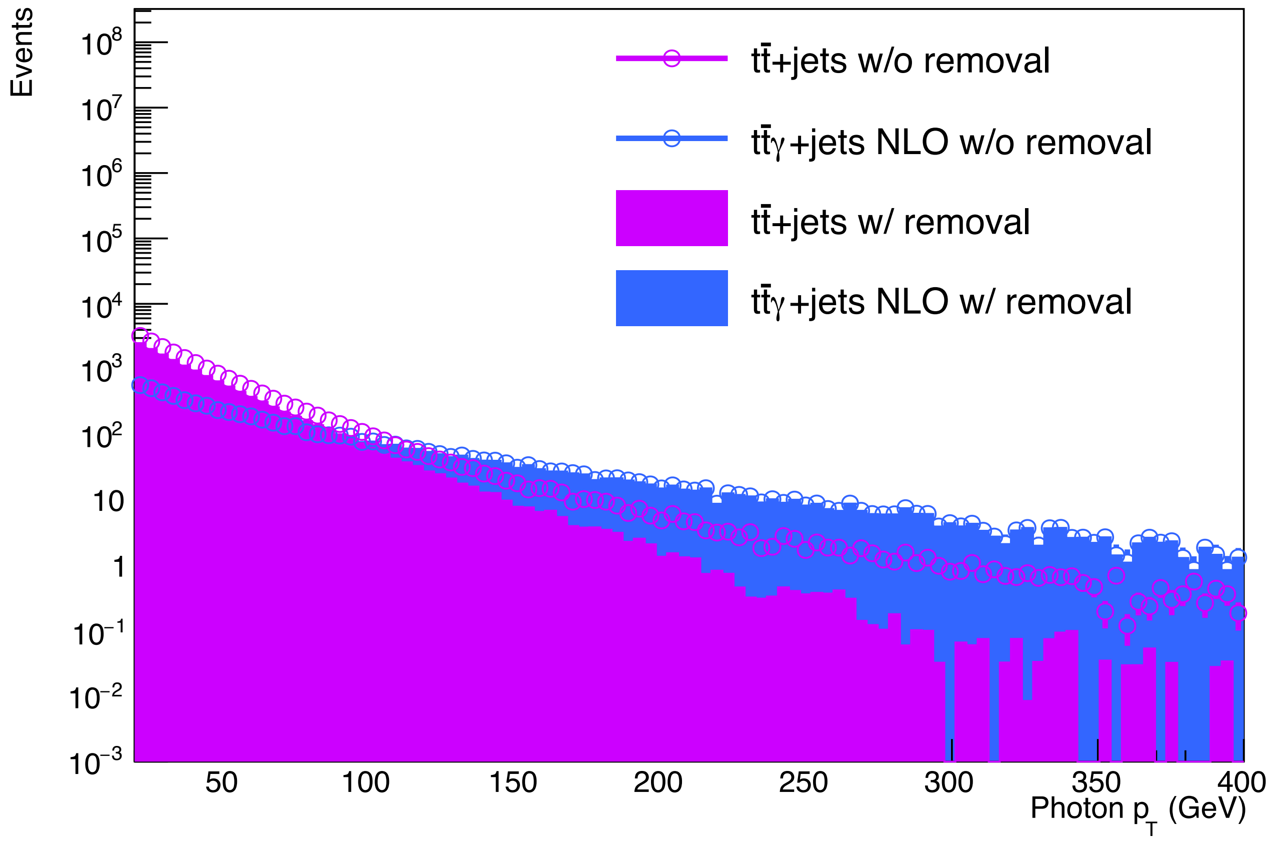
[back](#)

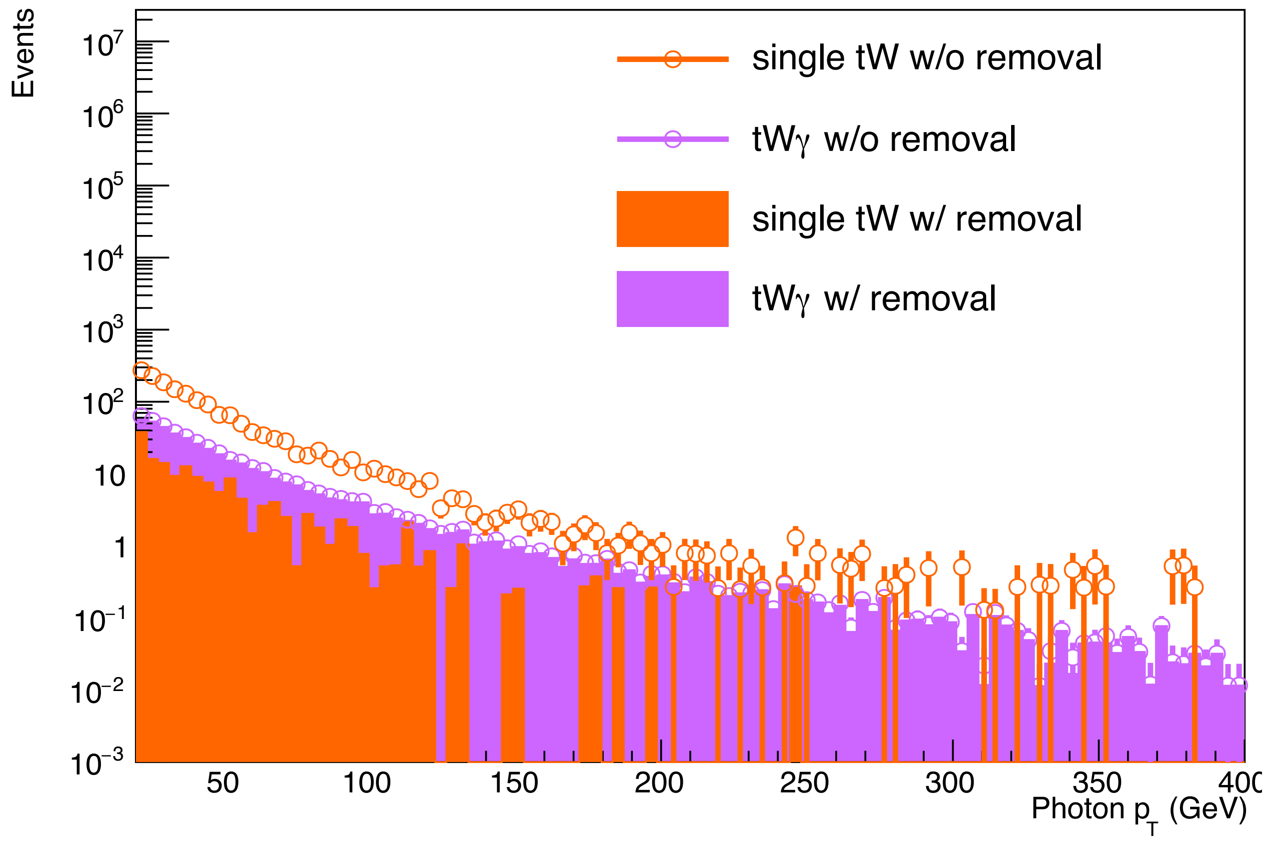
Distributions after removal

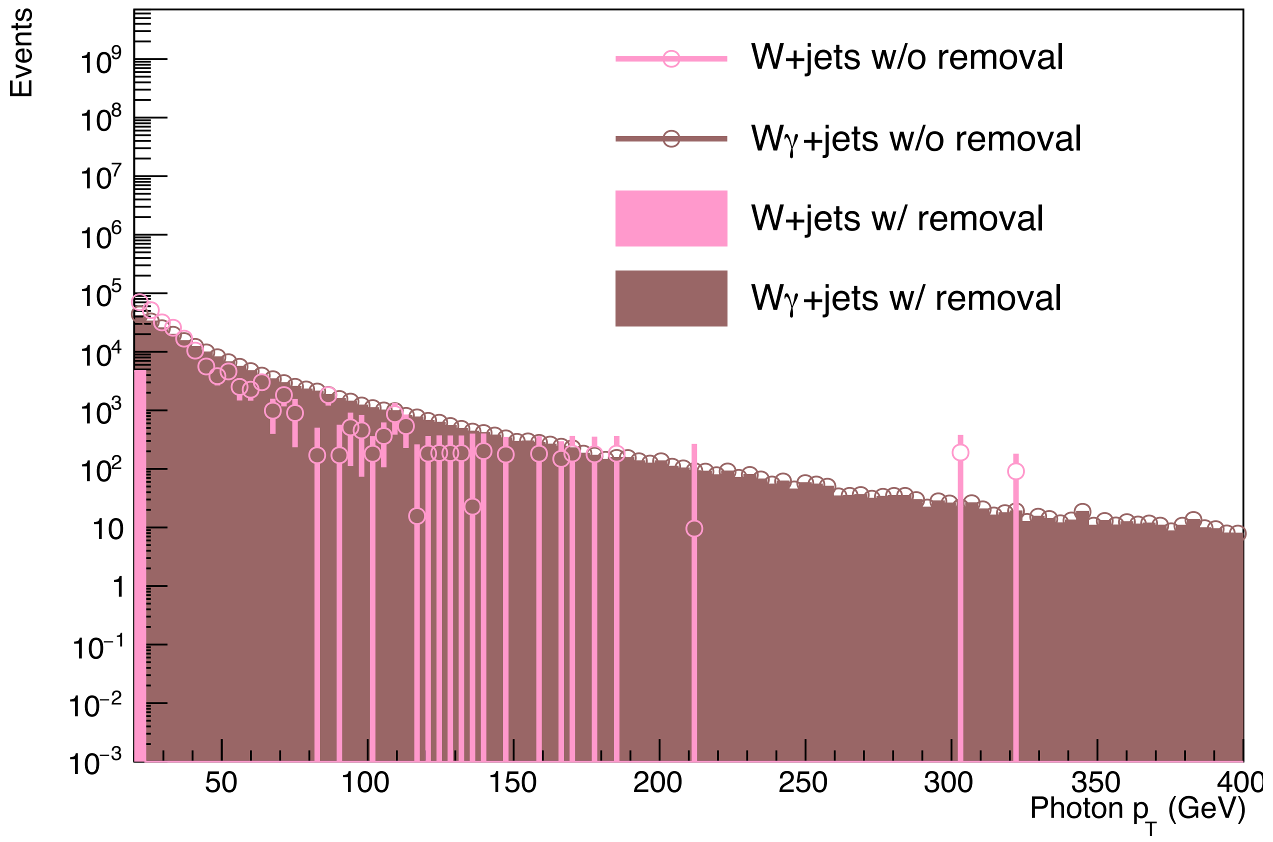


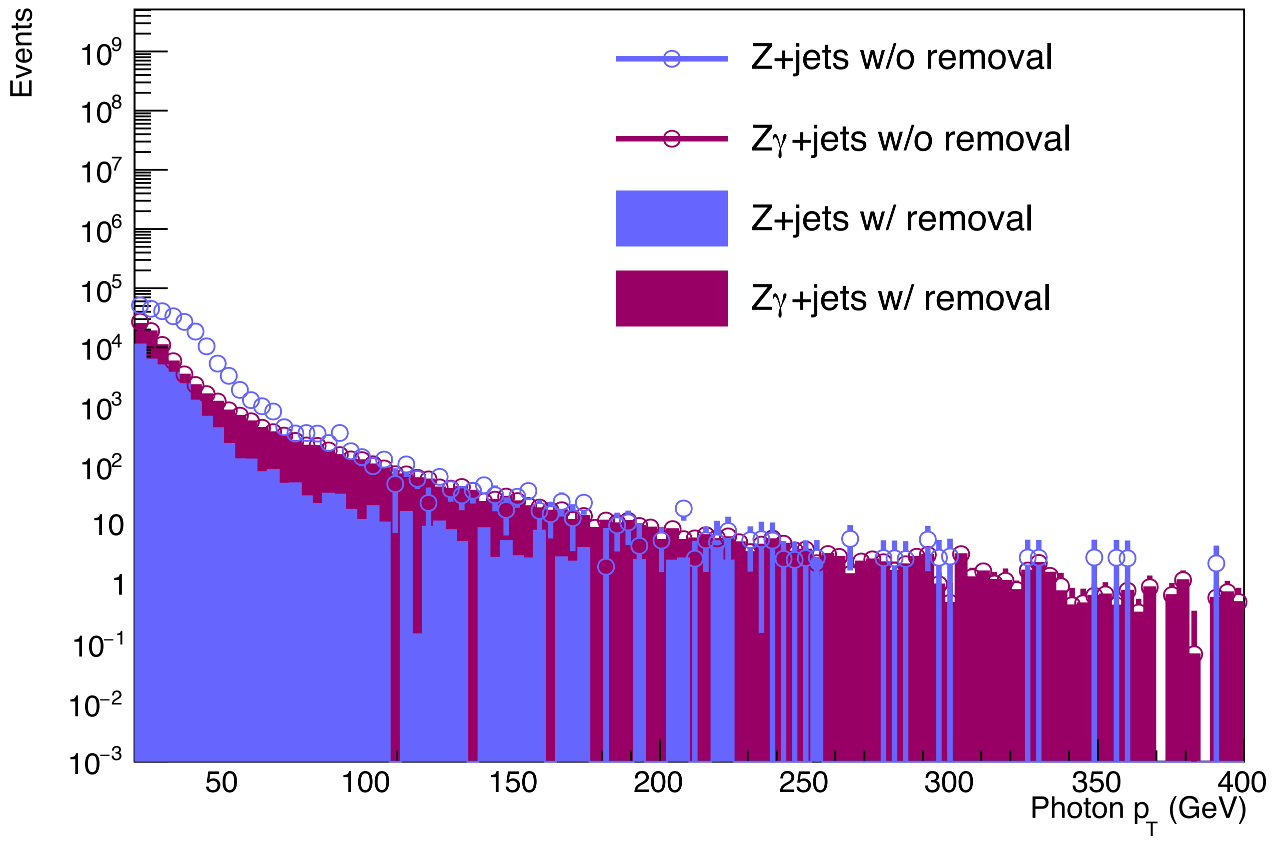


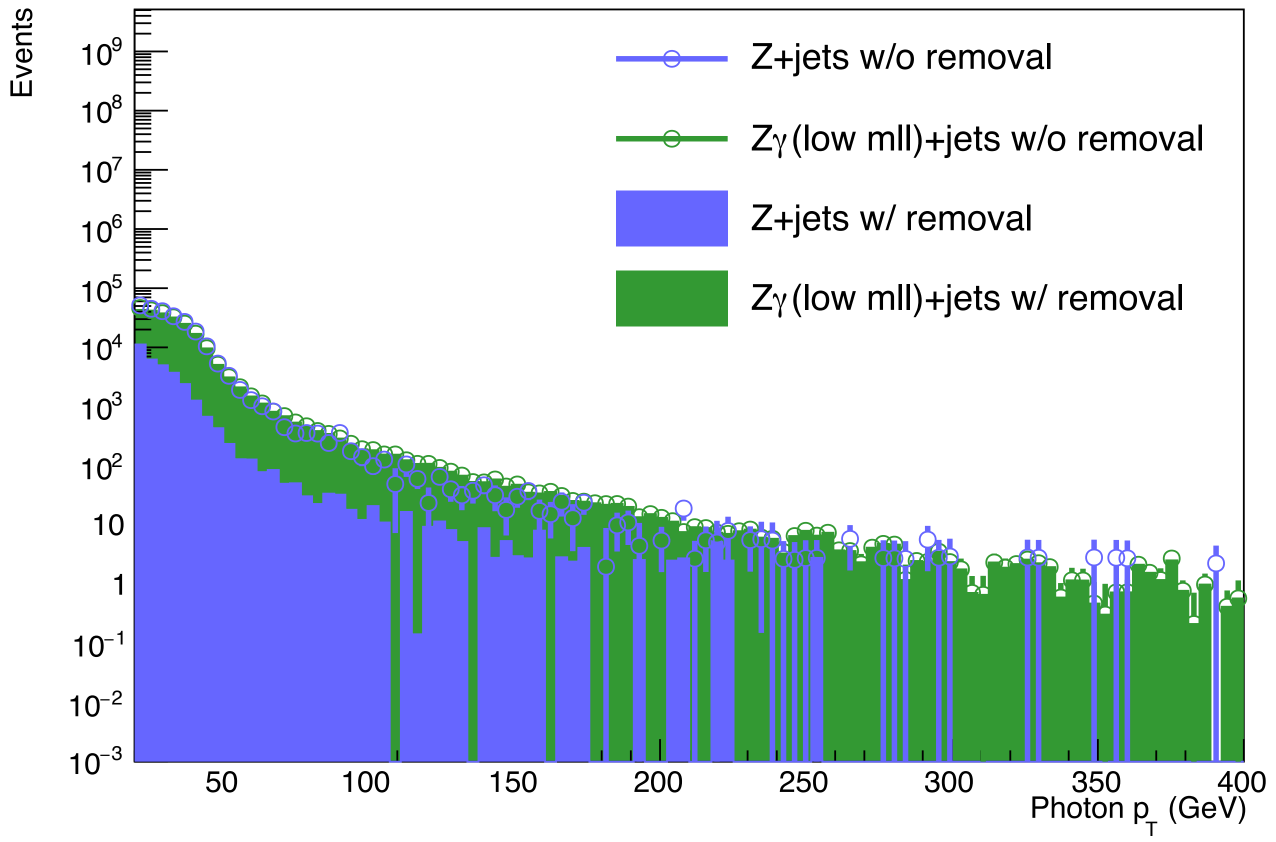












back

Distributions after removal

back

Signal tγq sample check

- Syntax like
  - p p > z, z > e+ e- (on-shell z decaying)
  - p p > e+ e- $ z (forbids s-channel z to be on-shell)
- Are linked to cut $|M^* - M| < BW_{cut} * \Gamma$
- Are more safer to use
- Prefer those syntax to the previous slides one

- Syntax Like
  - ➡ p p > z > e+ e- (ask one S-channel z)
  - ➡ p p > e+ e- / z (forbids any z)
  - ➡ p p > e+ e- $$ z (forbids any z in s-channel)
  - ARE NOT GAUGE INVARIANT !
  - forgets diagram interference.
  - can provides un-physical distributions.

1. generate **p p > t b~ j a**, (t > w+ b, w+ > l+ vl)

   \# 1 processes with 291 diagrams generated
   \# **original cross-section: 0.272 pb**
   \#   scale variation: +19.3% -15.2%
   \#   central scheme variation: + 0% -43%
   \# PDF variation: +2.56e-307% -2.55e-307%

2. generate **p p > t b~ j a $ w+ w-**, (t > w+ b, w+ > l+ vl)

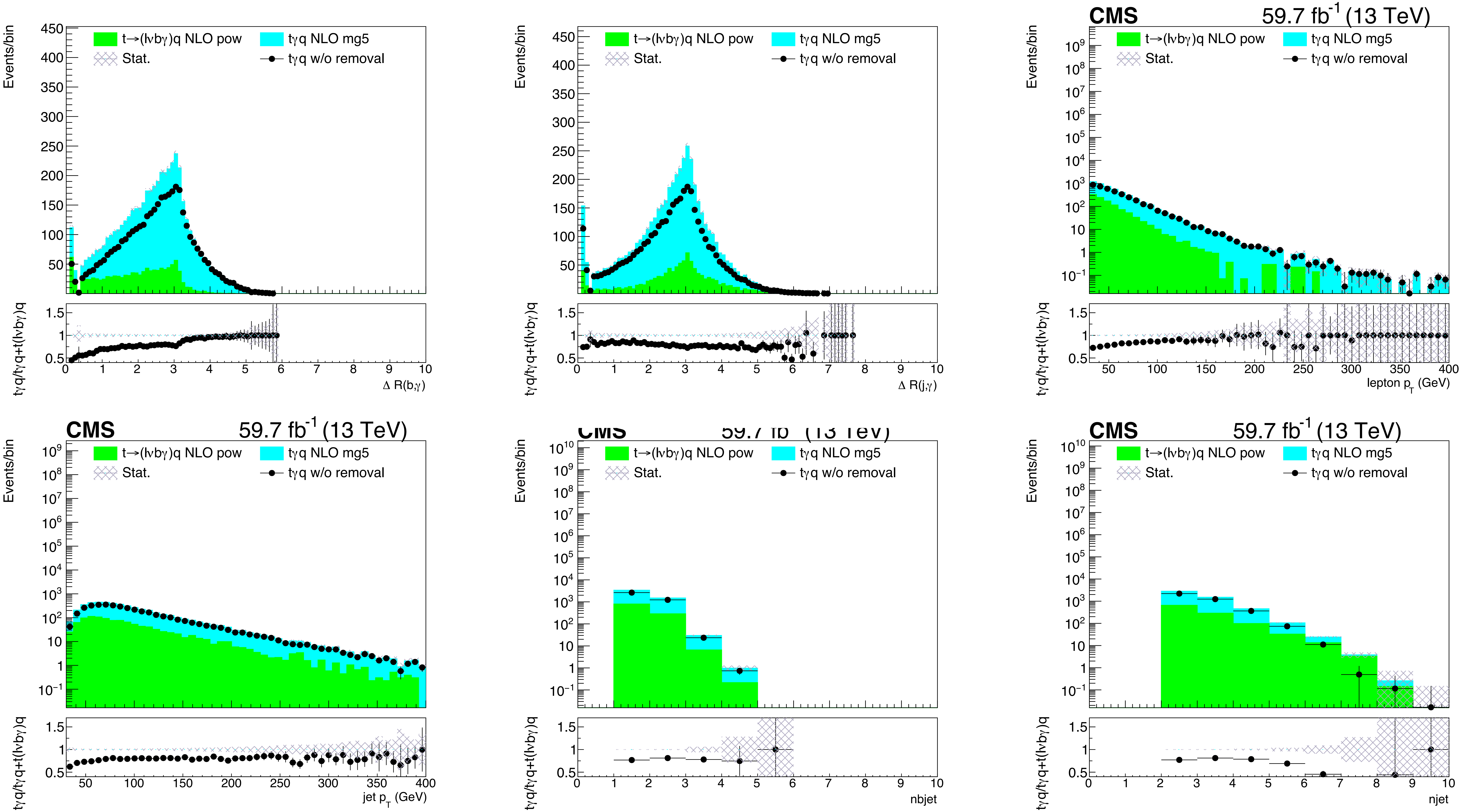
   \# 1 processes with 291 diagrams generated
   \# **original cross-section: 0.271 pb**
   \#   scale variation: +19.2% -15.2%
   \#   central scheme variation: + 0% -43%
   \# PDF variation: +2.56e-307% -2.55e-307%

3. generate **p p > t b~ j a $$ w+ w-**, (t > w+ b, w+ > l+ vl)

   \# 1 processes with 291 diagrams generated
   \# **original cross-section: 0.269 pb**
   \#   scale variation: +19.3% -15.3%
   \#   central scheme variation: + 0% -43.5%
   \# PDF variation: +2.58e-307% -2.58e-307%

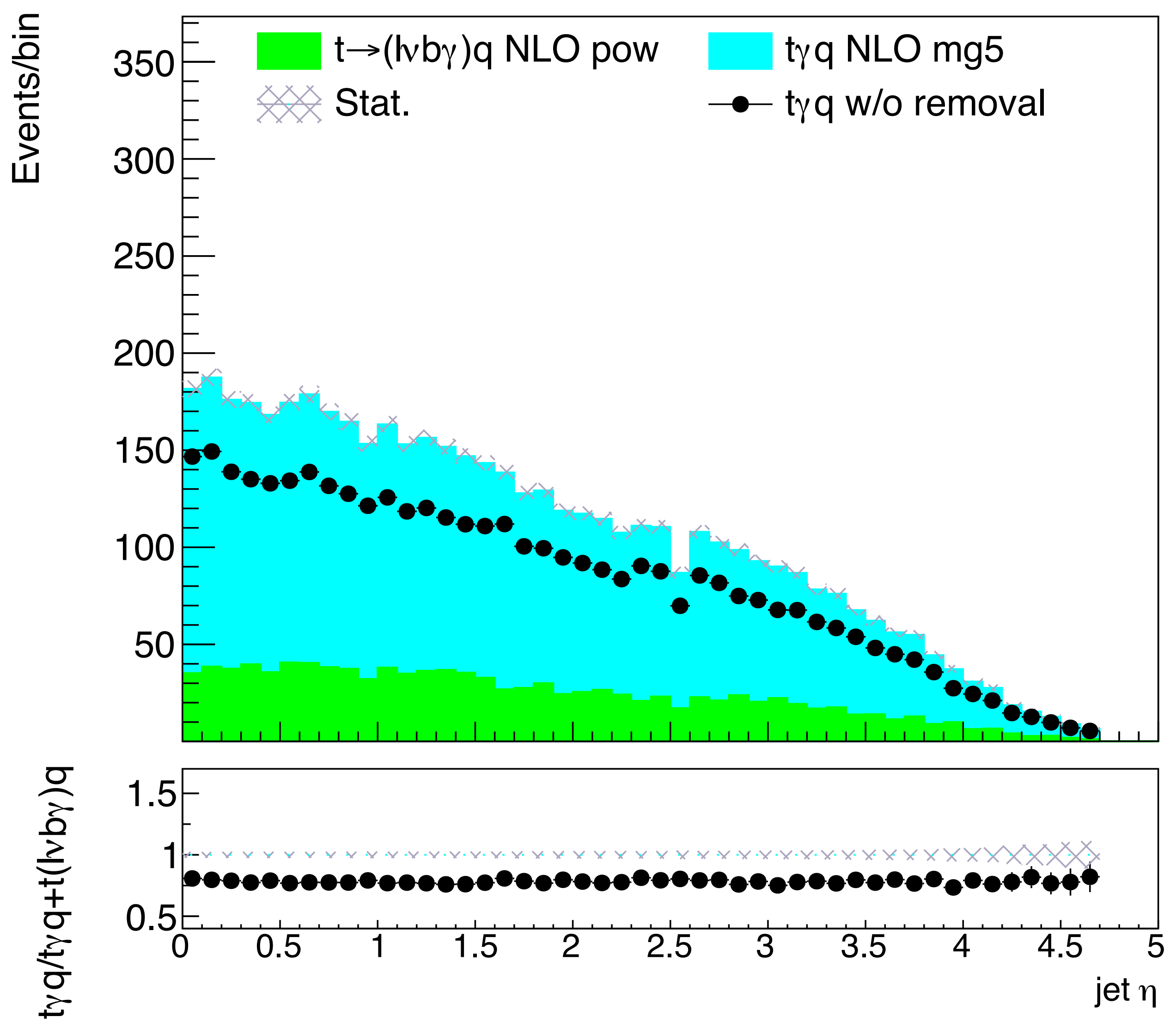
back

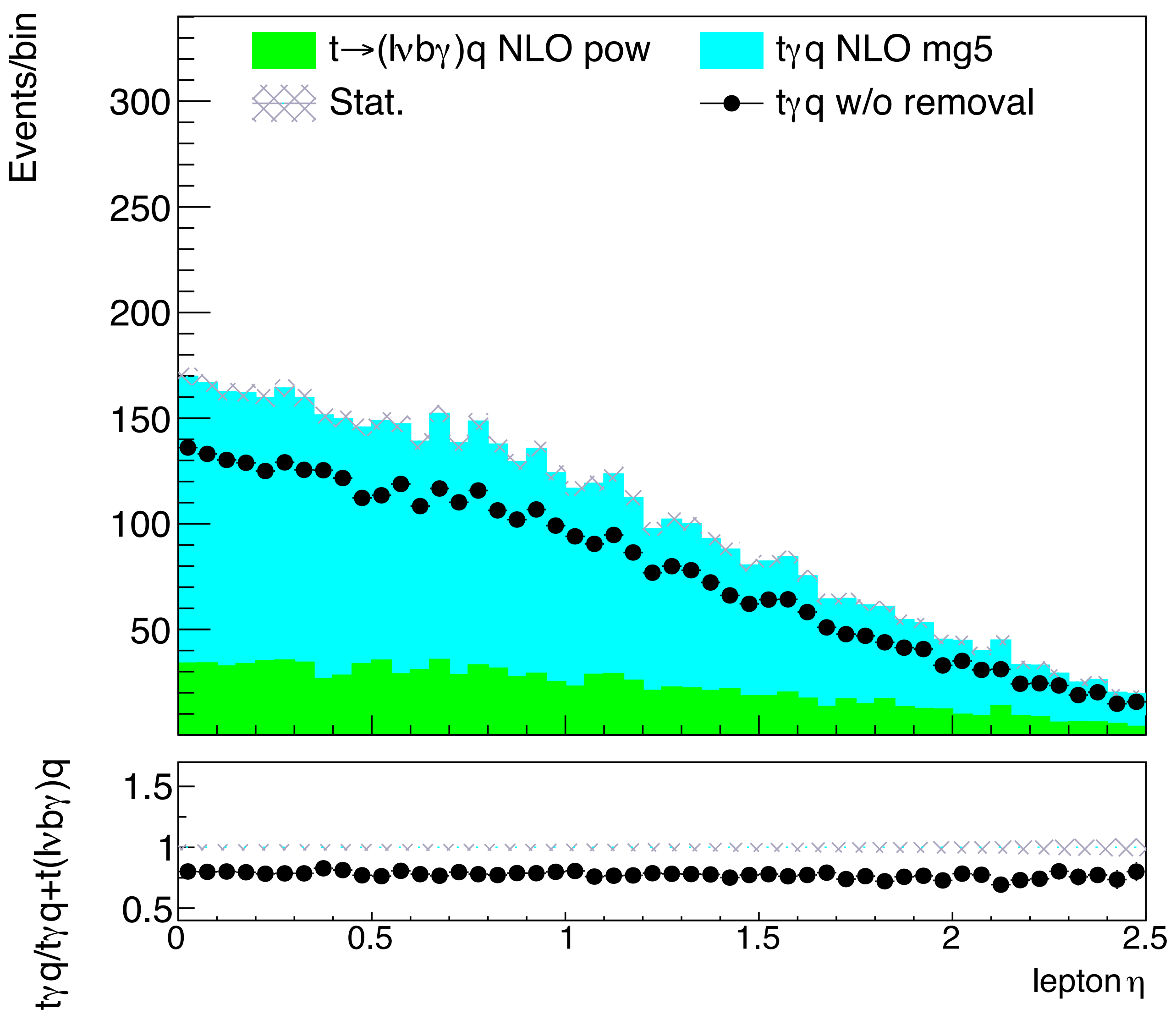
Signal *t*γq gen-level distributions

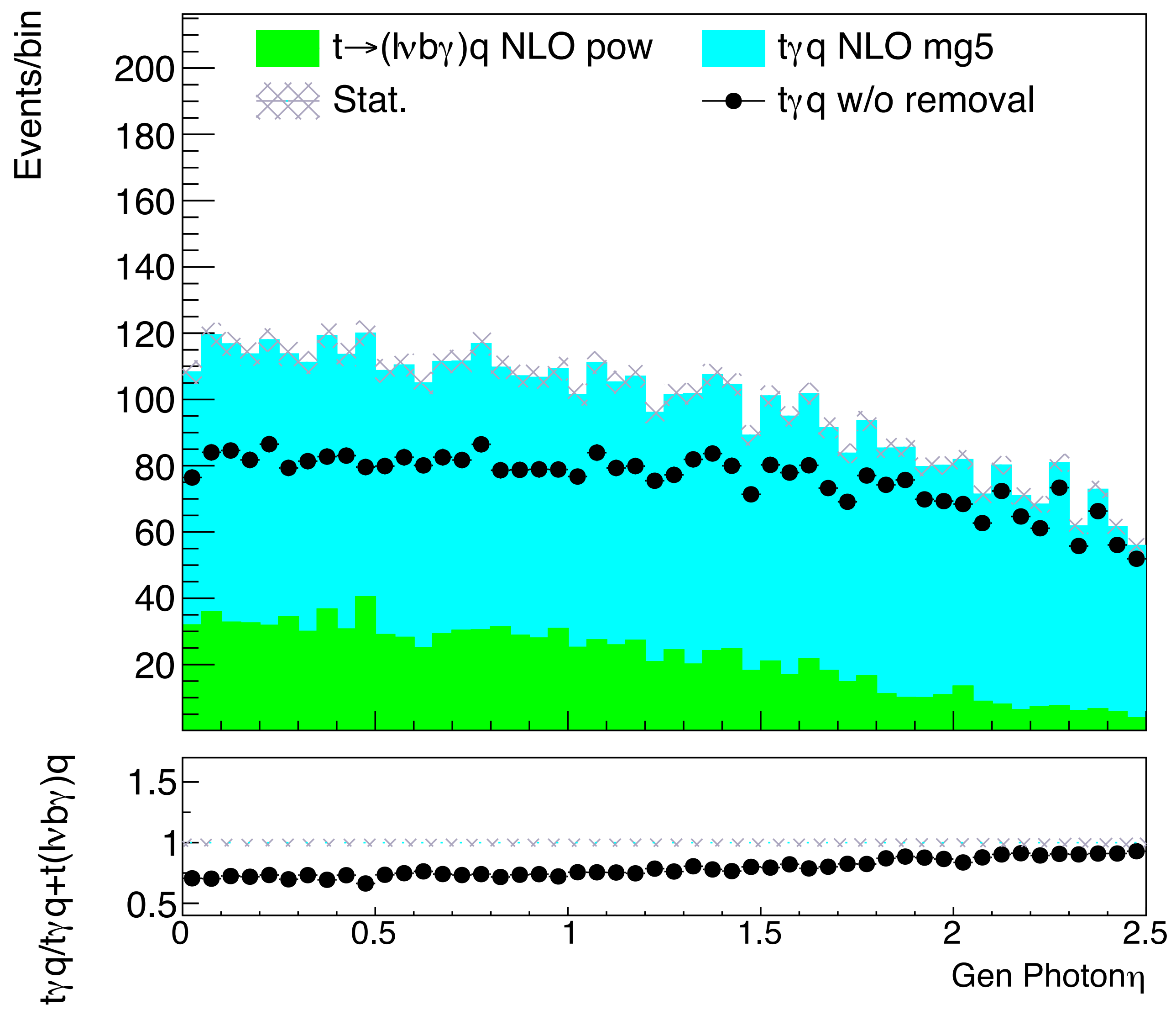


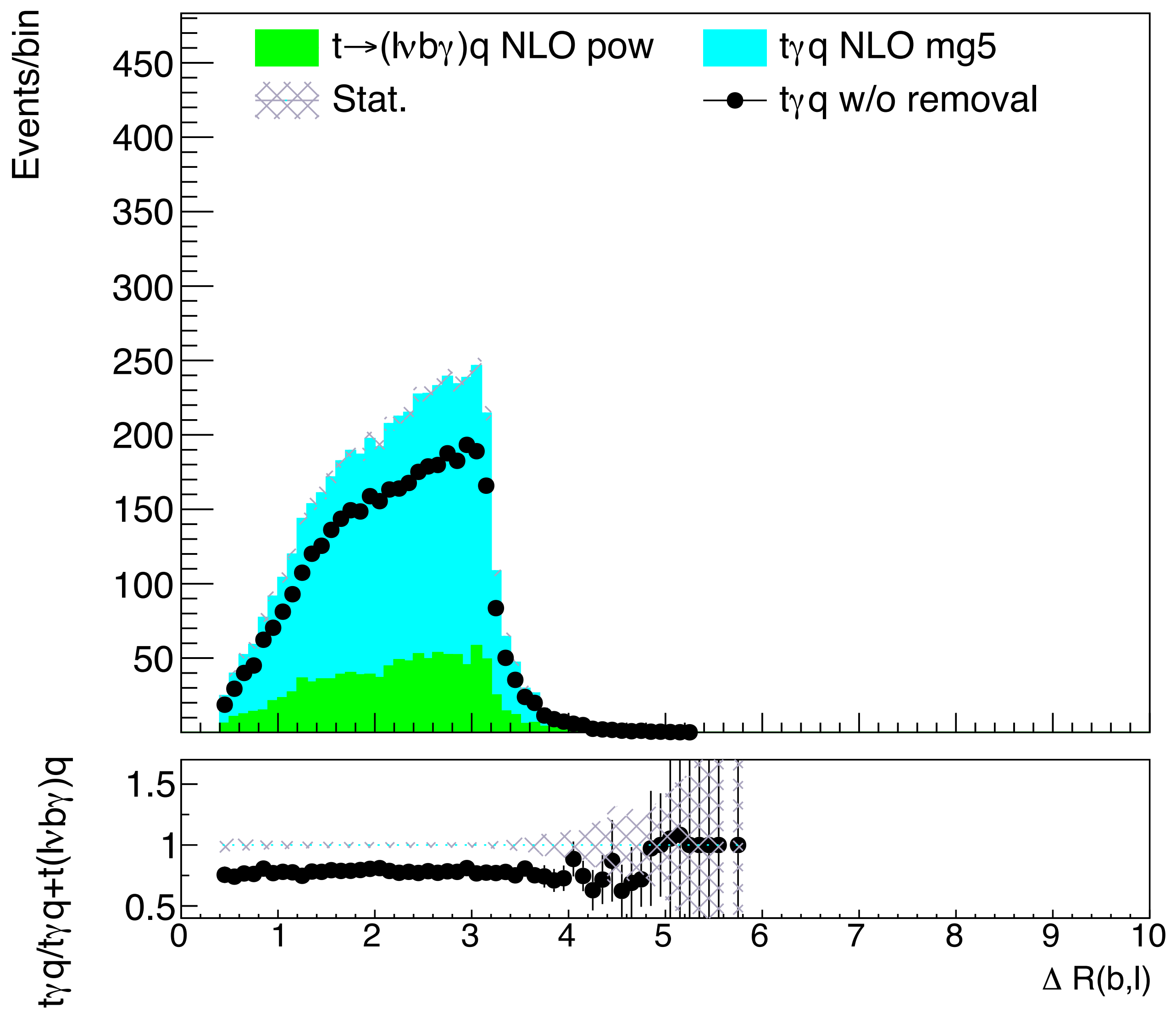
back

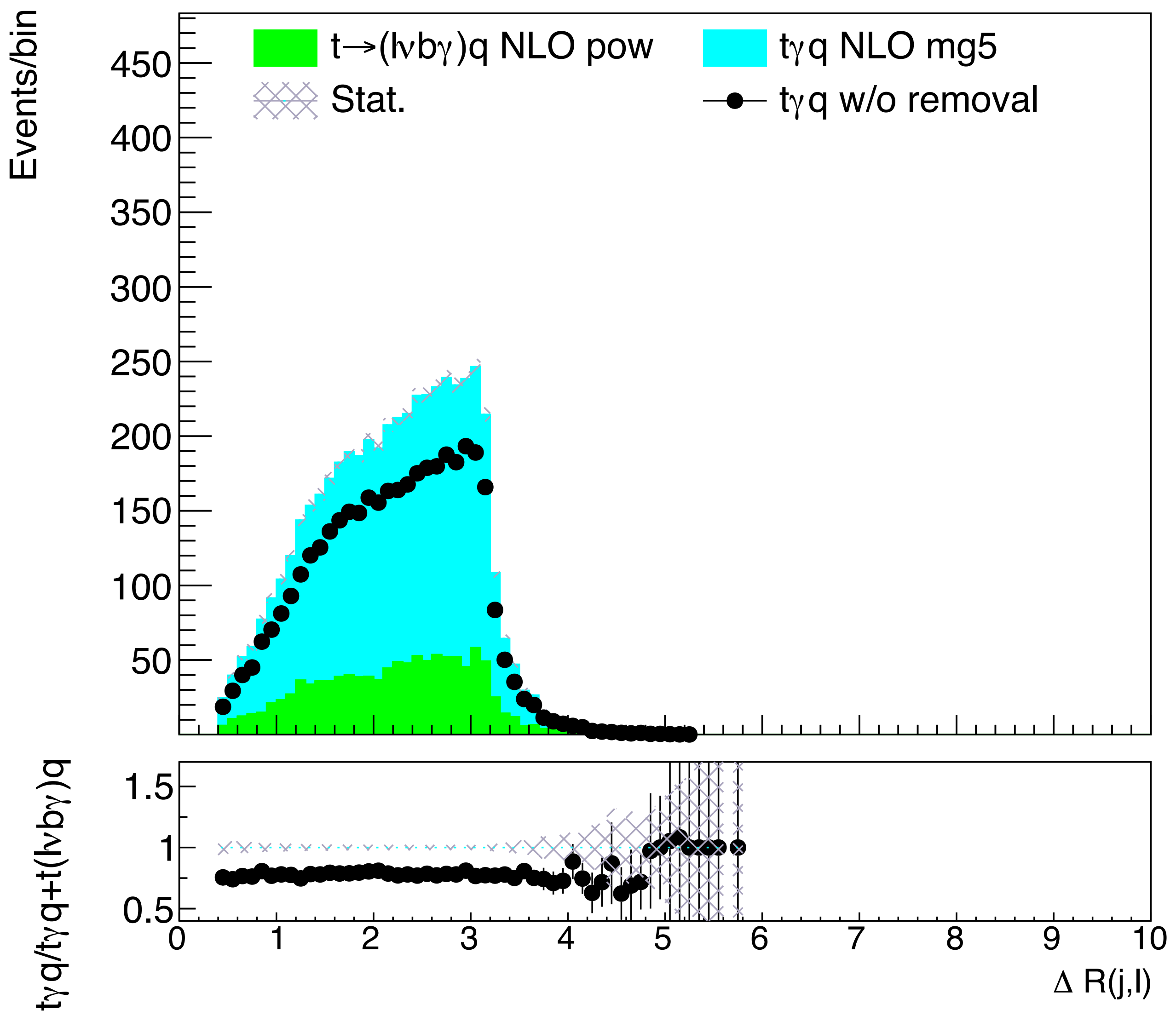
Signal $t\gamma q$ gen-level distributions











back

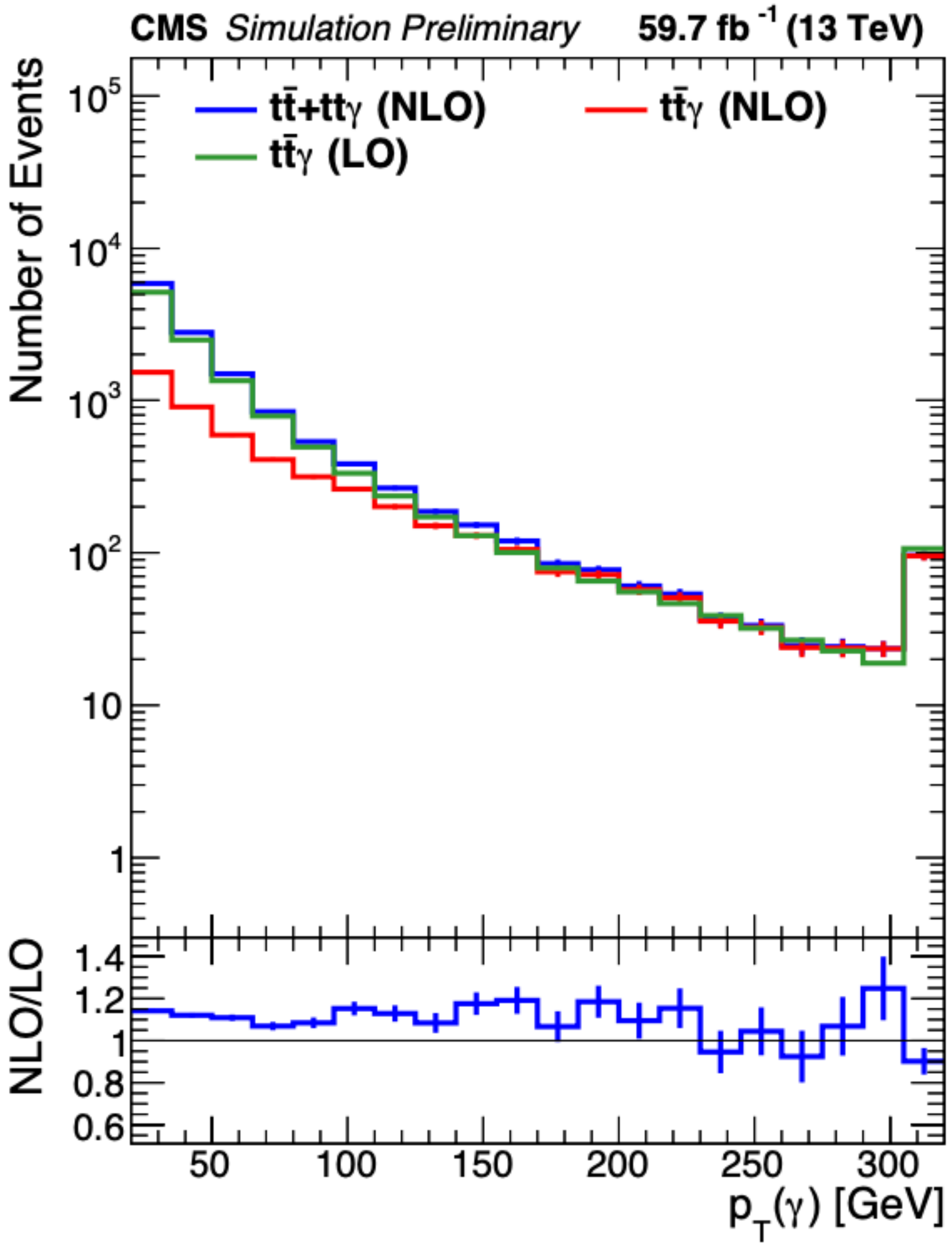
Signal simulation — $t\gamma q$

- **Reconstruction level** with SR requirements
- N_ℓ=1, $N_\gamma \geq 1$, $N_j \geq 2$, $N_b \geq 1$ and pass removal requirement

The loss due to incomplete simulation in tγq NLO sample is not large which accounts for ≈25%, mainly concentrating on low photon p_T and low $\Delta R(\ell,\gamma)$

back

$t\bar{t}\gamma$ LO k-factor



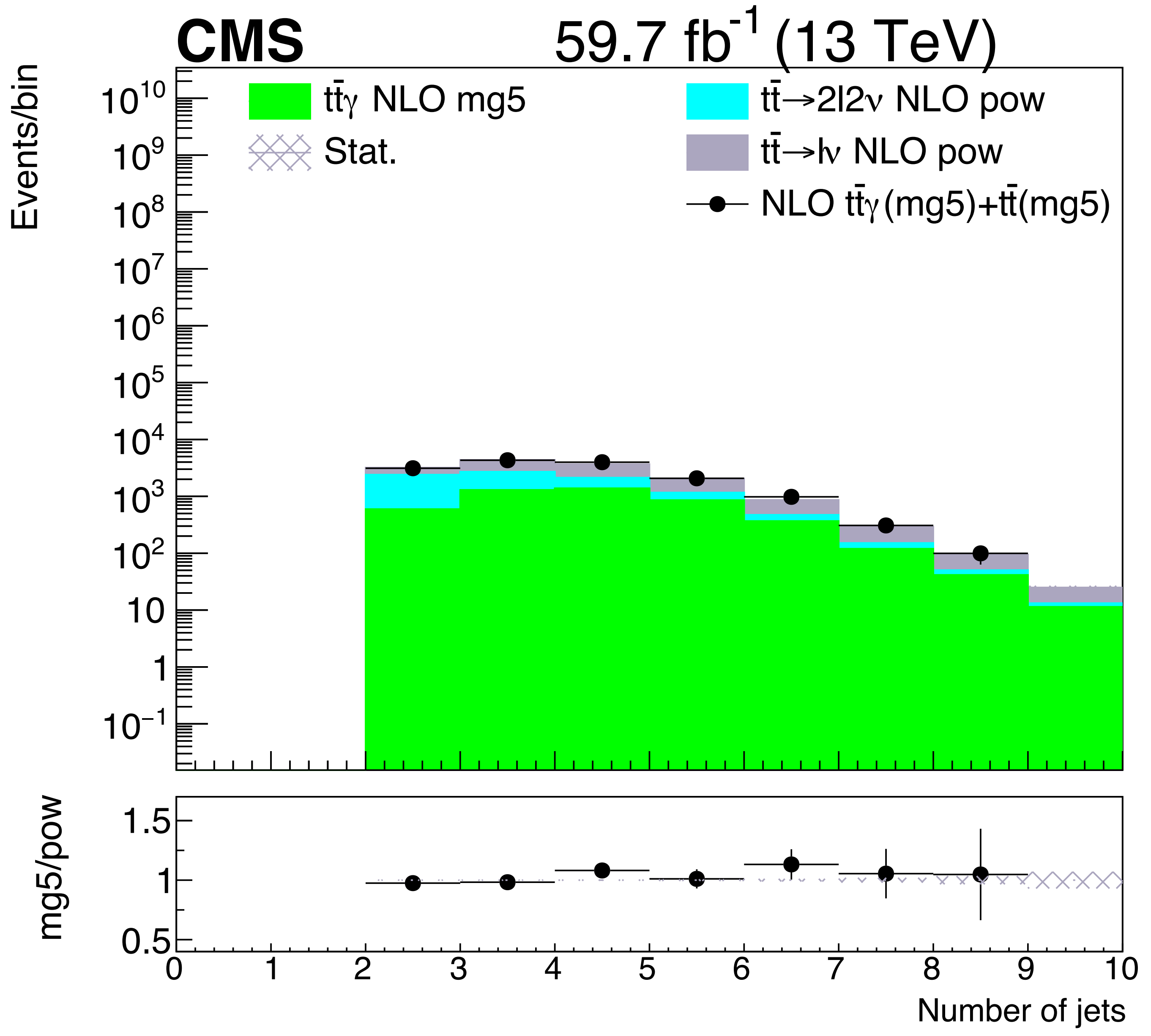
- The analysis TOP-18-010 measured the $t\bar{t}\gamma$ LO production k-factor **1.4852** in a fiducial region of N_γ =1, N_ℓ=1, $N_j\geq3$, $N_b\geq1$
- Normalised $t\bar{t}\gamma$ LO was compared with the $t\bar{t}\gamma$ NLO (MG5) + $t\bar{t}$ NLO (Powheg)
- The photon p_T distribution in the gen-level shows an agreement within ~20%

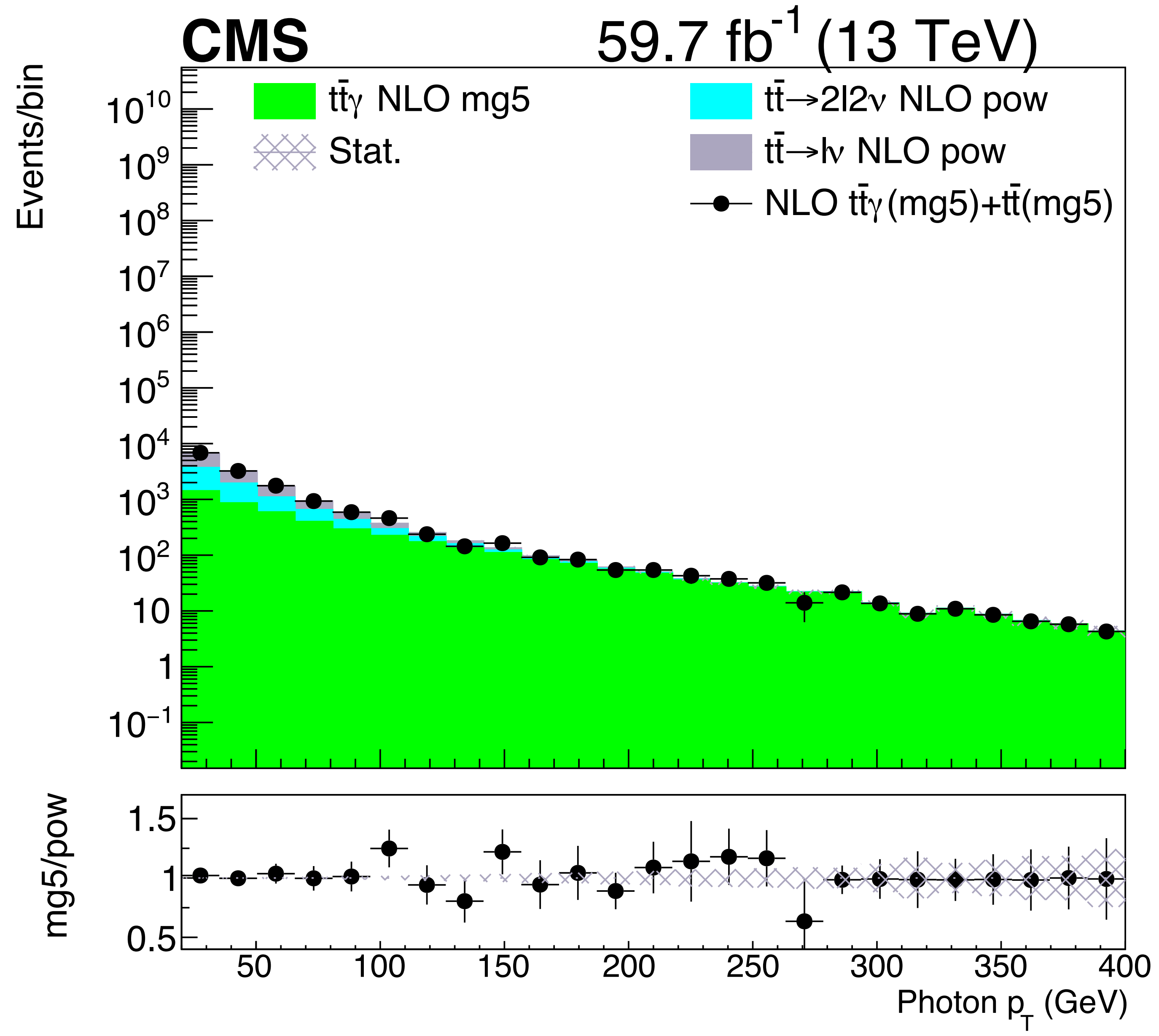
Cut	gen-Photon	gen-Electron (dressed)	gen-Muon (dressed)	gen-Jet	gen-b-Jet
p_T [GeV]	$\geq$20	$\geq$35	$\geq$30	$\geq$30	$\geq$30
$\lvert\eta\rvert$	$\leq$1.4442	$\leq$2.4	$\leq$2.4	$\leq$2.4	$\leq$2.4
|pdgID|	20	11	13		
status	1	1	1		
other	no meson mother isolated(*) min $\Delta R(\gamma, \ell)$>0.4	no meson mother	no meson mother	min ΔR(jets, ℓ)>0.4 min ΔR(jets, γ)>0.1	|partonFlavour| = 5 min ΔR(b-jets, ℓ)>0.4 min ΔR(b-jets, γ)>0.1

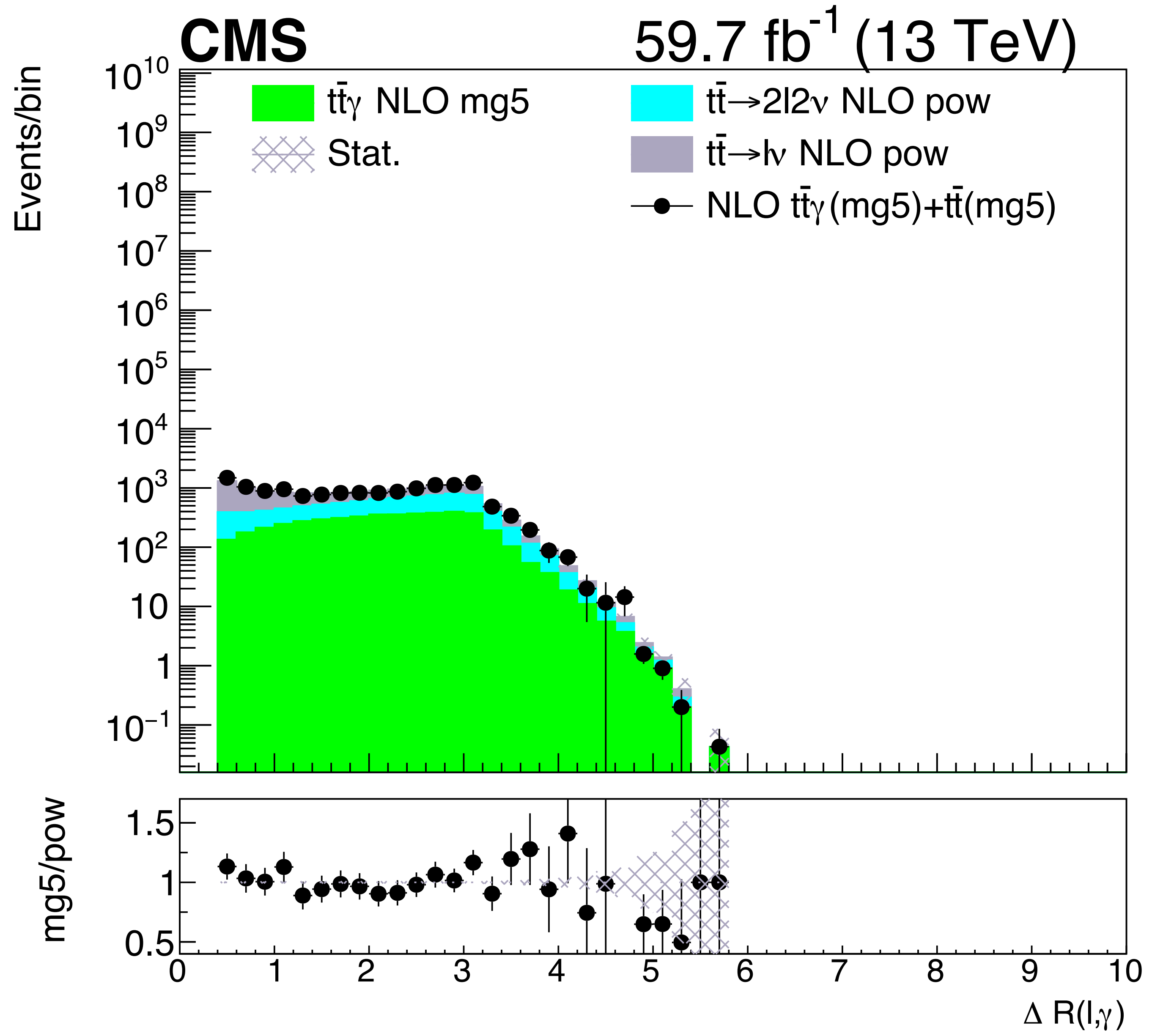
(*)Photon Isolation requirement: No status 1 gen-particles (except neutrinos) with p_T >5 GeV in a ΔR cone of 0.1

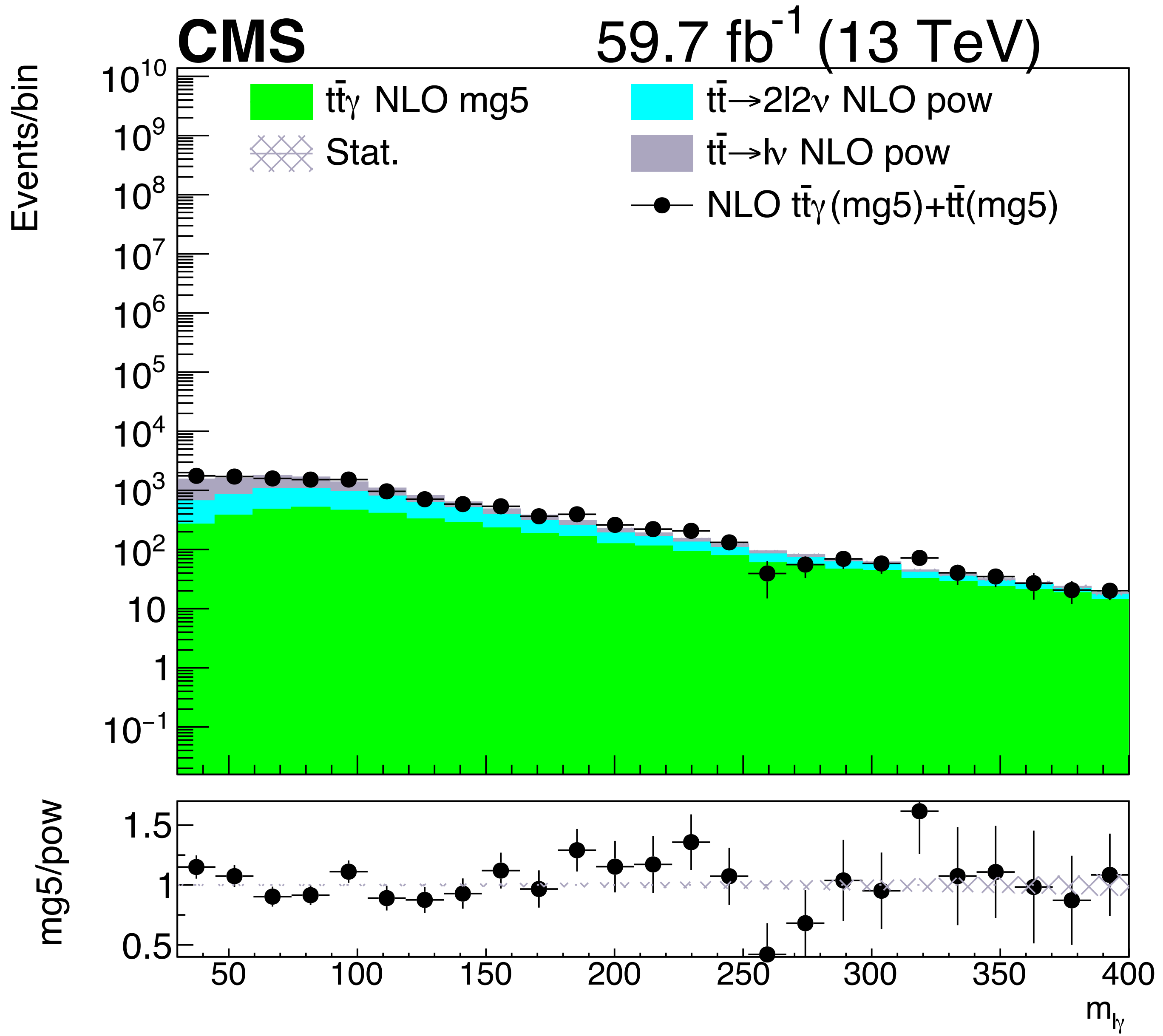
back

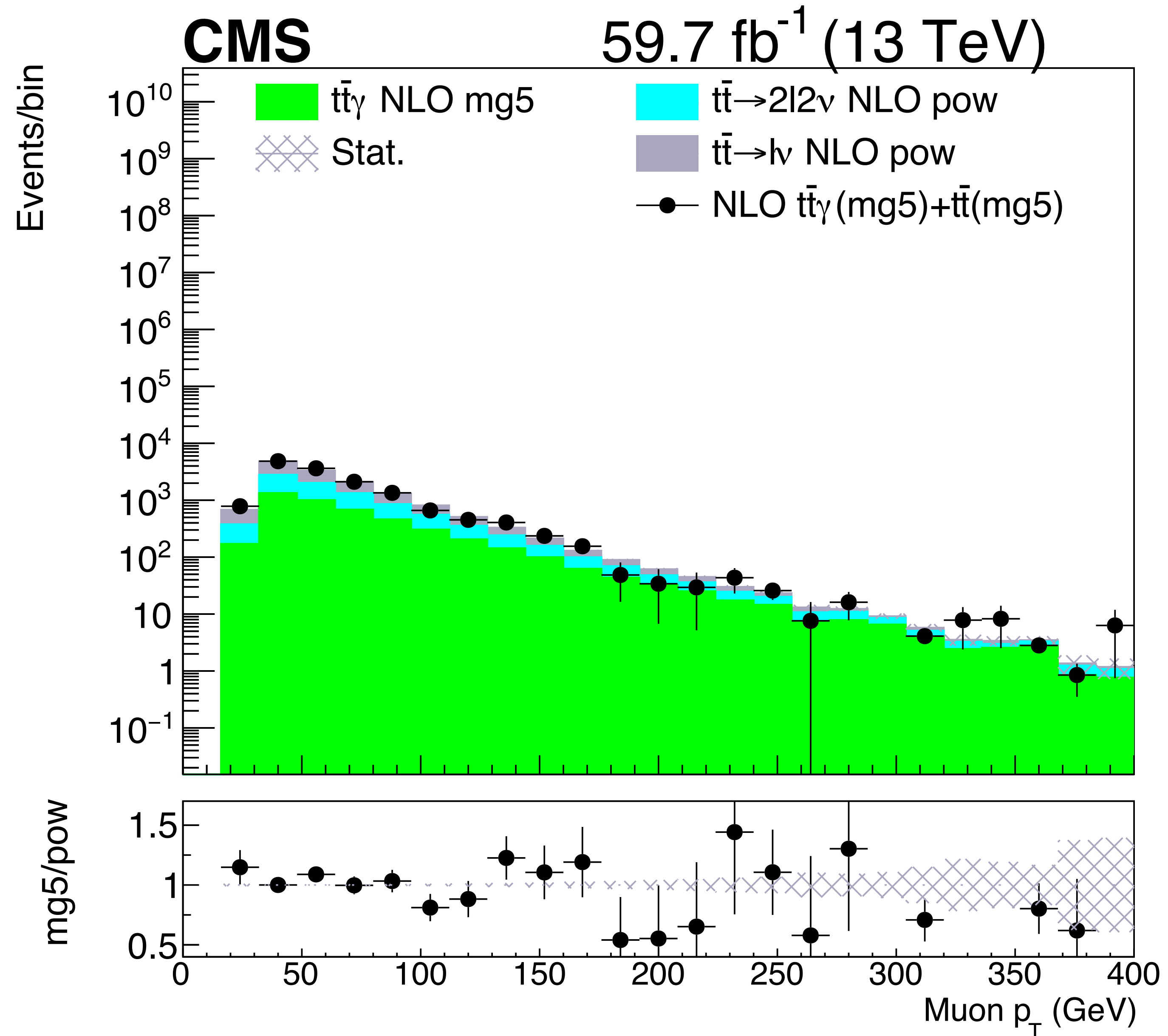
Removal in different generators

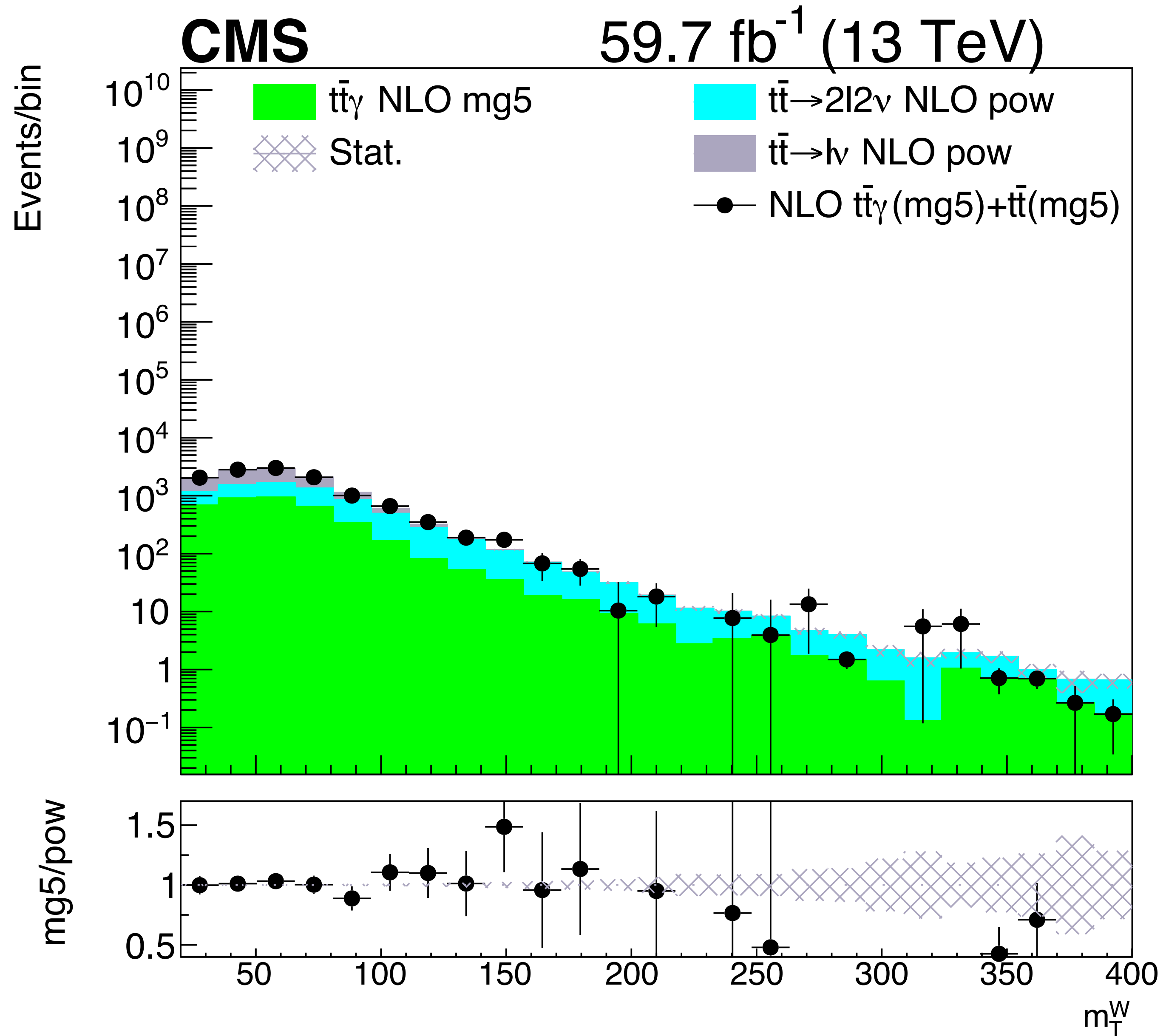






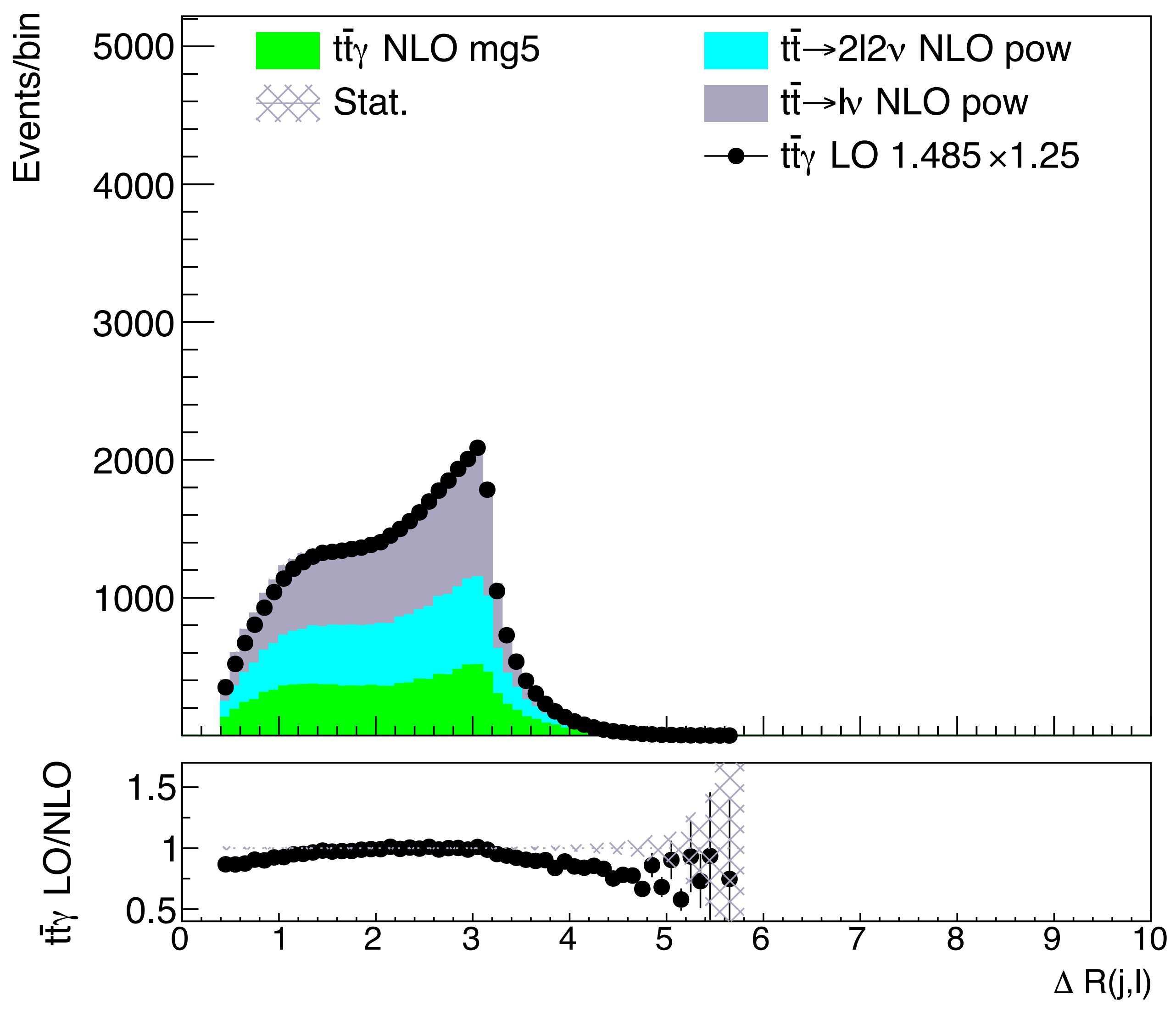


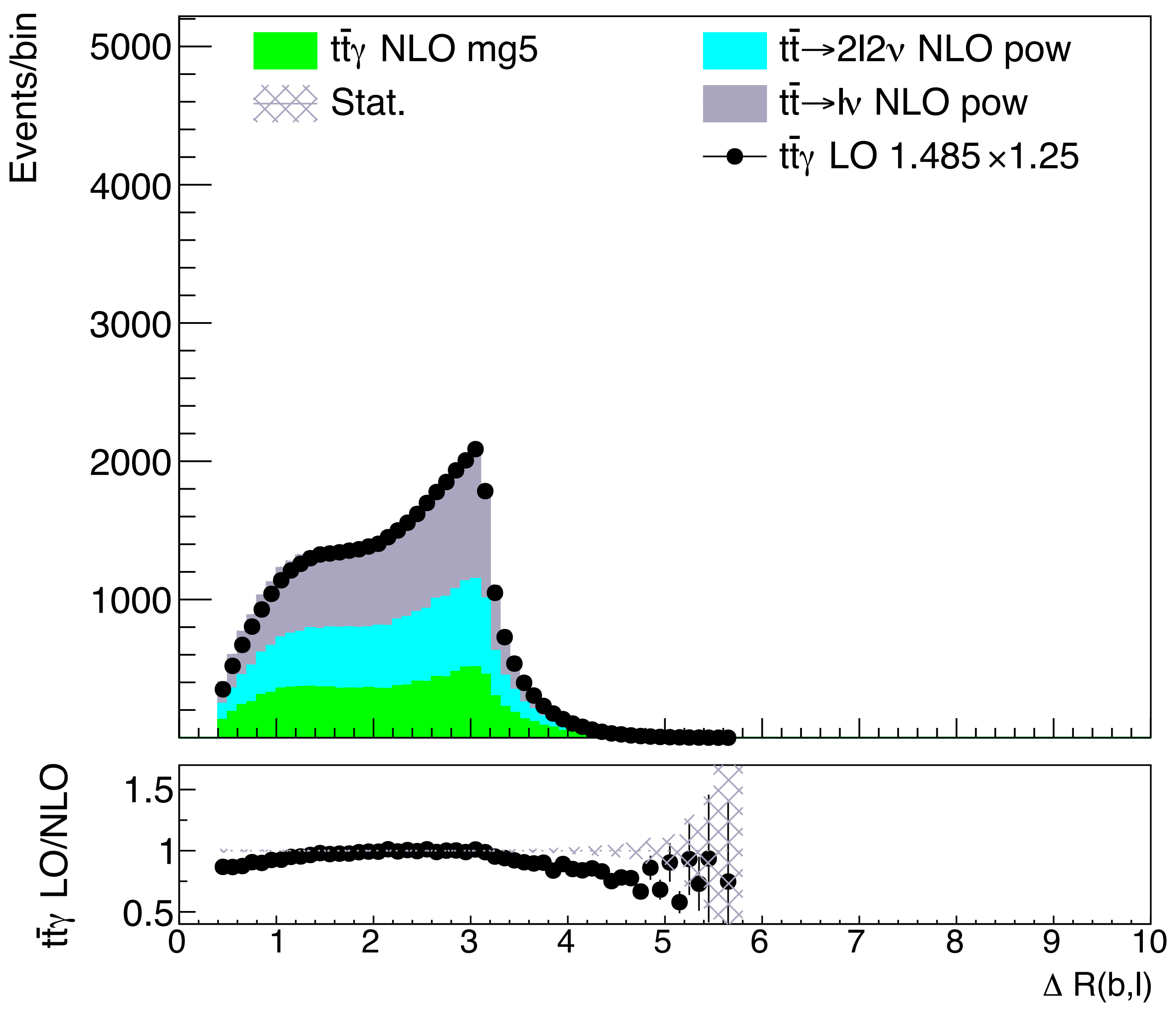


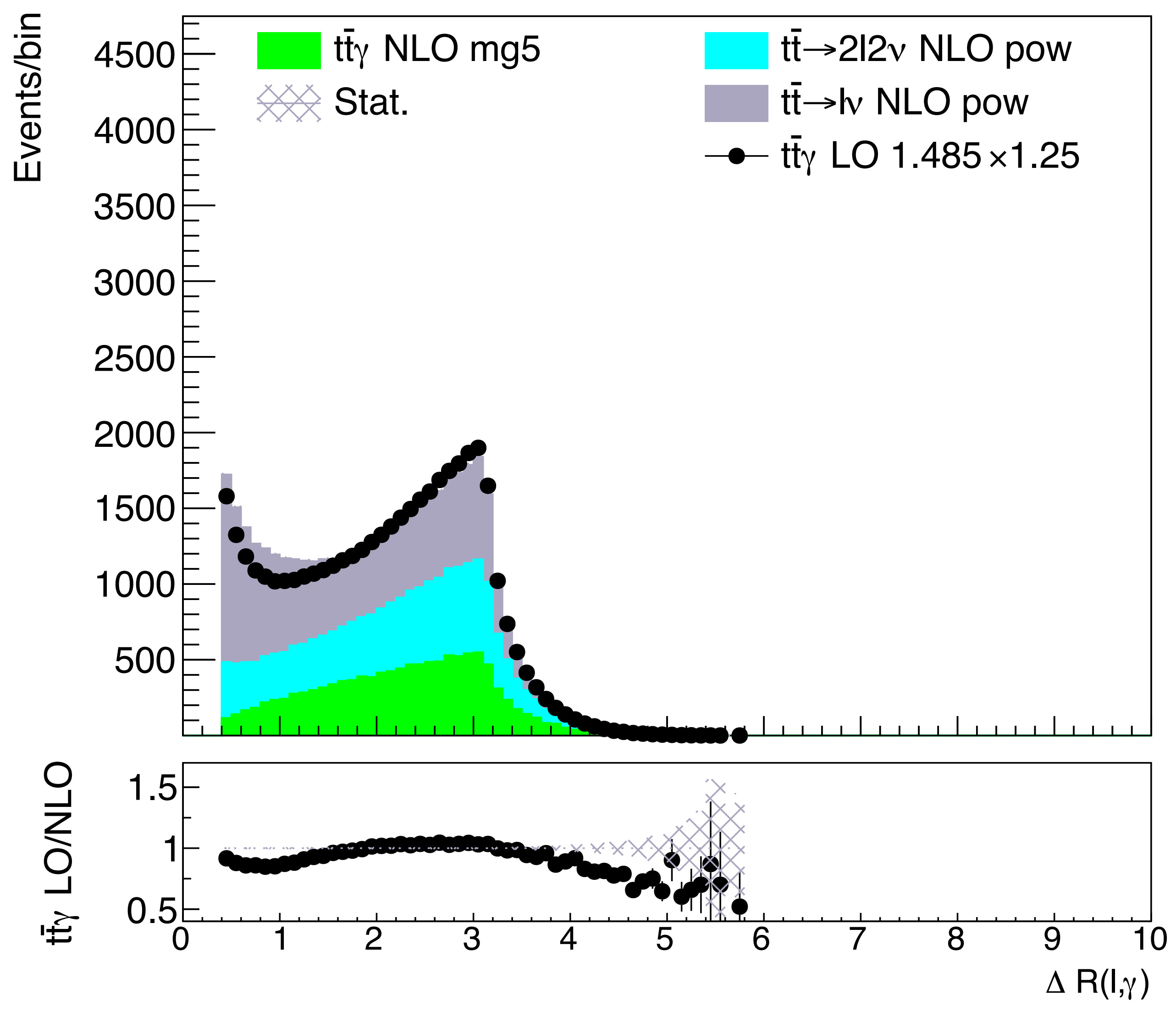


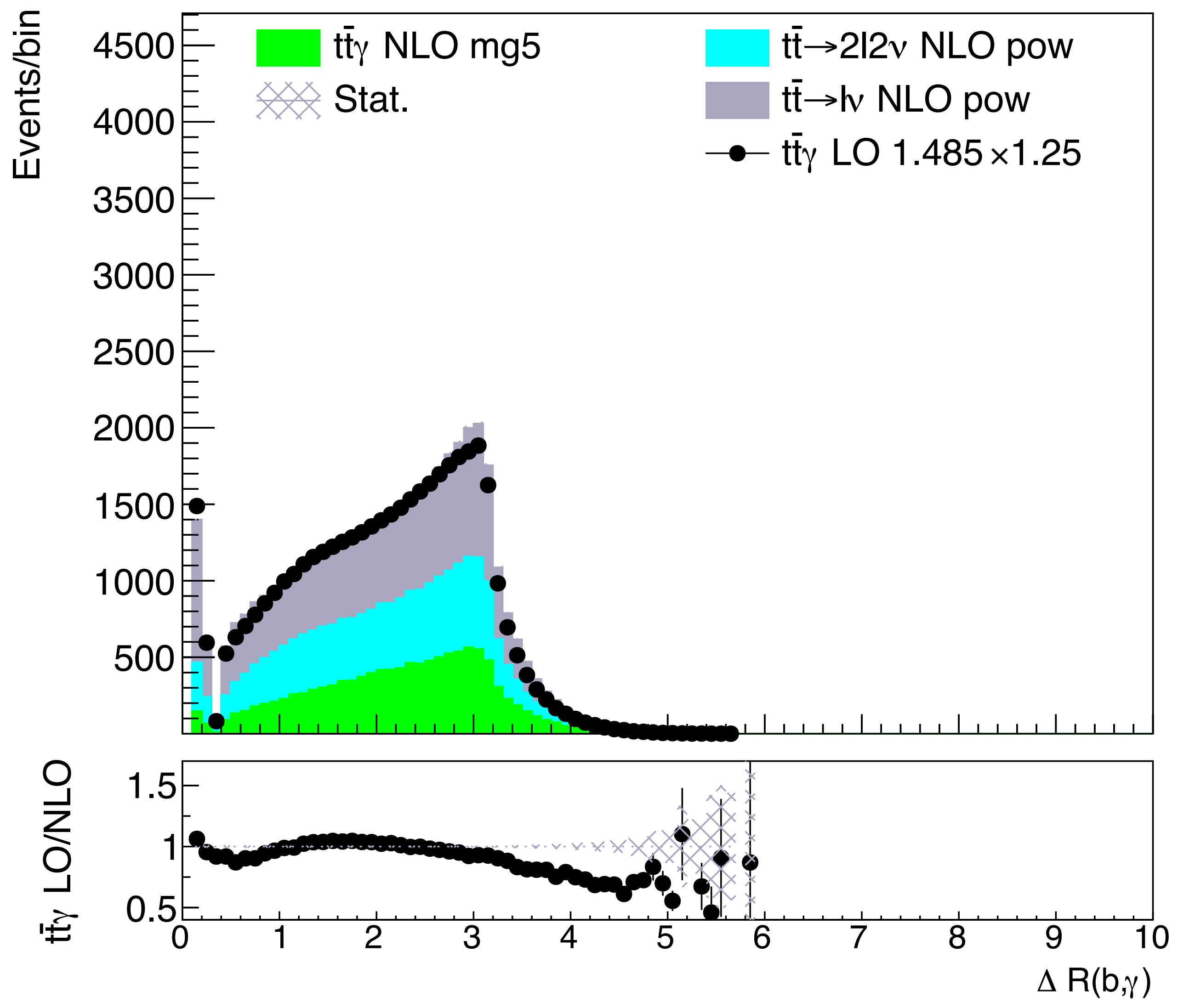
back

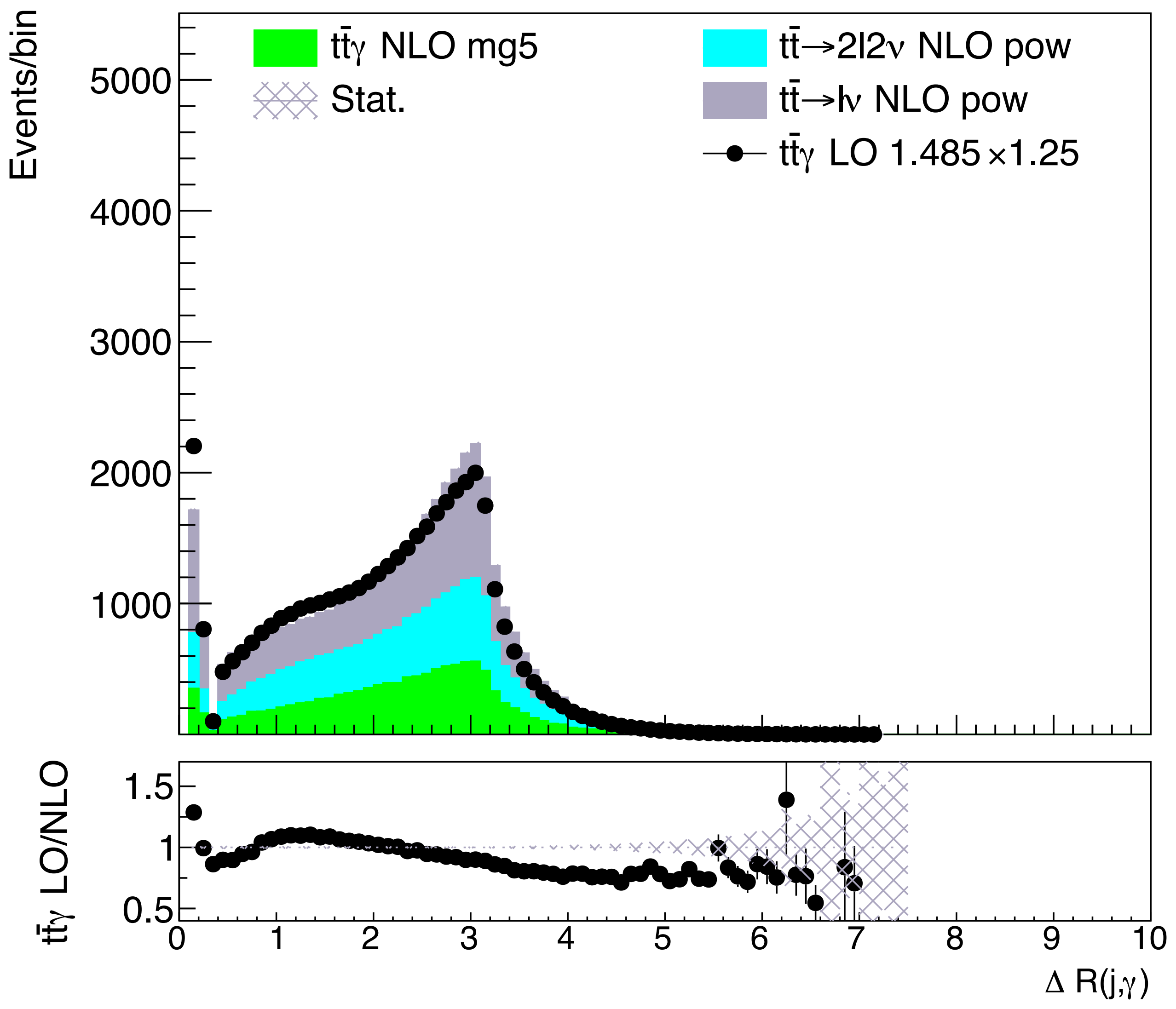
$t\bar{t}\gamma$ gen-level distributions

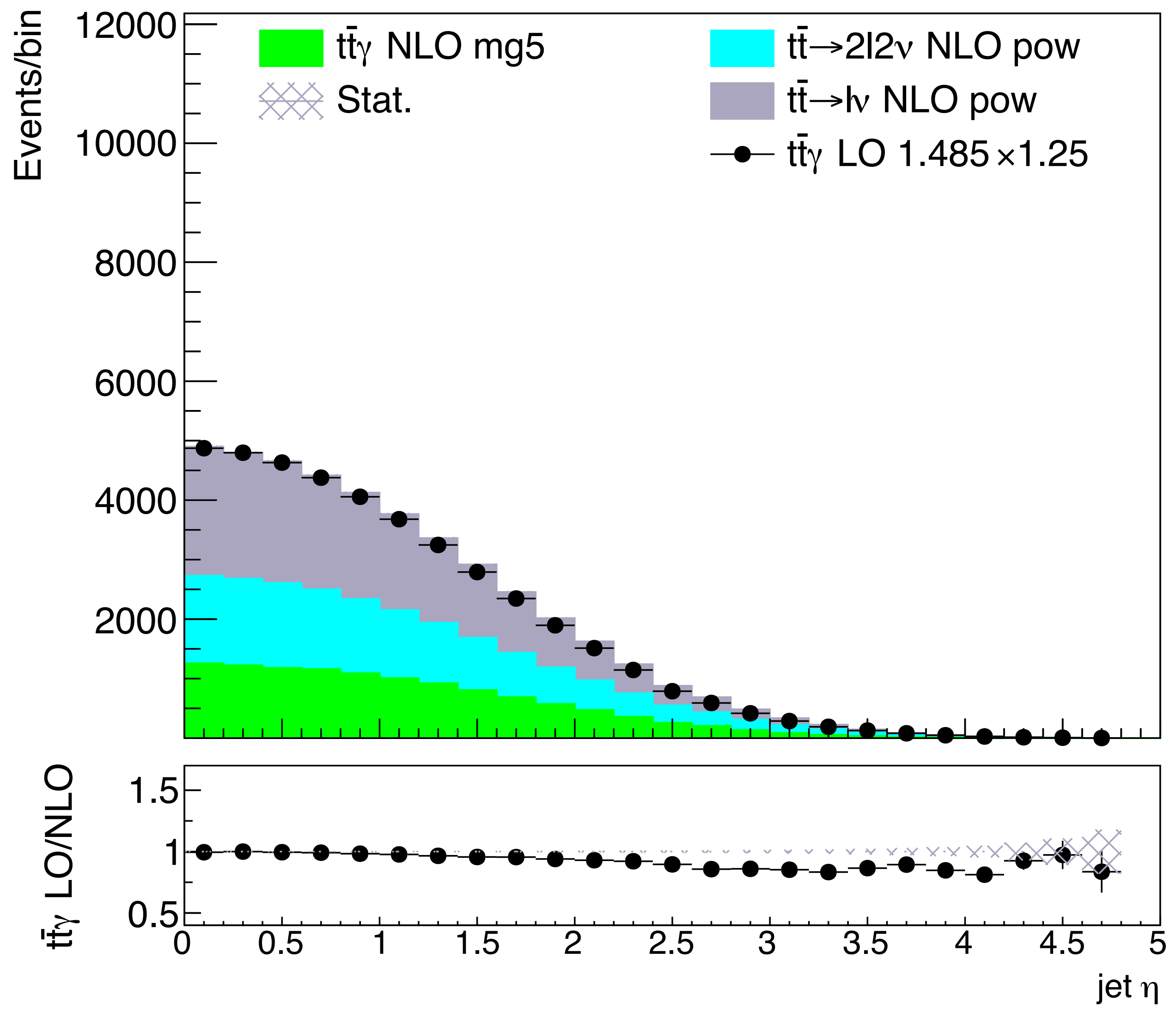






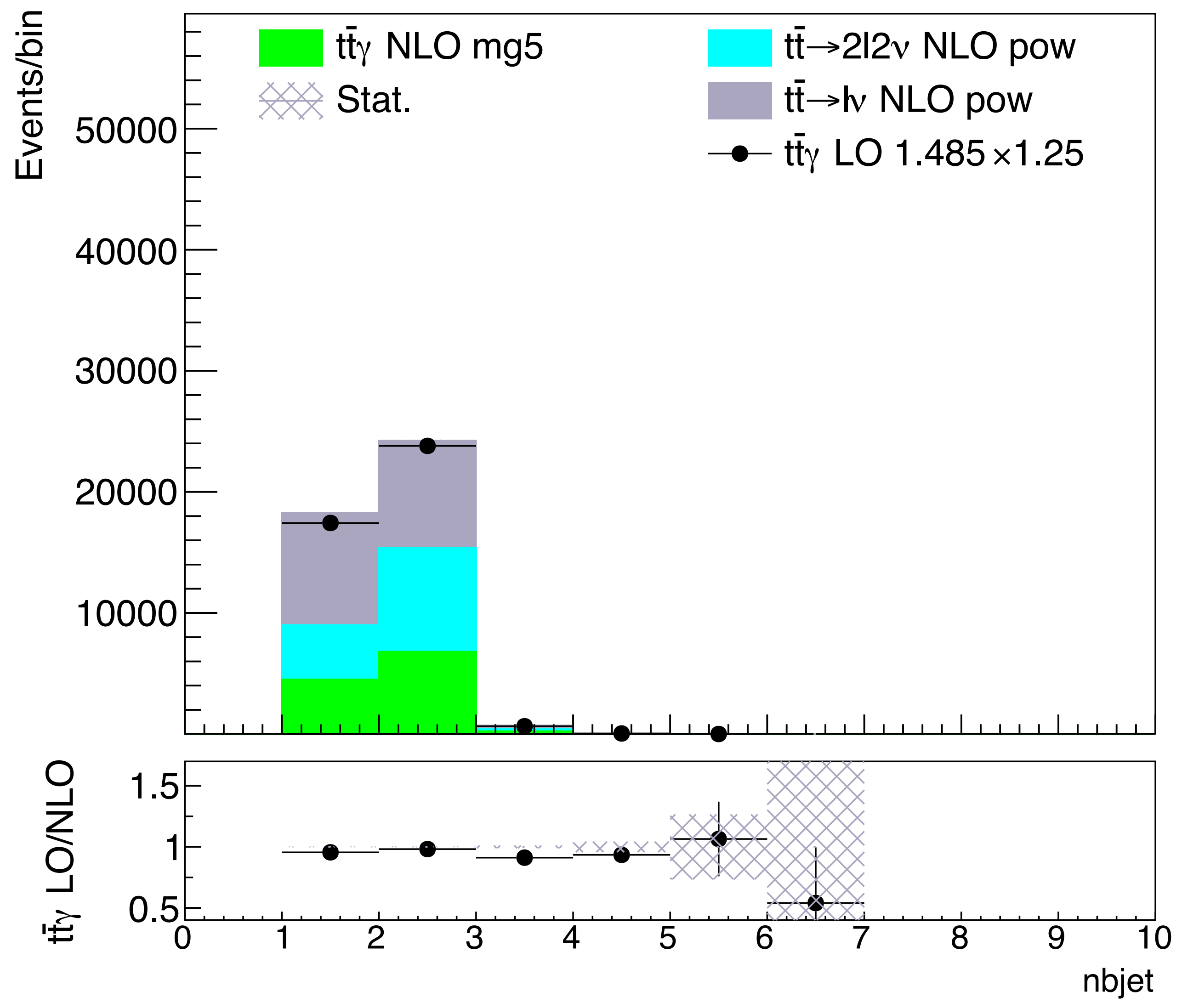


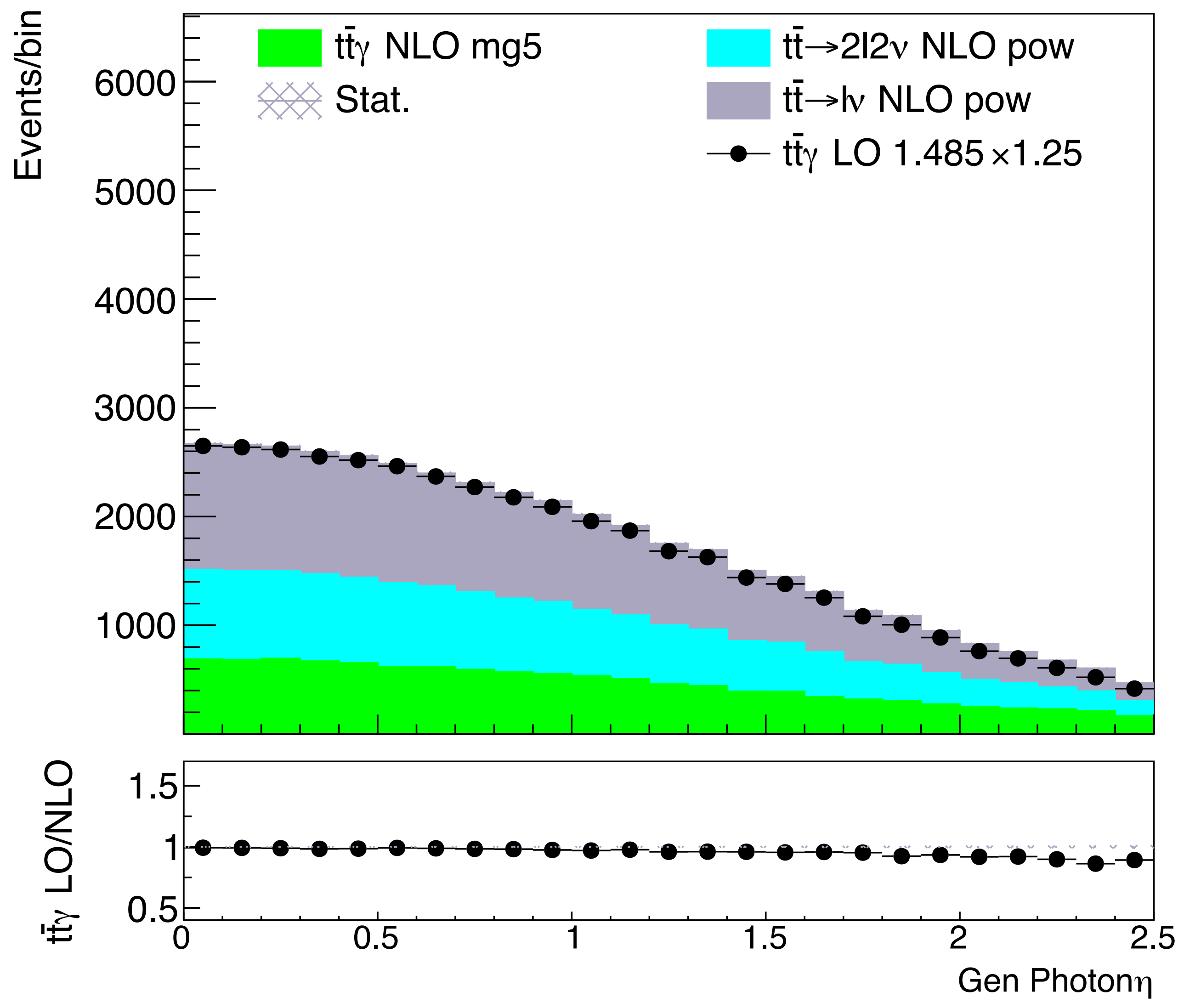


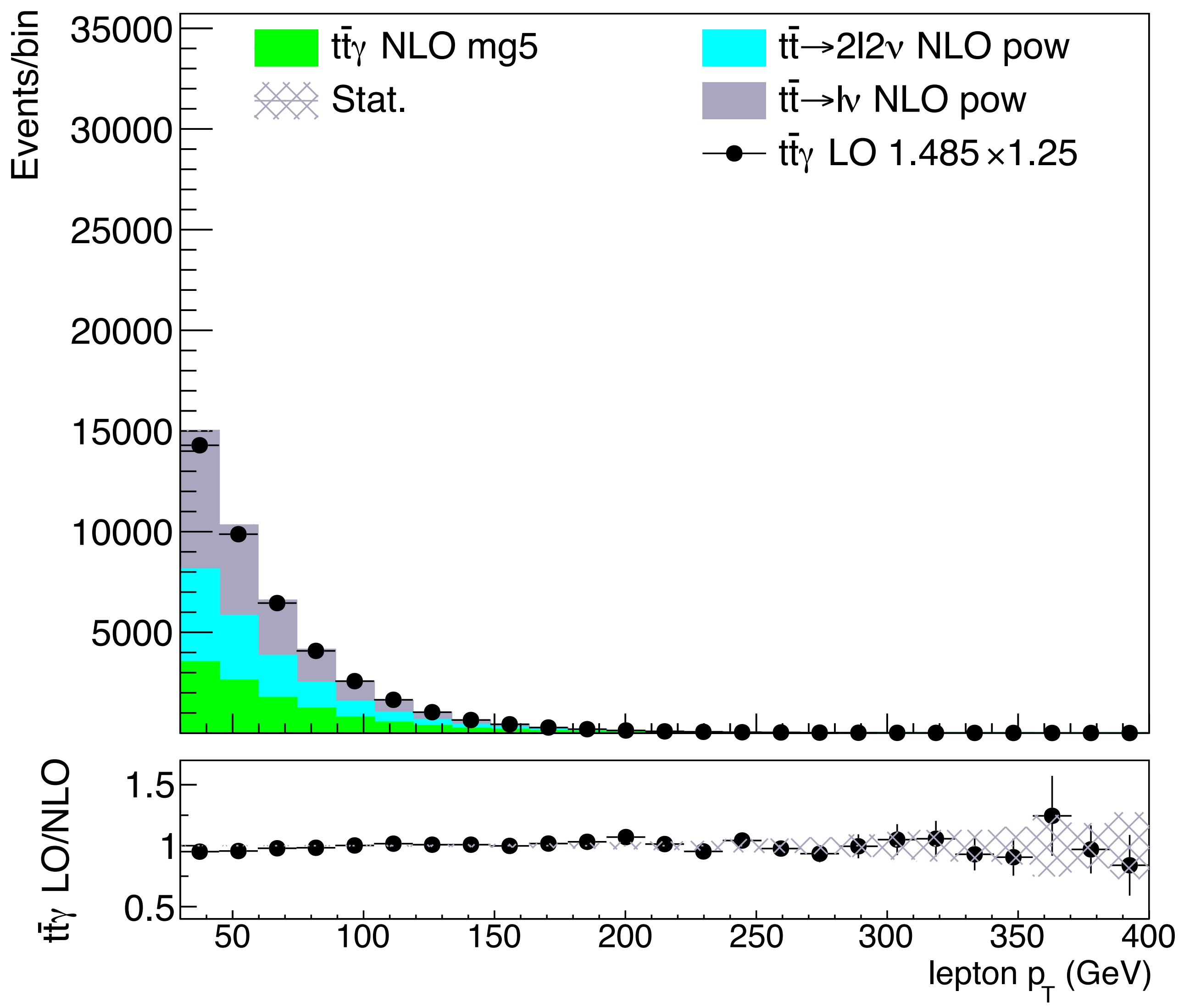


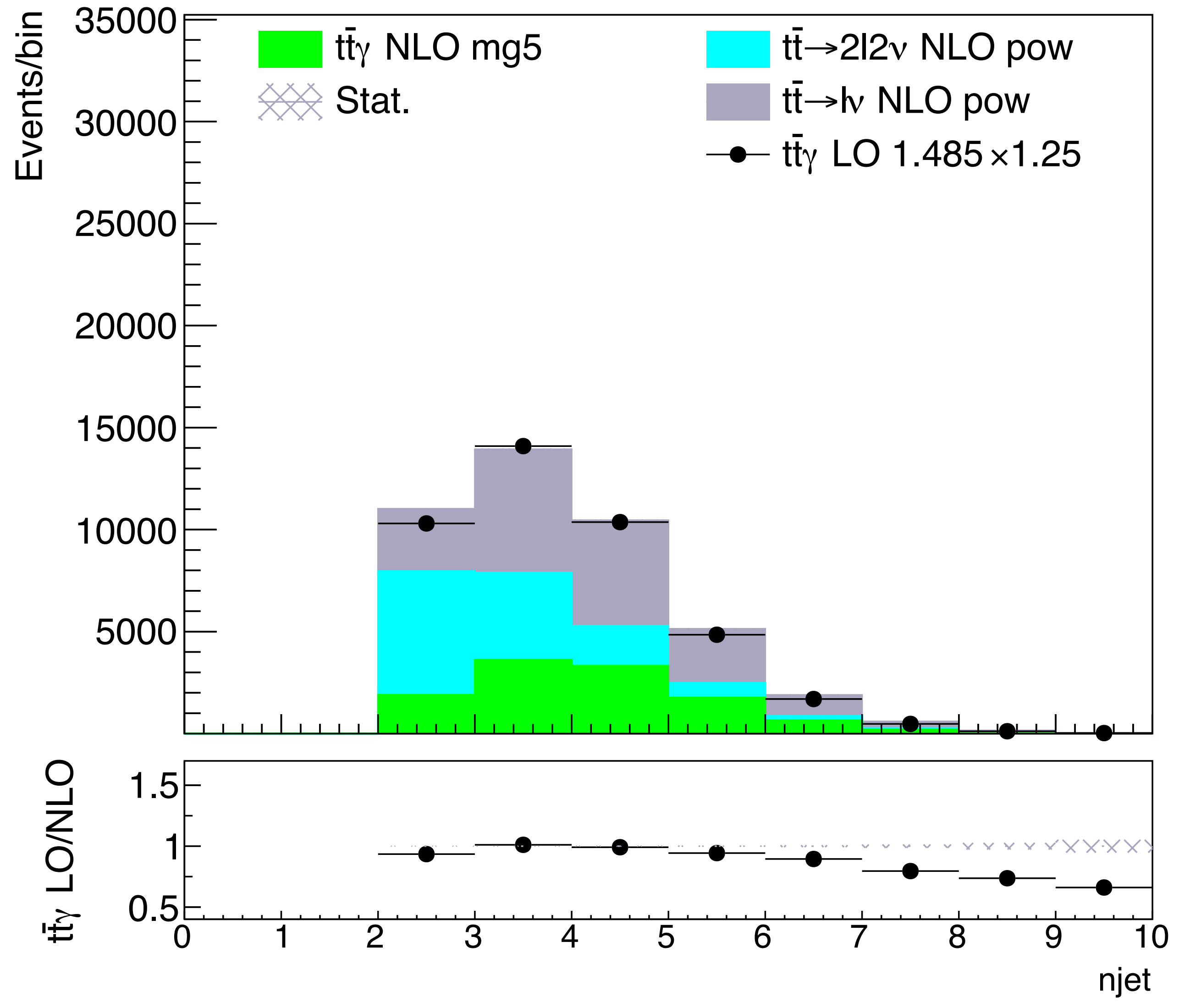
back

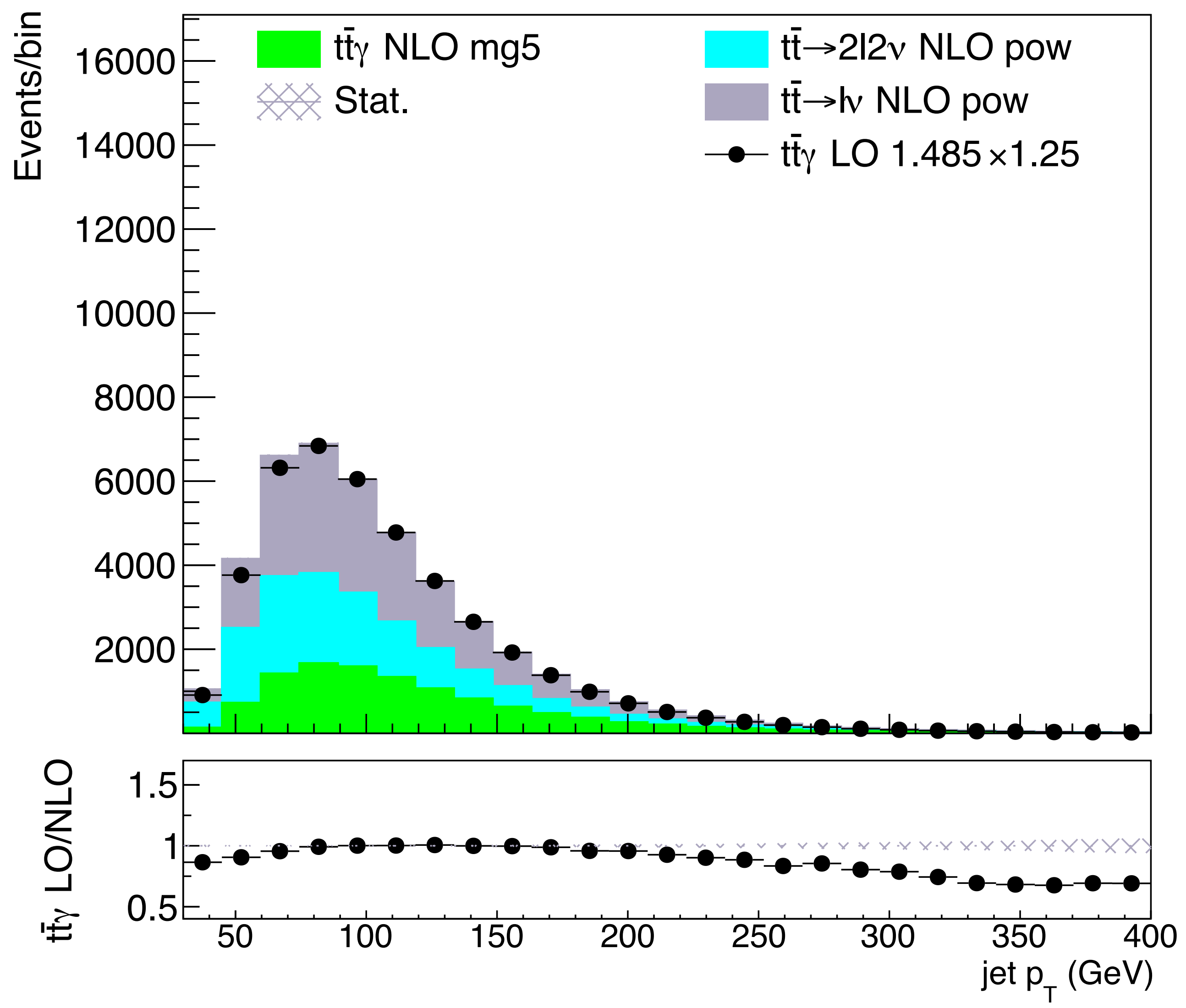
$t\bar{t}\gamma$ gen-level distributions

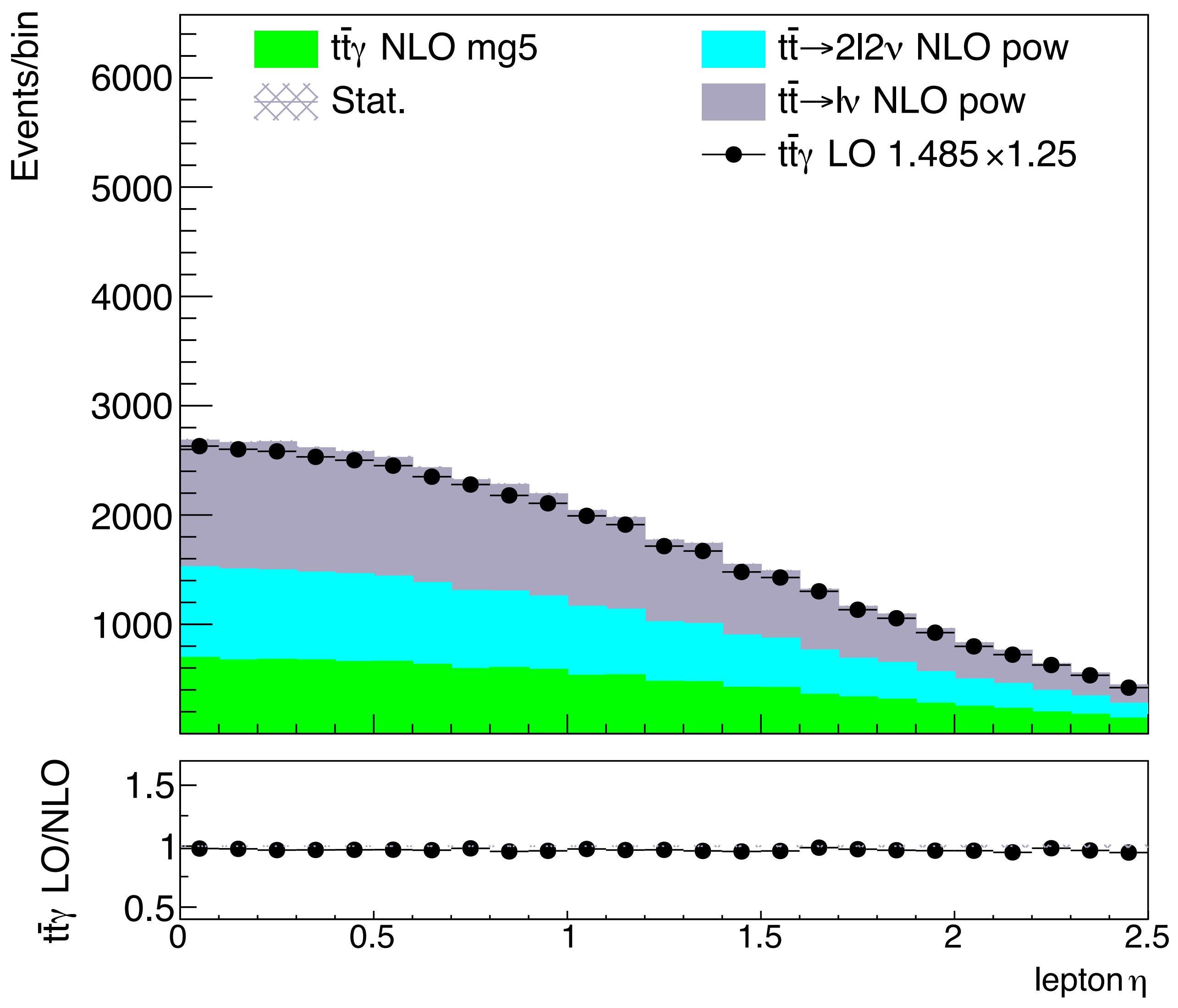










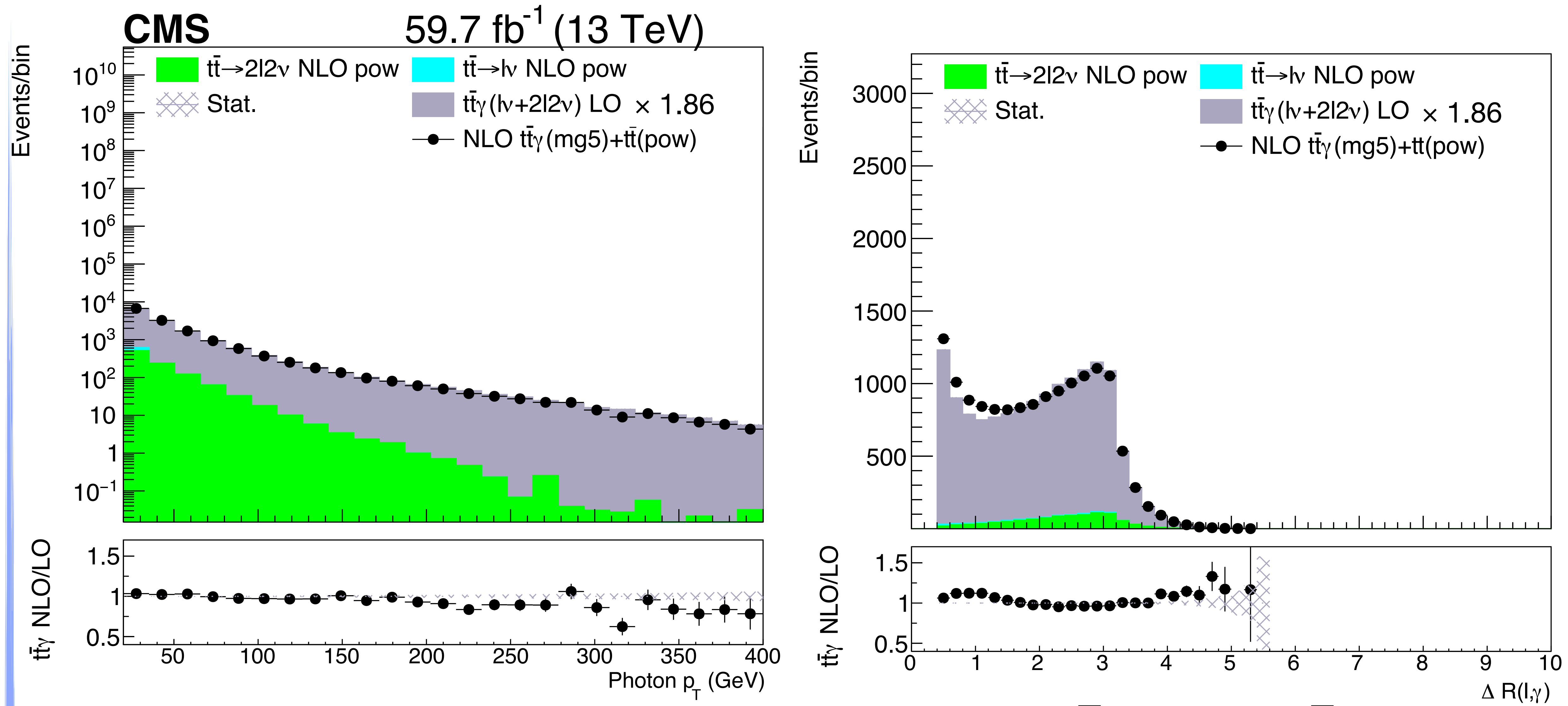


back

(Signal) simulation — $t\bar{t}\gamma$

- **Reconstruction level** with SR requirements
- $N_{\ell}=1$, $N_{\gamma} \geq 1$, $N_j \geq 2$, $N_b \geq 1$

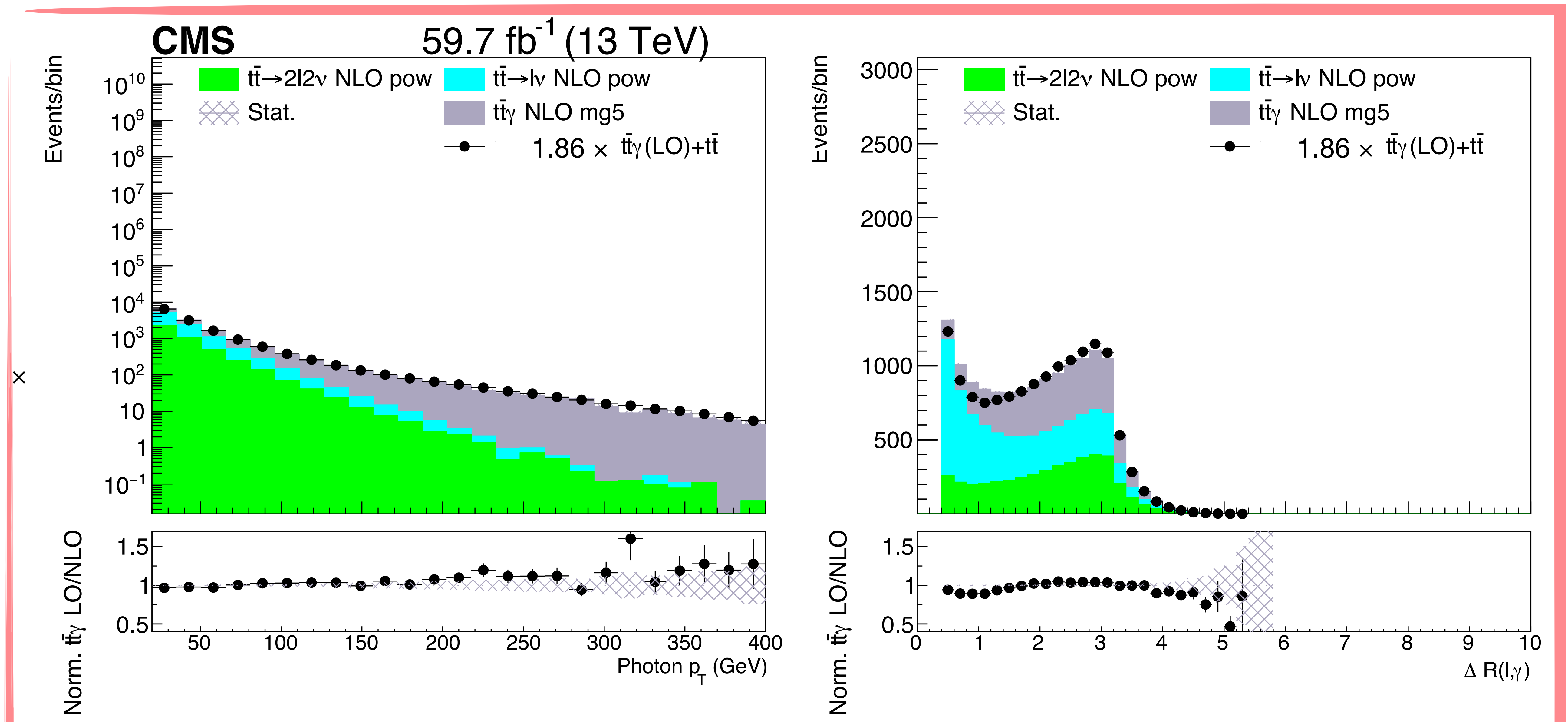
The k-factor ≈1.86 works well here but is not suitable in region of $N_b=0$ (plots backup)



Stacked histograms: Normalised $t\bar{t}\gamma$ LO + $t\bar{t}$ after overlap removal

Black points: $t\bar{t}\gamma$ NLO+ $t\bar{t}$ after overlap removal

From stacked histograms, most contribution comes from $t\bar{t}\gamma$ LO, $t\bar{t}$ contribution is small after removal



Stacked histograms: $t\bar{t}\gamma$ NLO + $t\bar{t}$ after overlap removal

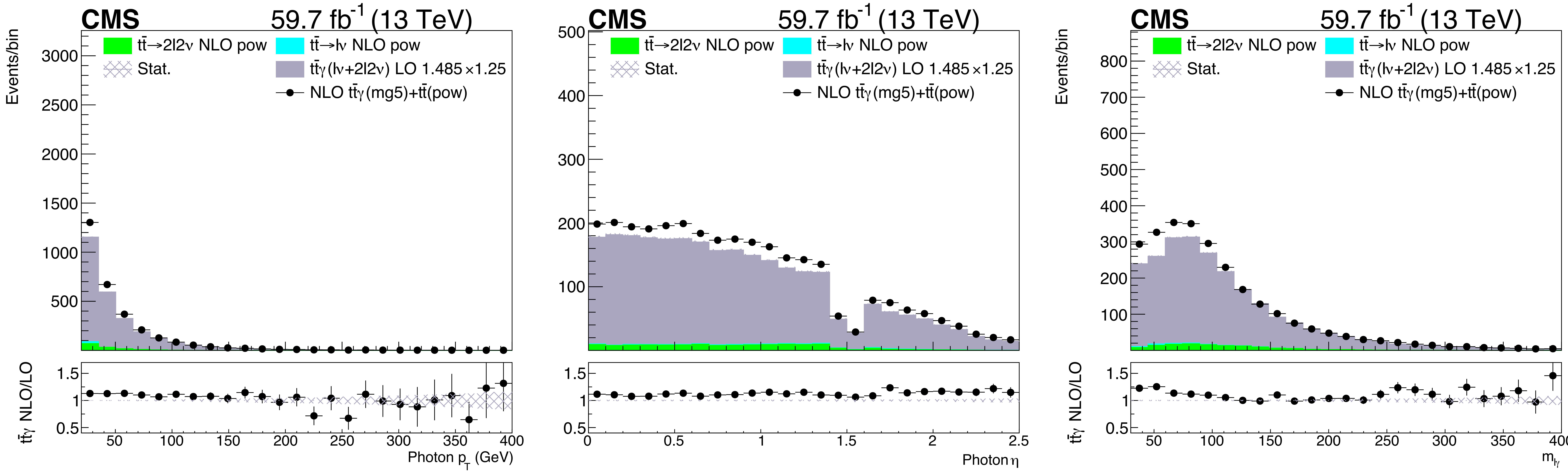
Black points: Normalised $t\bar{t}\gamma$ LO + $t\bar{t}$ after overlap removal

From stacked histograms, contribution from $t\bar{t}\gamma$ NLO and $t\bar{t}$ contribution is comparable

back

$t\bar{t}\gamma$ reco-level distributions

When N_b=0



Since the $t\bar{t}$ contribution is small/suppressed when require N_b=0, this deviation might be fine

back

Background simulation — Zγ+jets

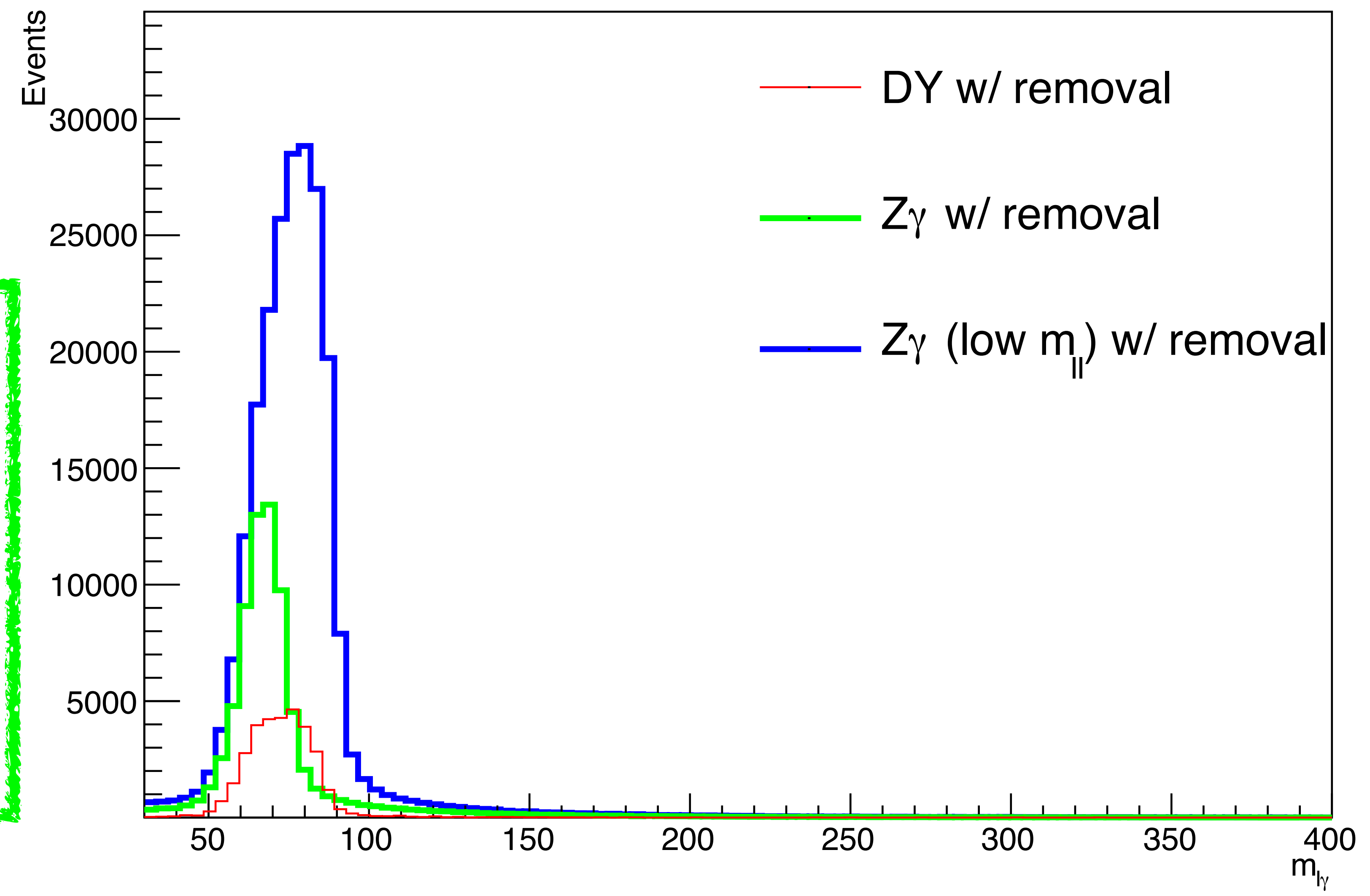
- DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8: **6077.22 pb**
- ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8: **51.48 pb**
- ZGToLLG_01J_5f_lowMLL_lowGPt_TuneCP5_13TeV-amcatnloFXFX-pythia8: **174.1 pb**

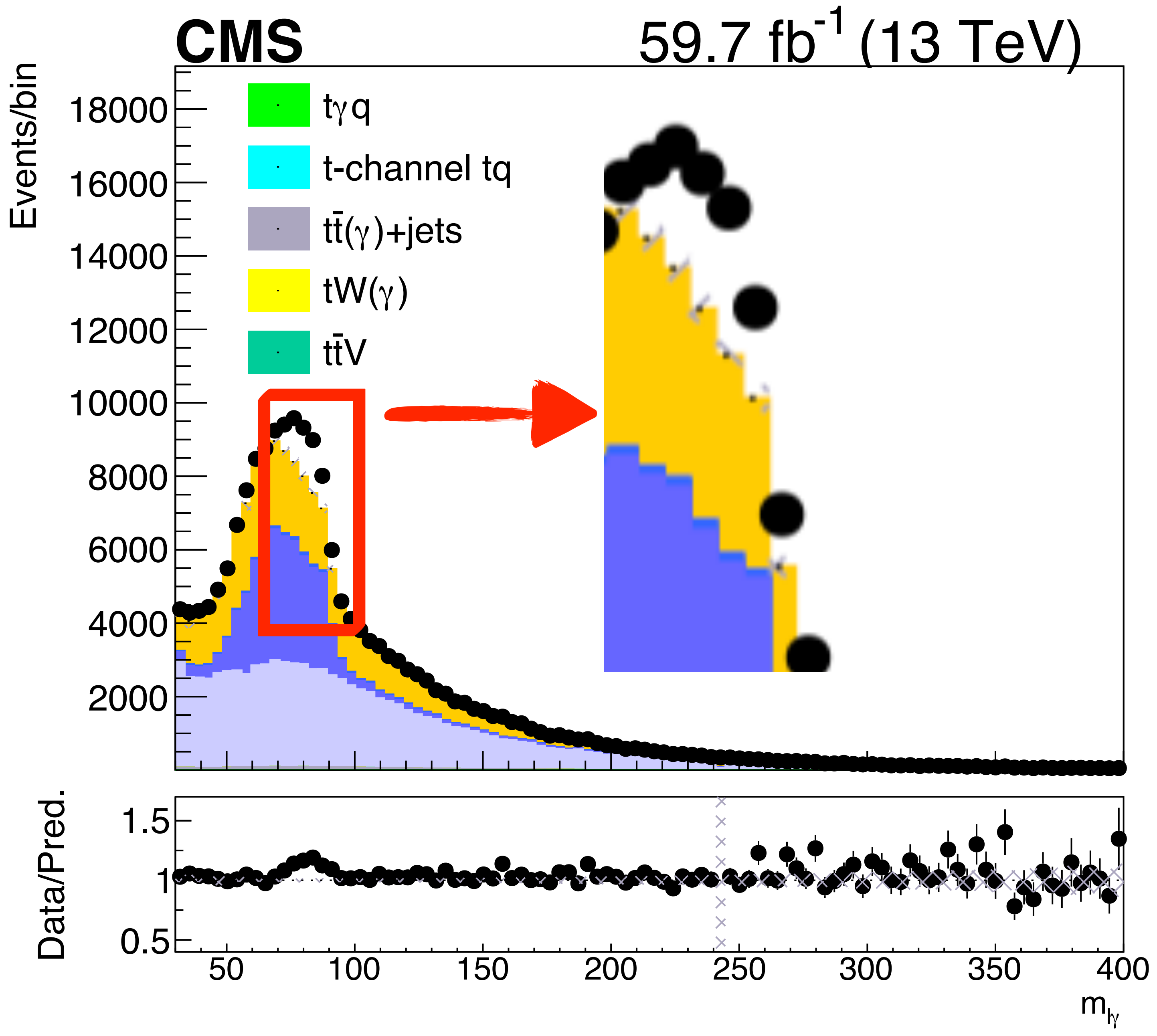
Zγ production process card:
generate p p > lep lep a [QCD] @0
add process p p > lep lep j a [QCD] @1

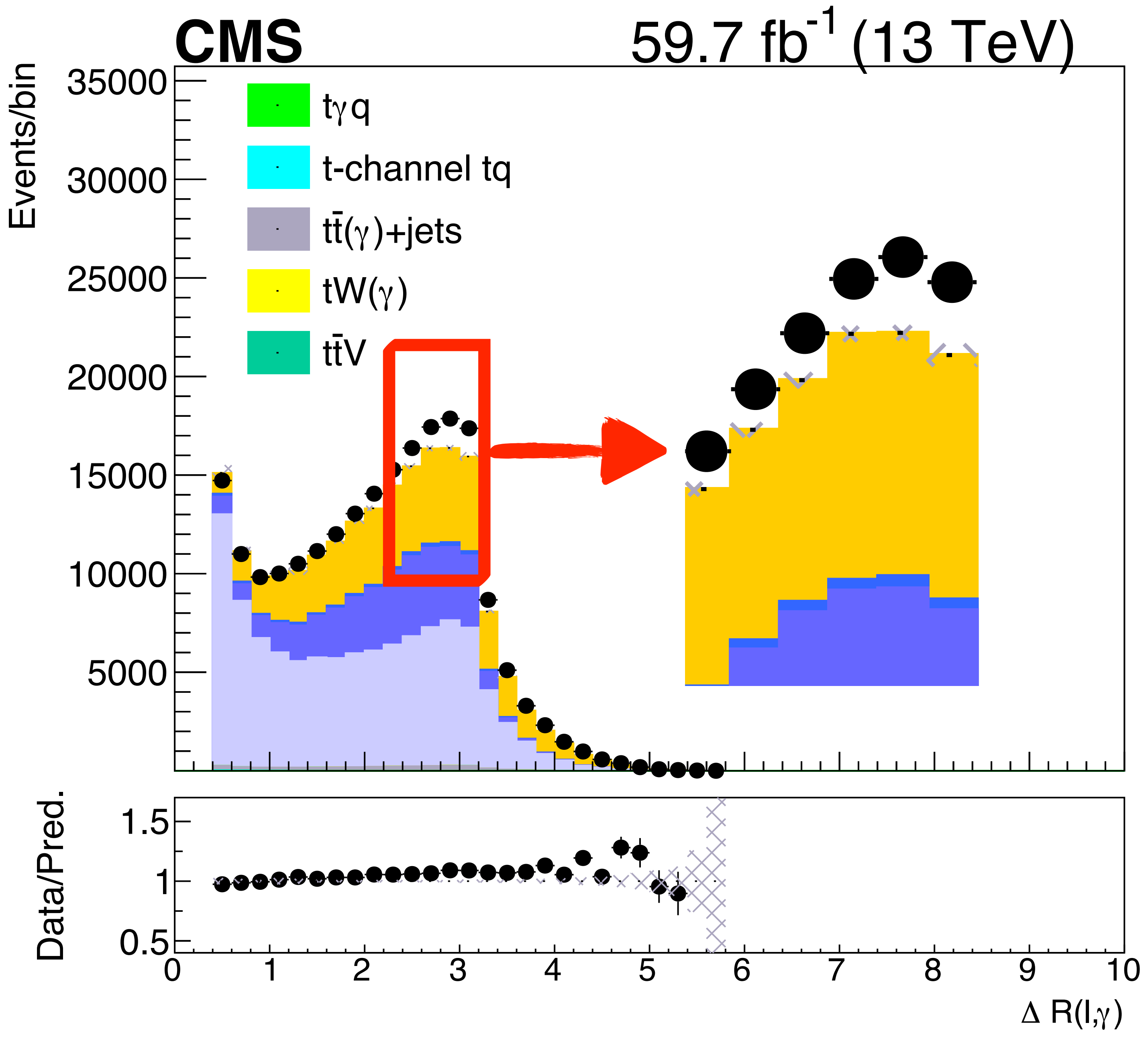
run card
ptl = 1
mll_sf = 10
ptgmin = 9
R0gamma = 0.05

run card
ptl = 15
mll_sf = 30
ptgmin = 15
R0gamma = 0.05

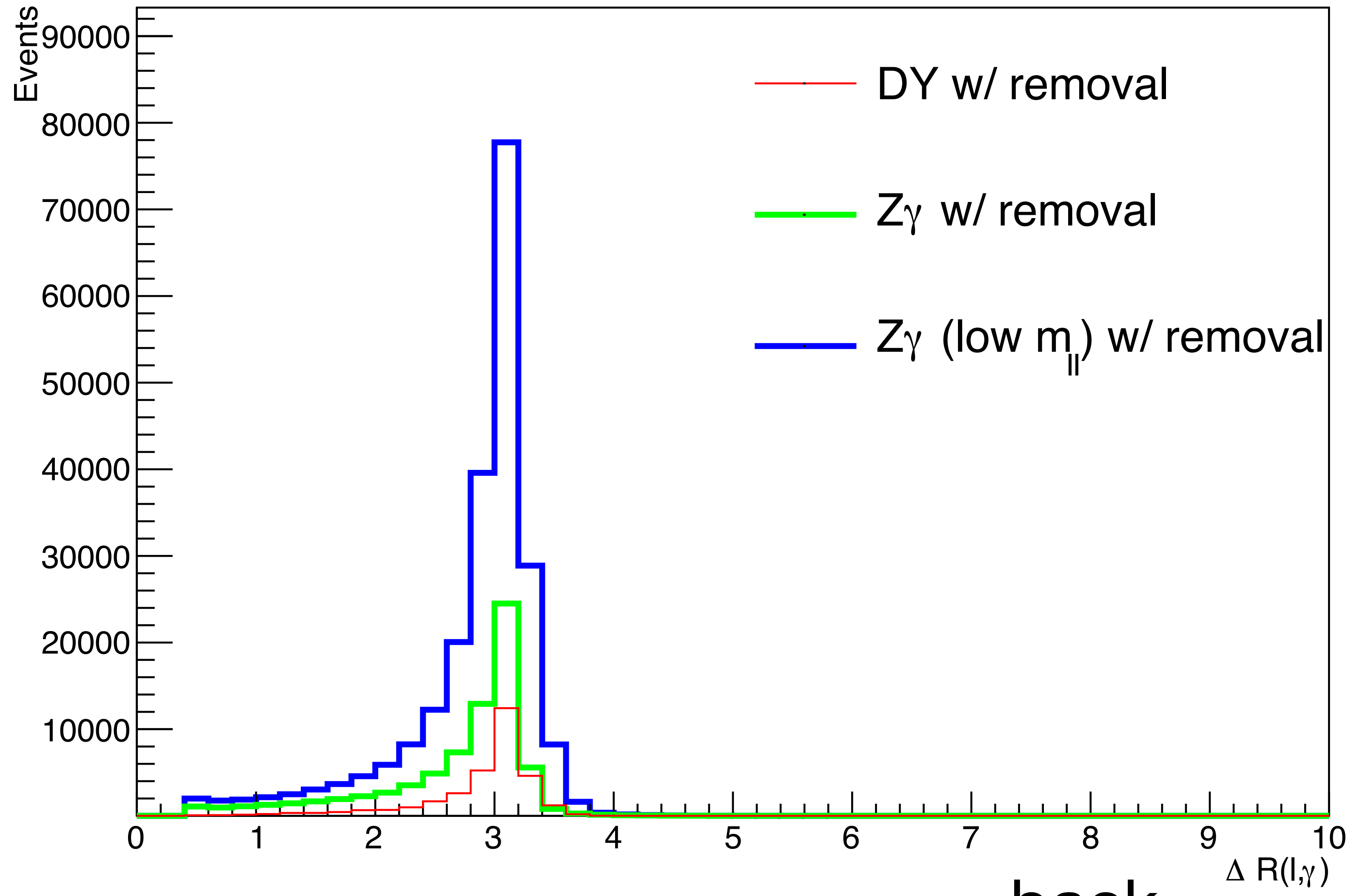
- $N_\ell=1$, $N_\gamma \geq 1$







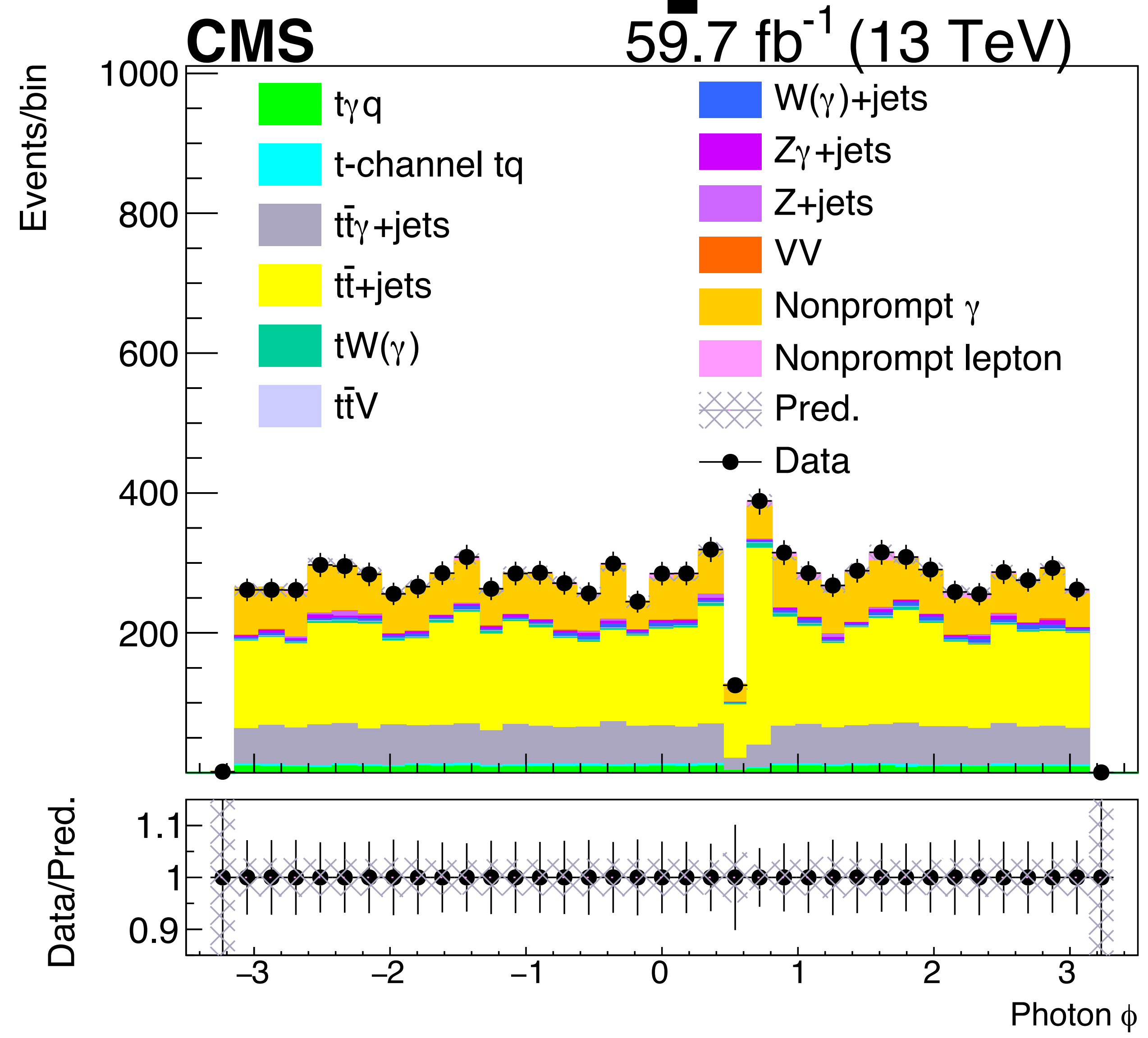
- $N_\ell=1$, $N_\gamma \geq 1$

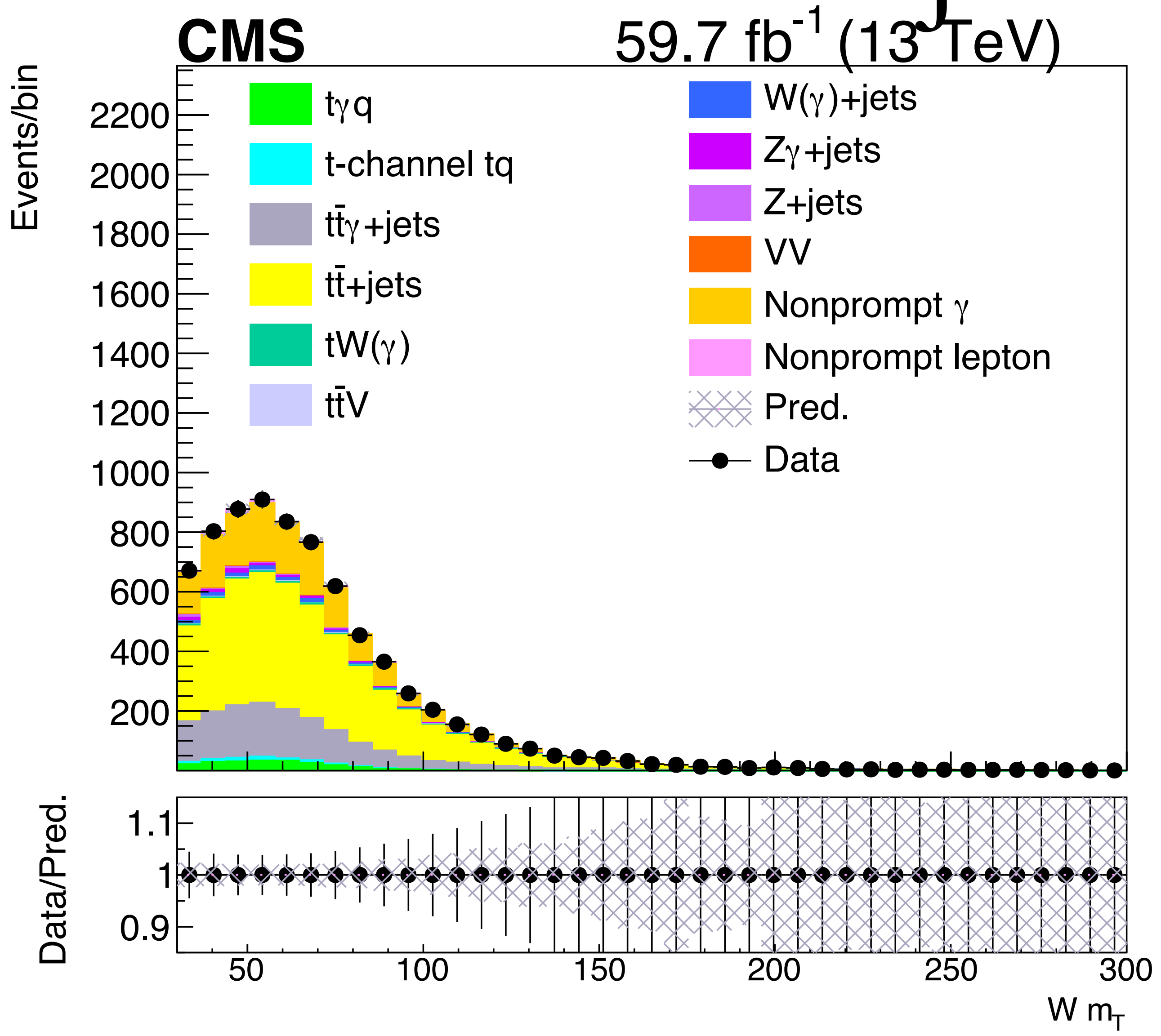


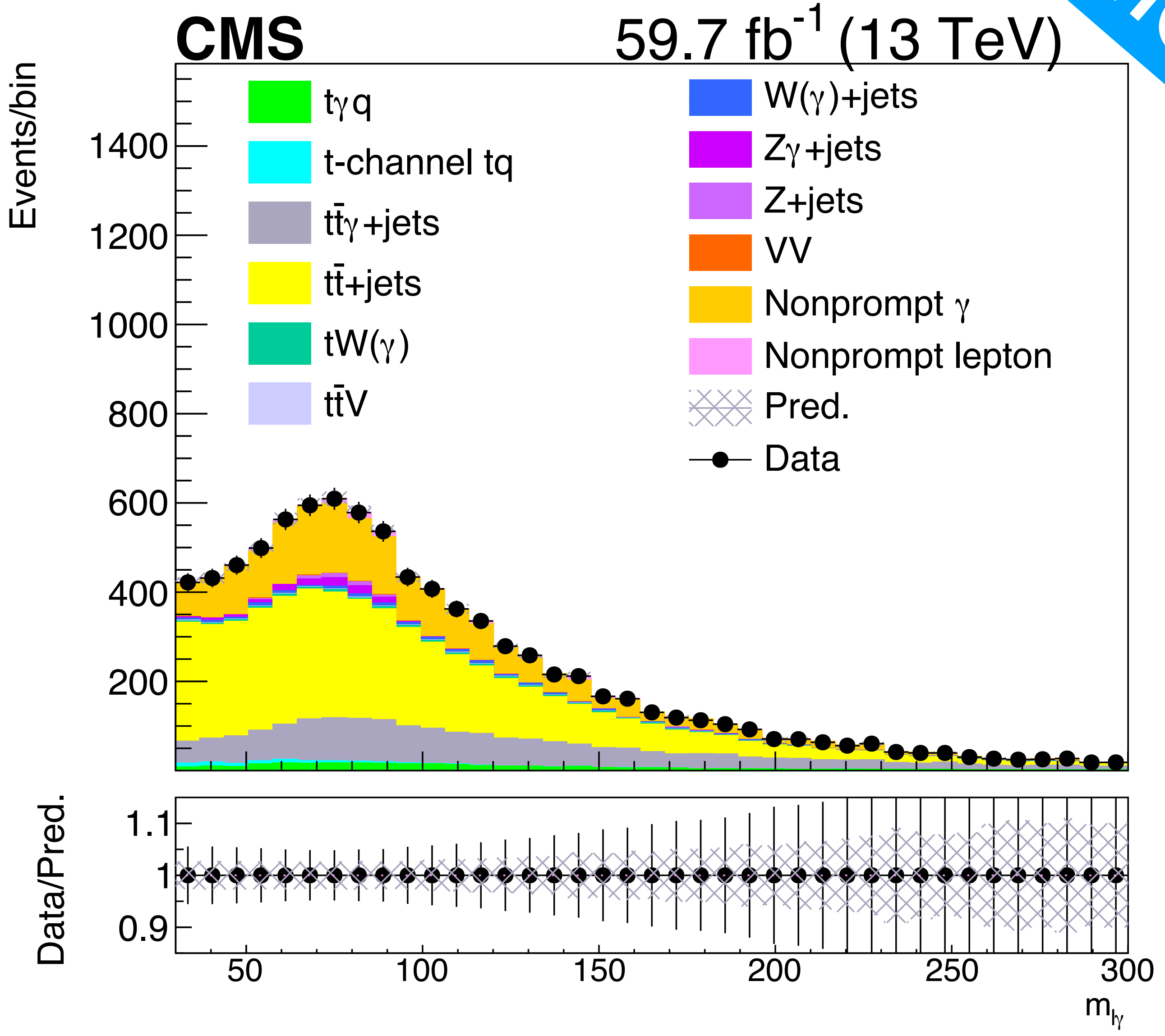
back

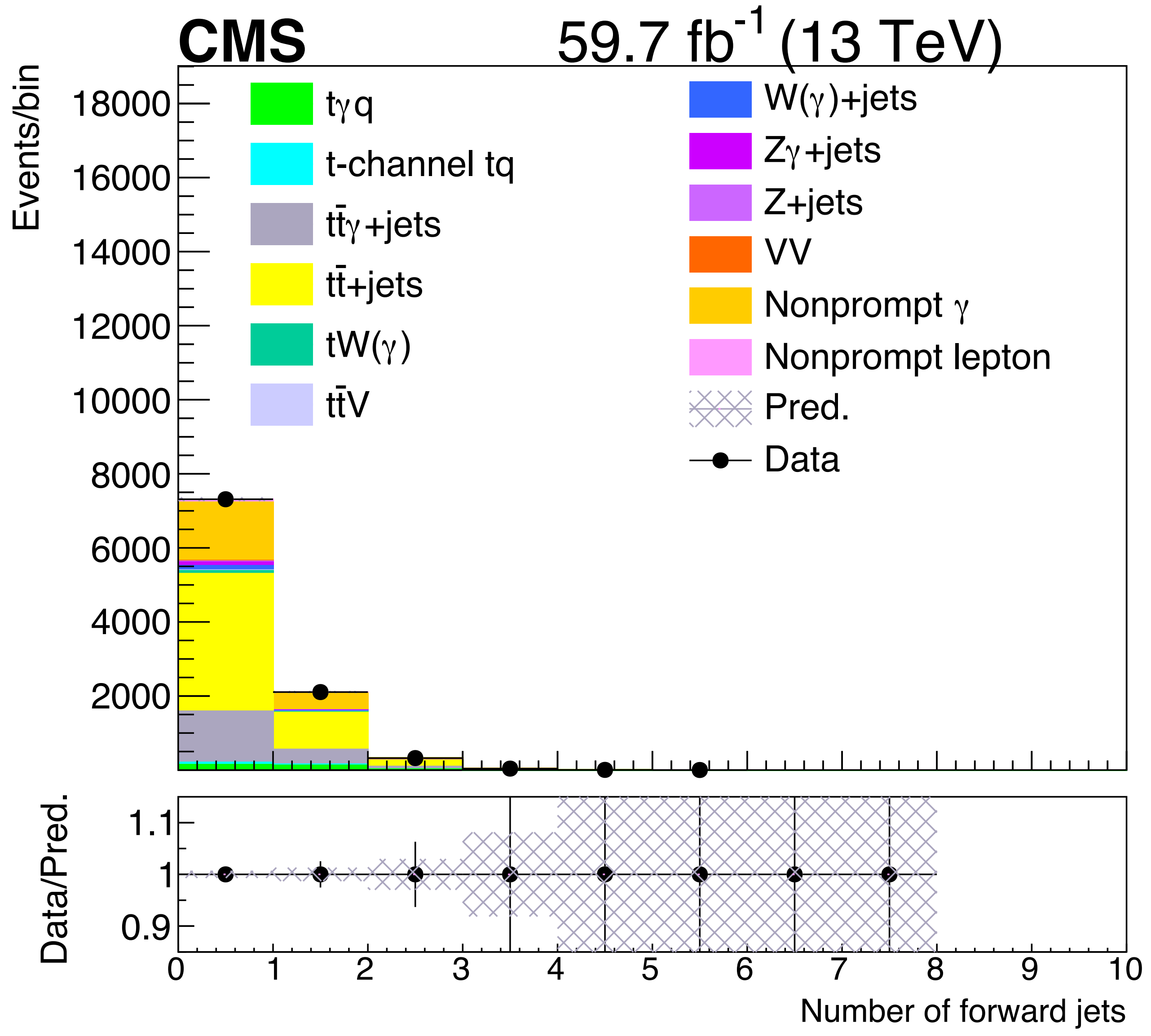
$t\bar{t}\gamma$ SR plots — $N_{\mathrm{j}} \geq 2 \;\; N_{\mathrm{b-jets}} = 2$

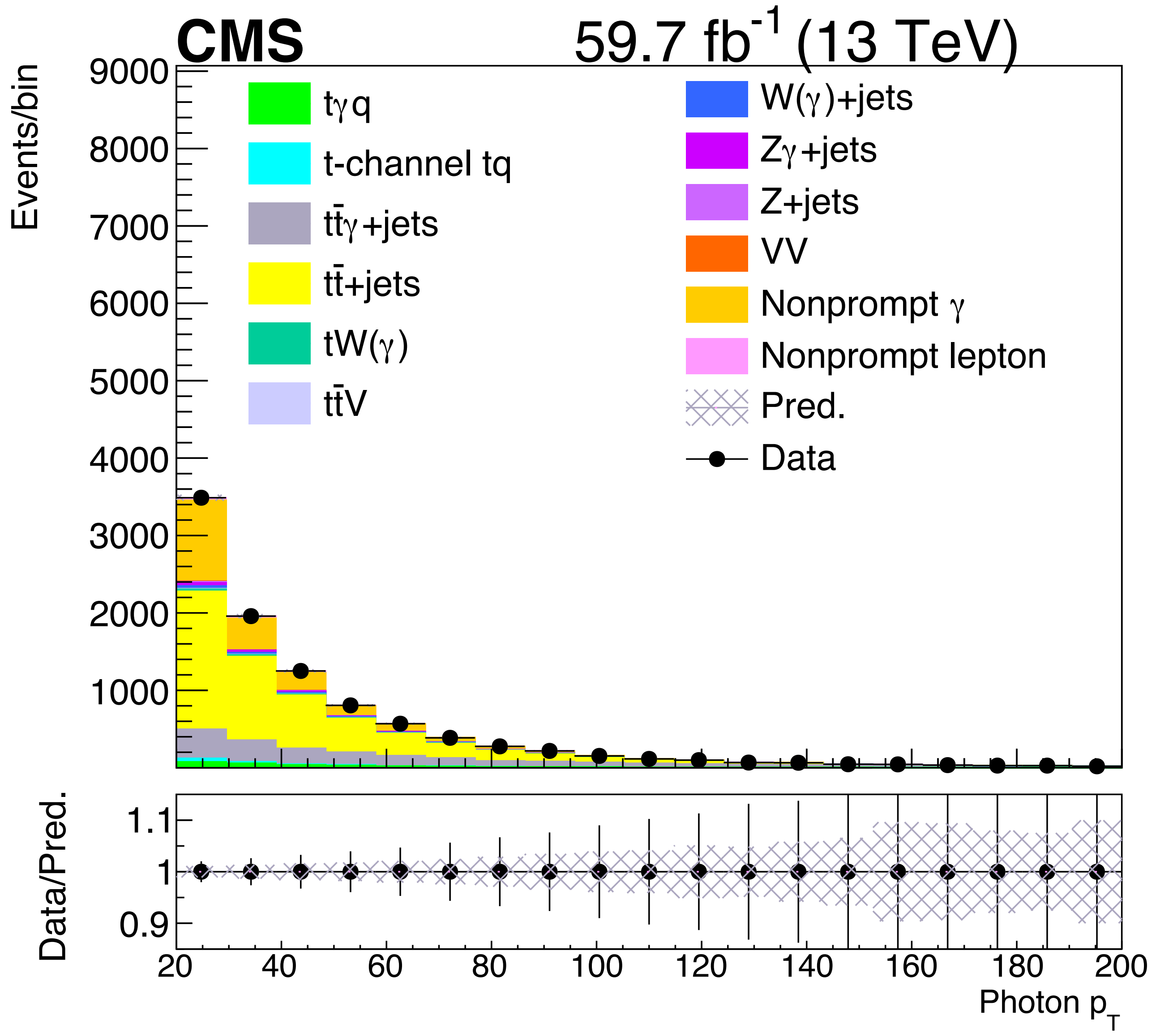
2018 MC only

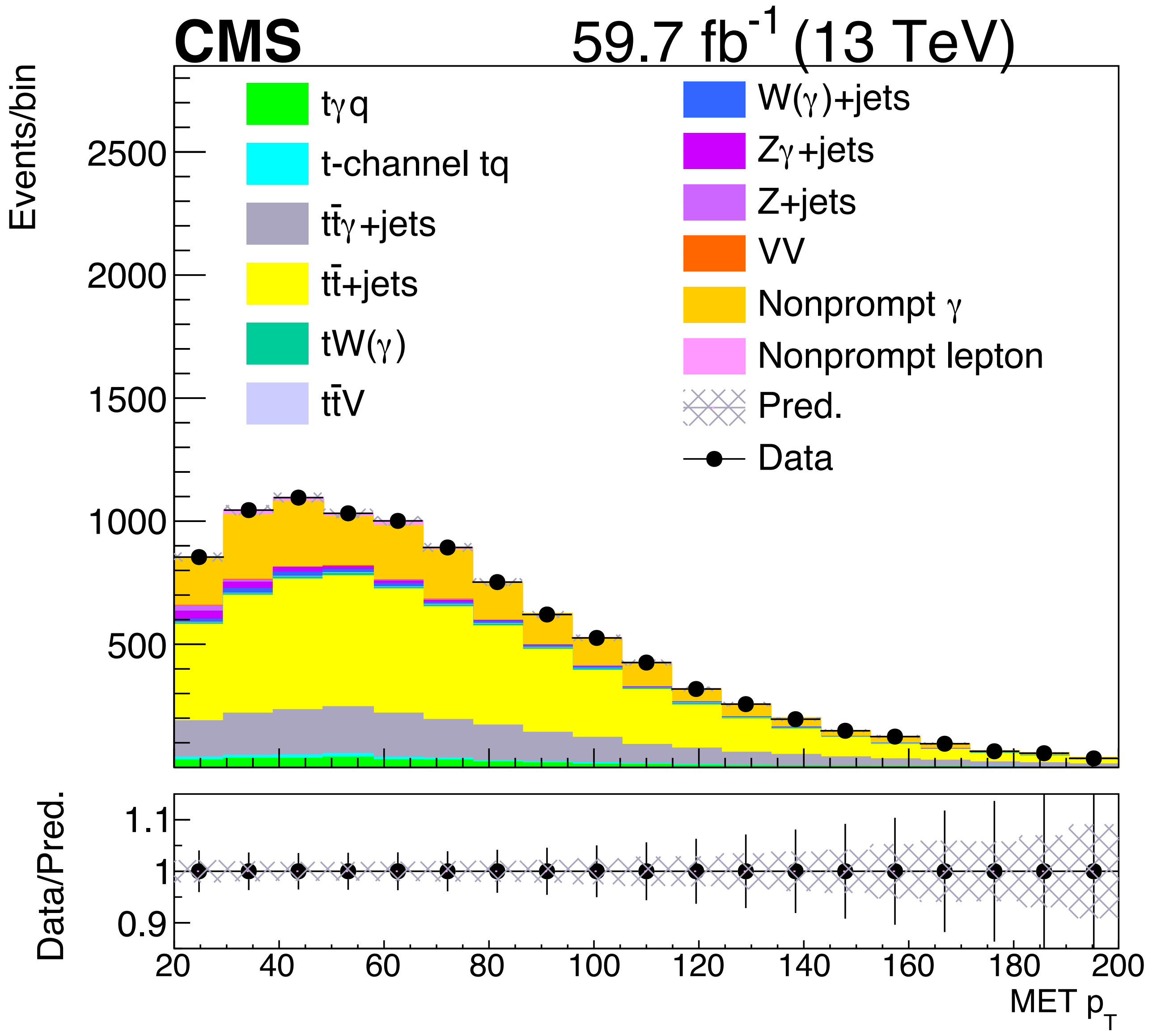








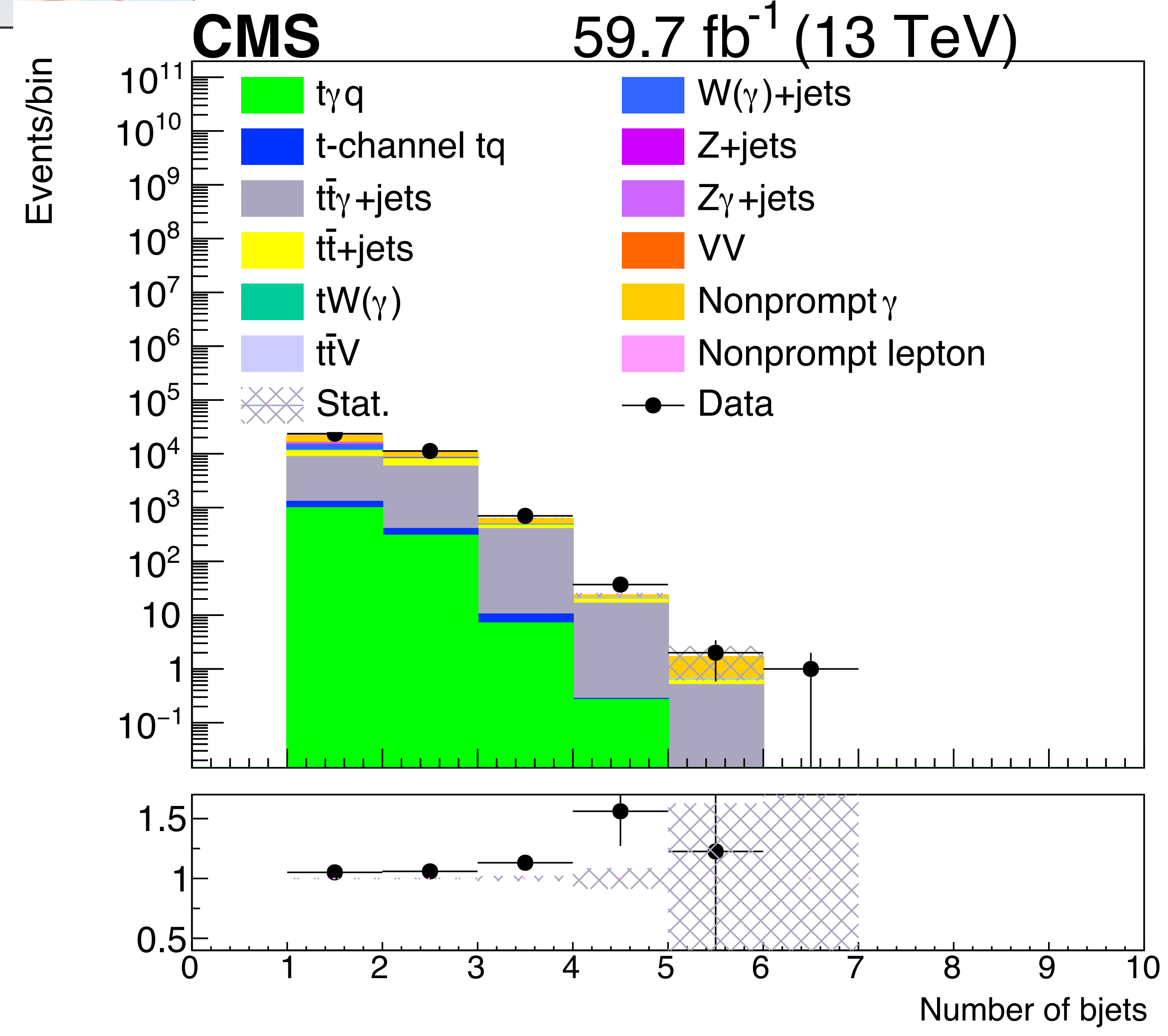


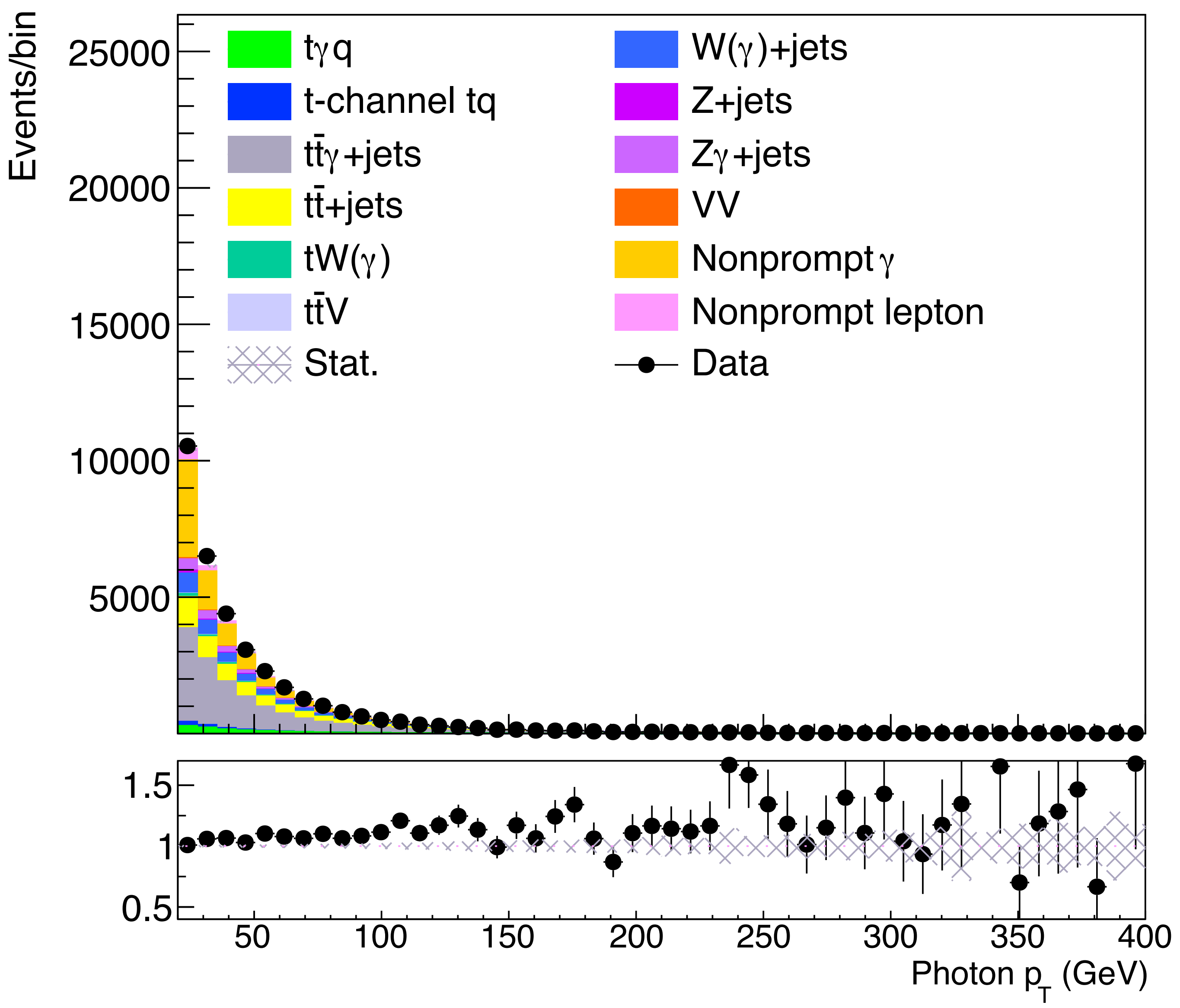


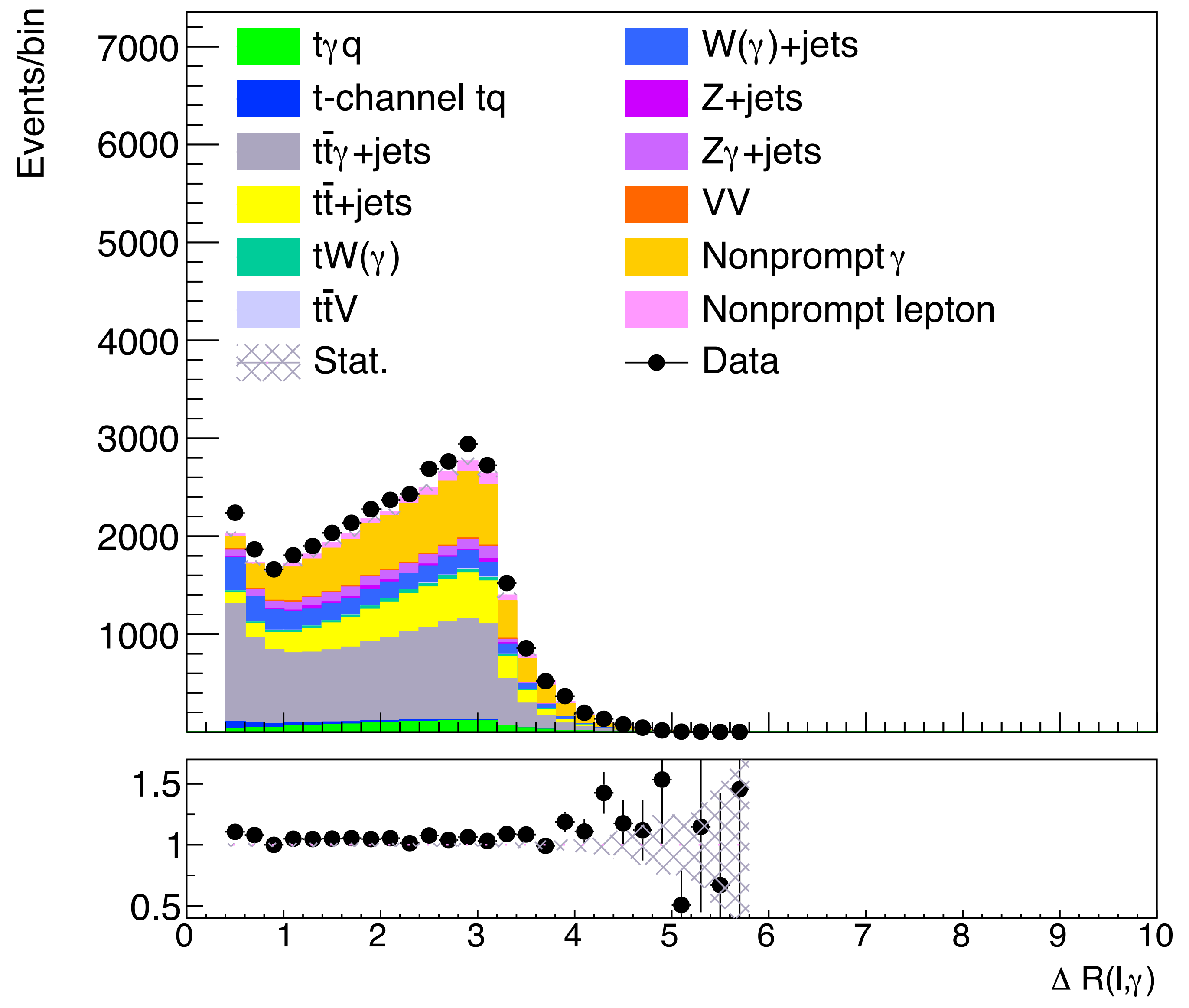
Use $t\bar{t}\gamma$ NLO, so the $t\bar{t}$ (yellow) has large contribution

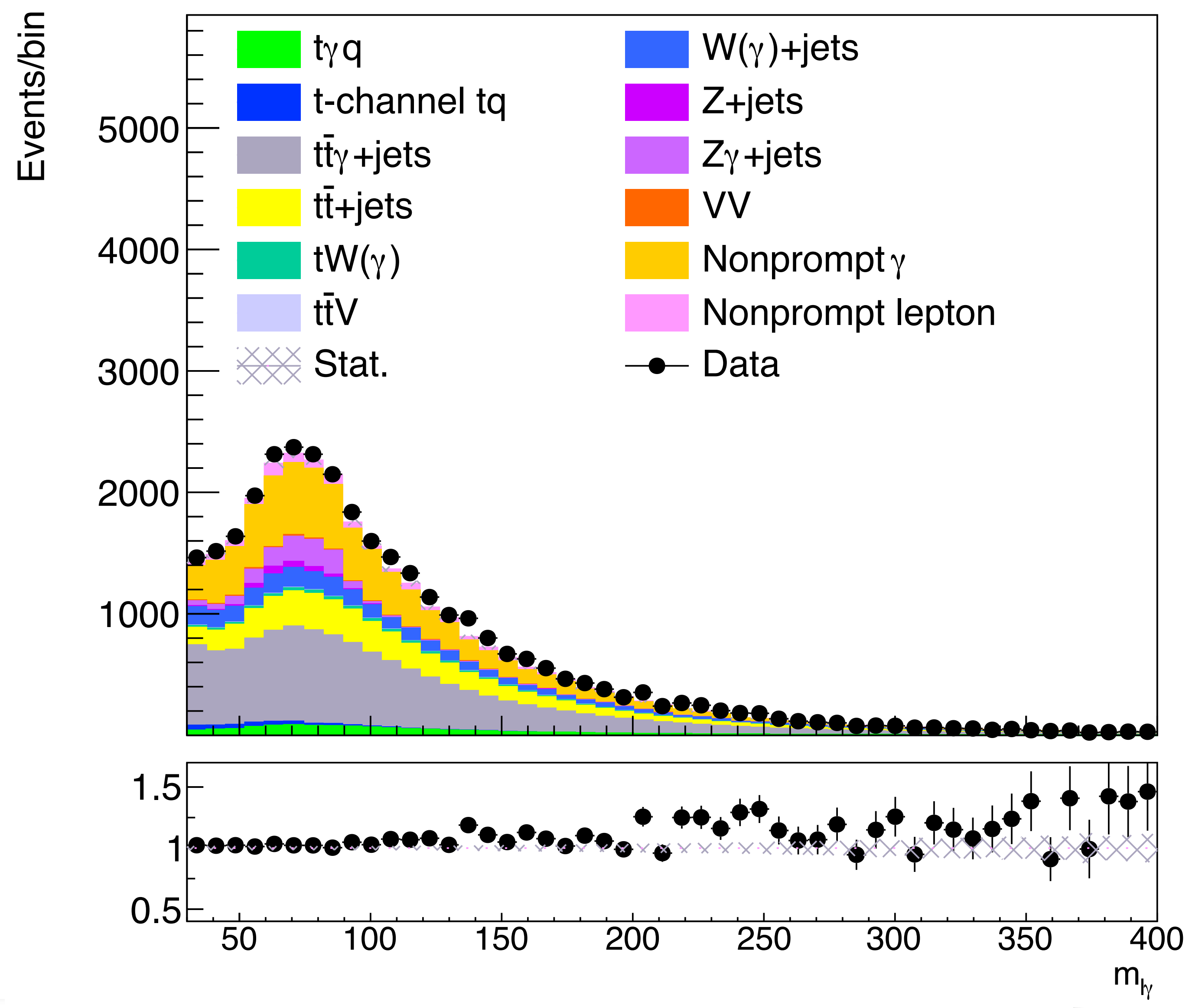
back

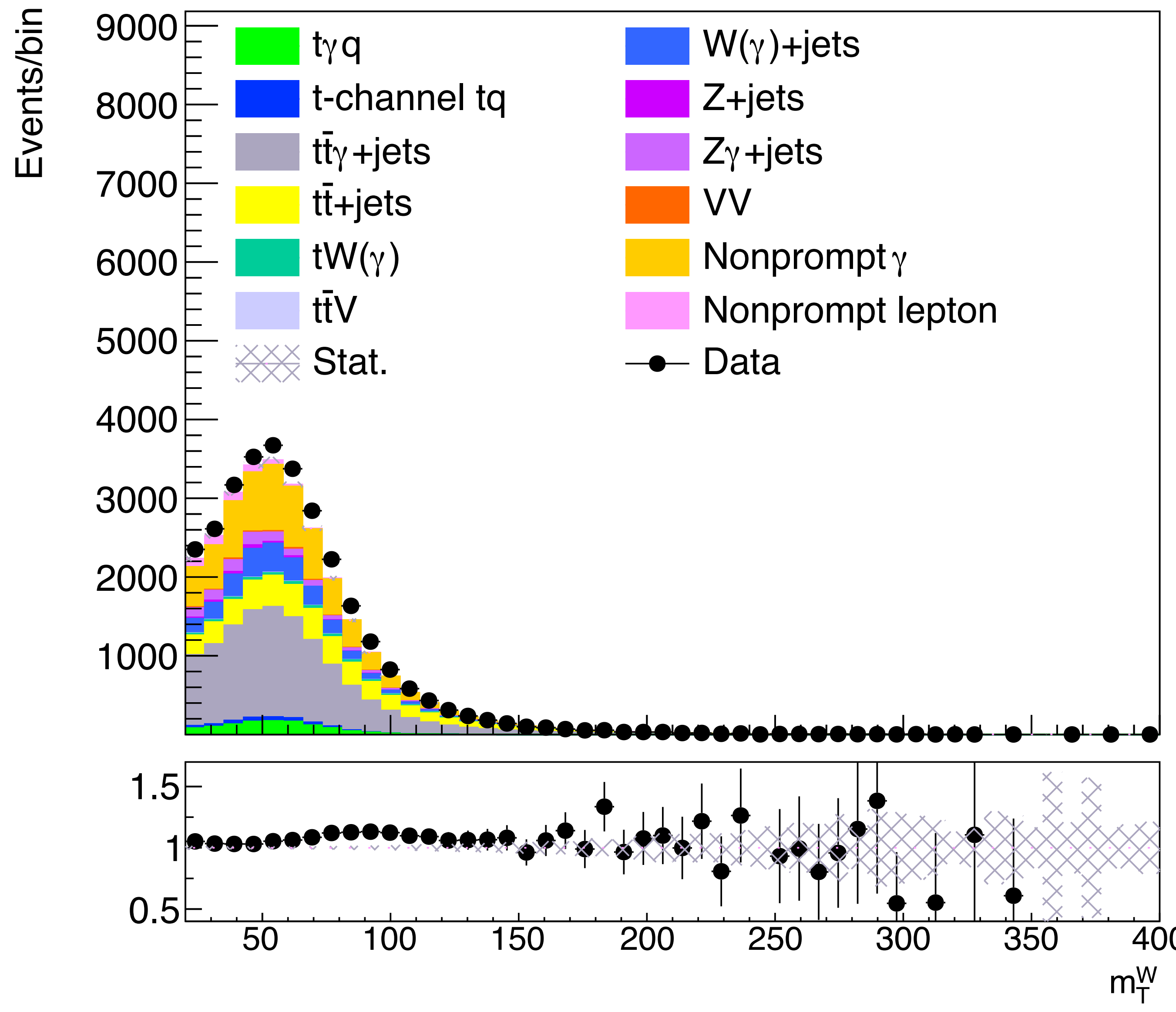
SR plots — $N_{\mathrm{j}} \geq 2$ $N_{\mathrm{b-jets}} \geq 1$

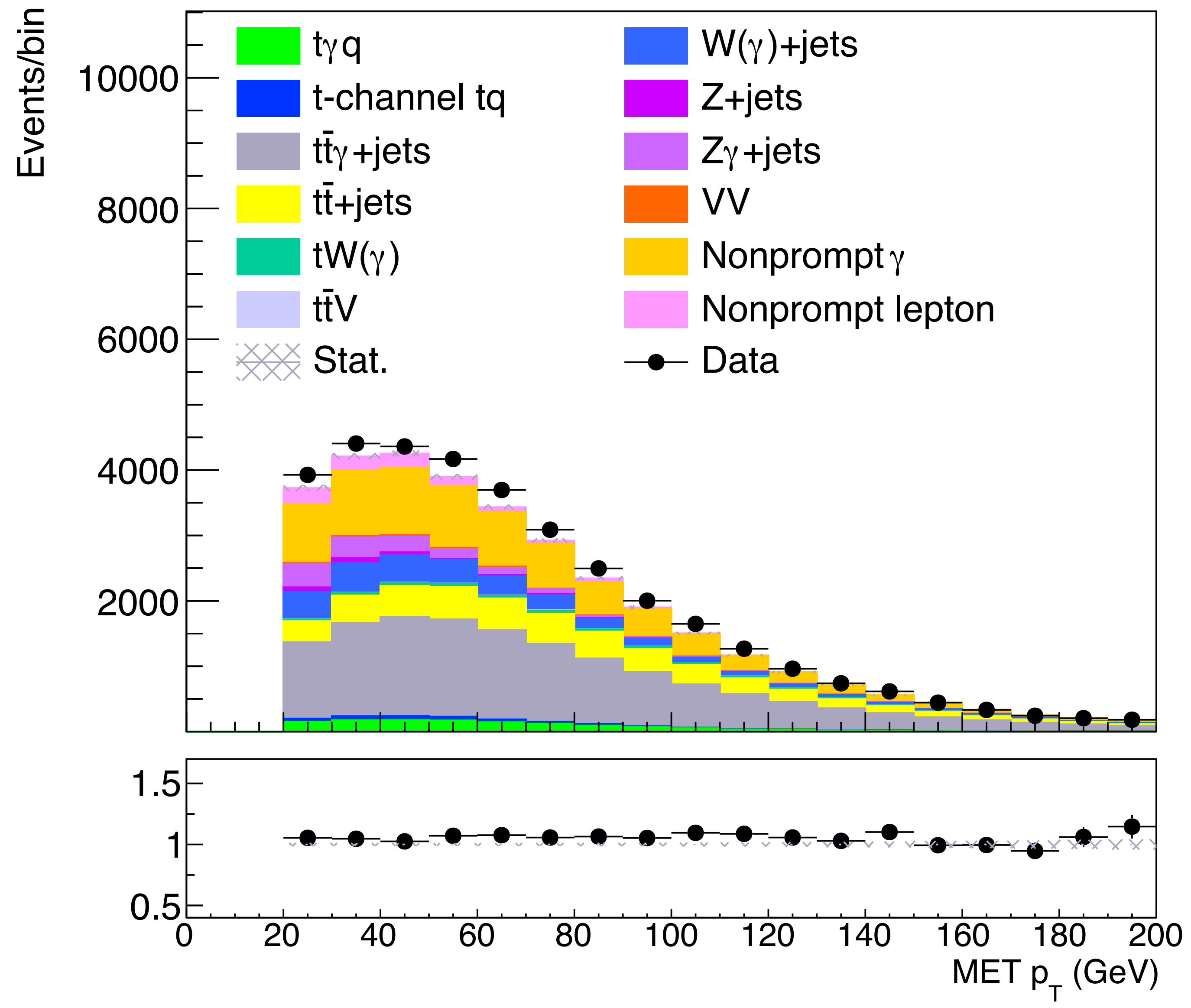








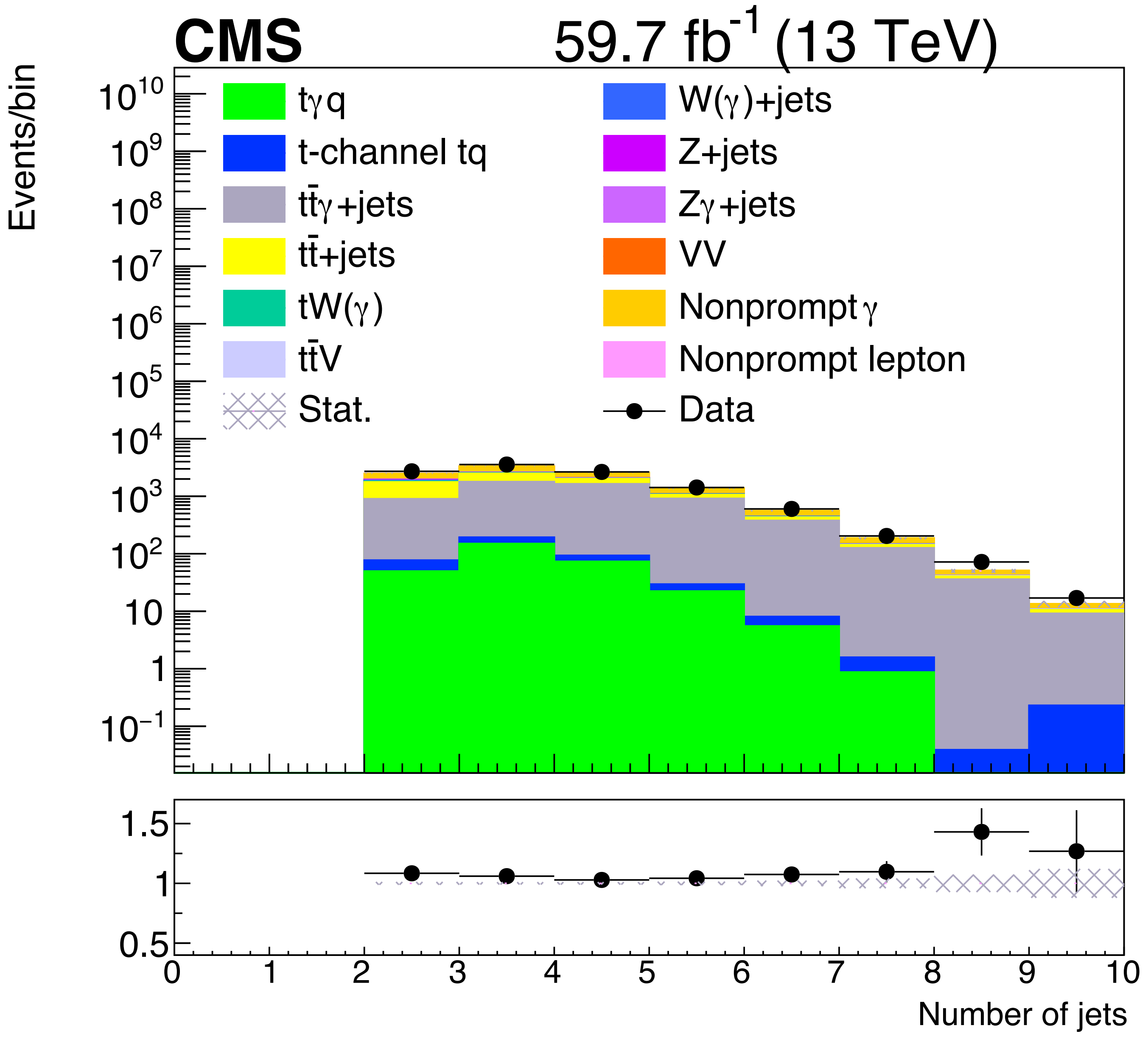


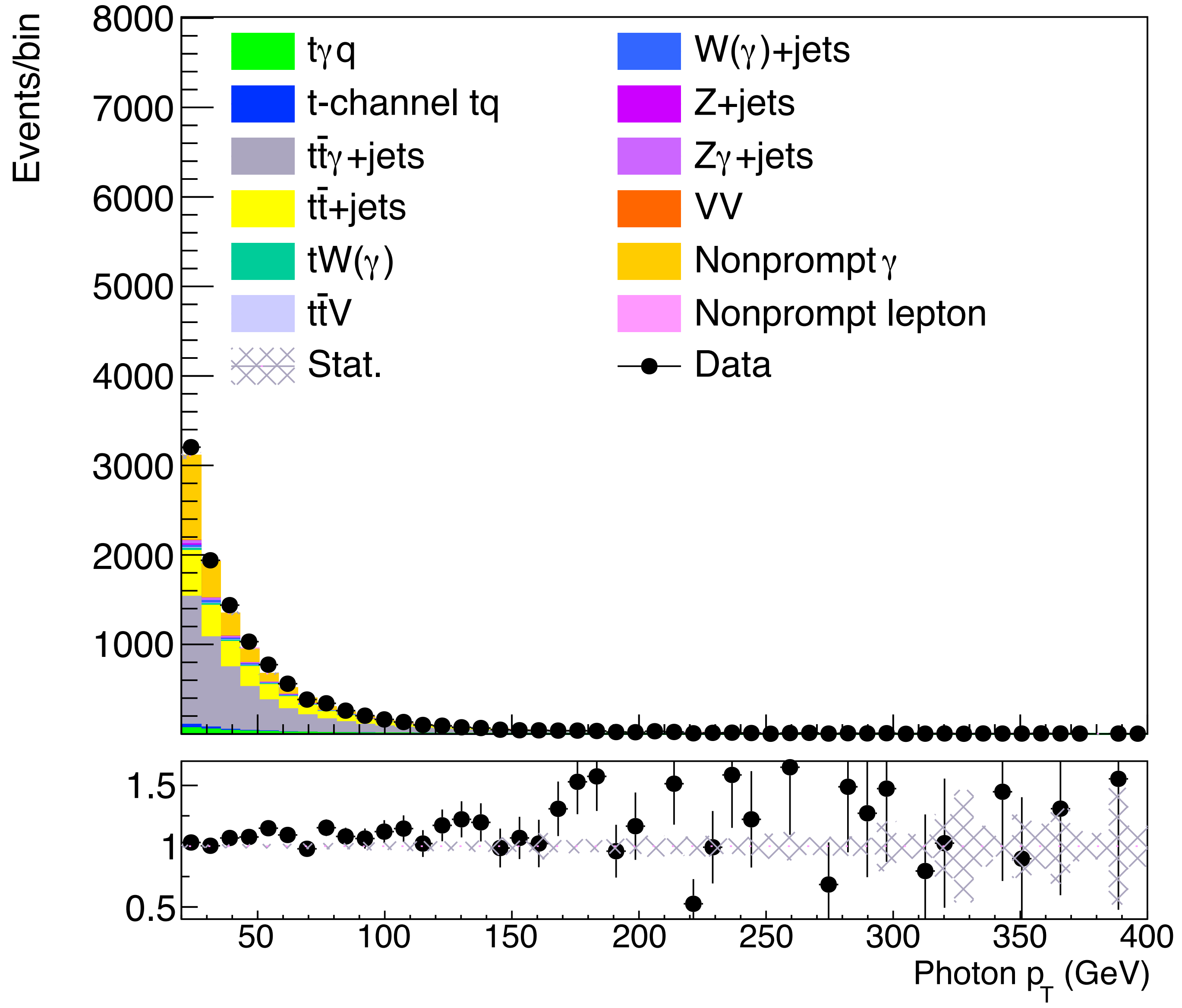


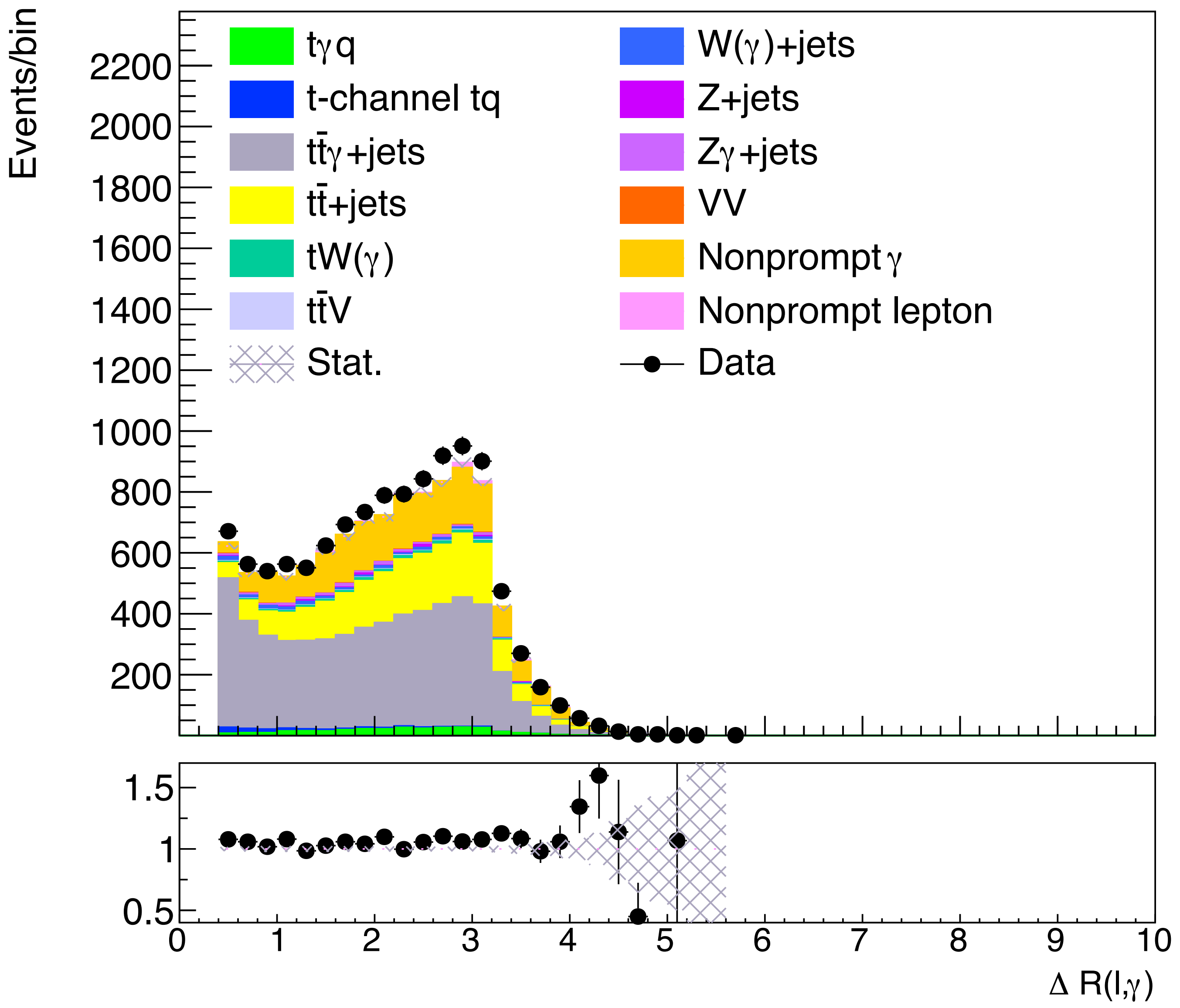
Use tt̄γ LO with k-factor 1.86

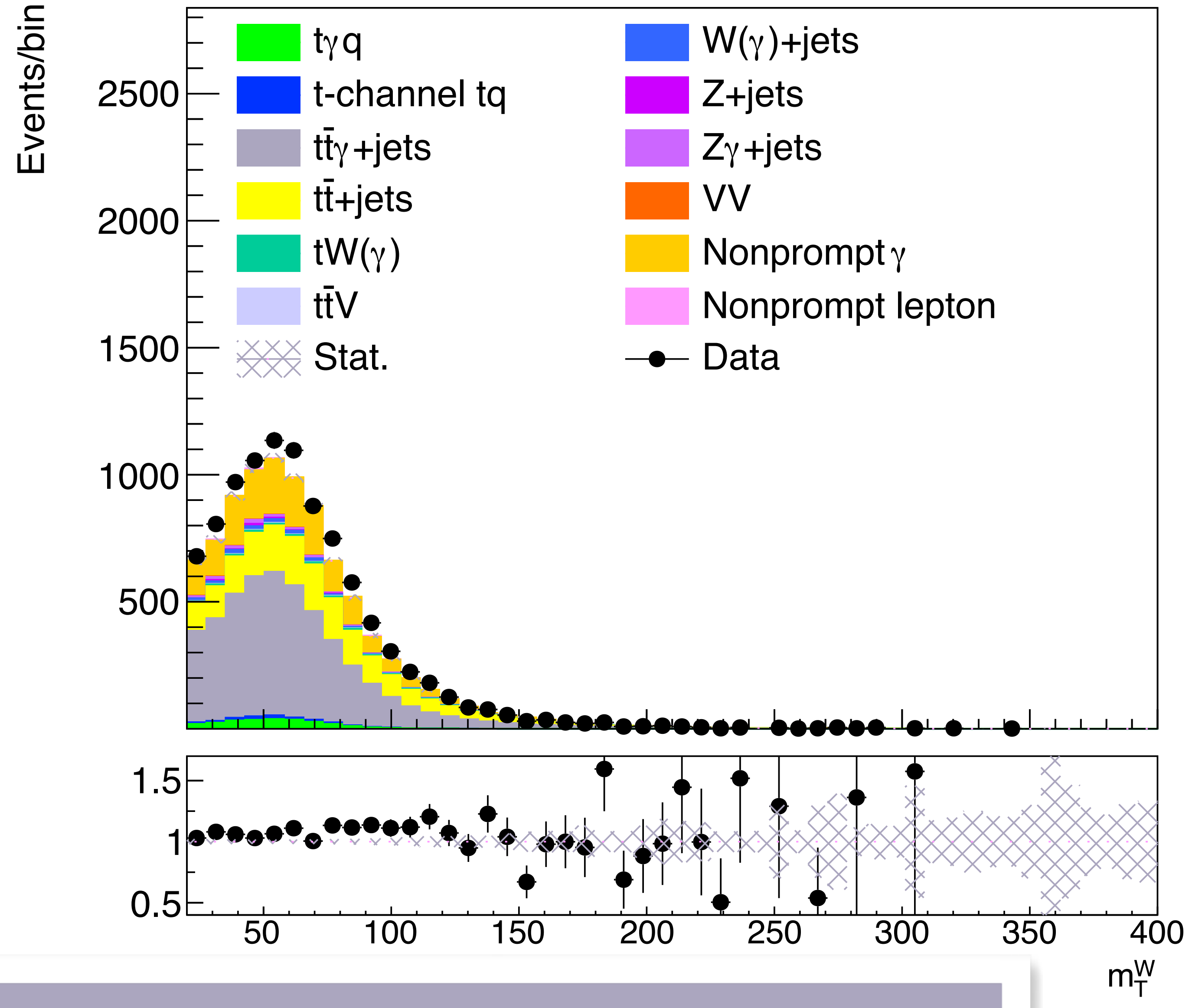
back

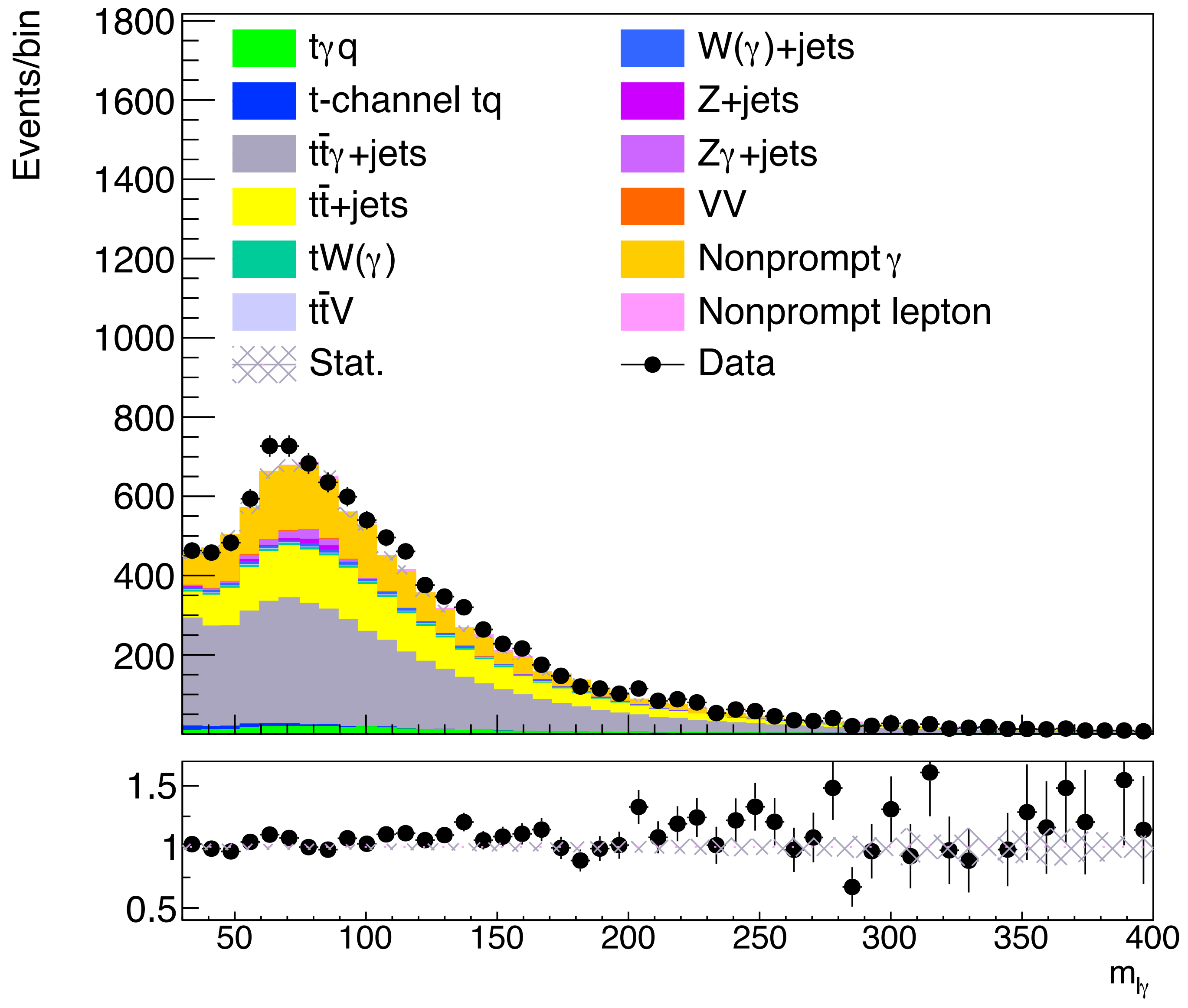
$t\bar{t}\gamma$ SR plots — $N_j \geq 2$ $N_{b-jets} = 2$

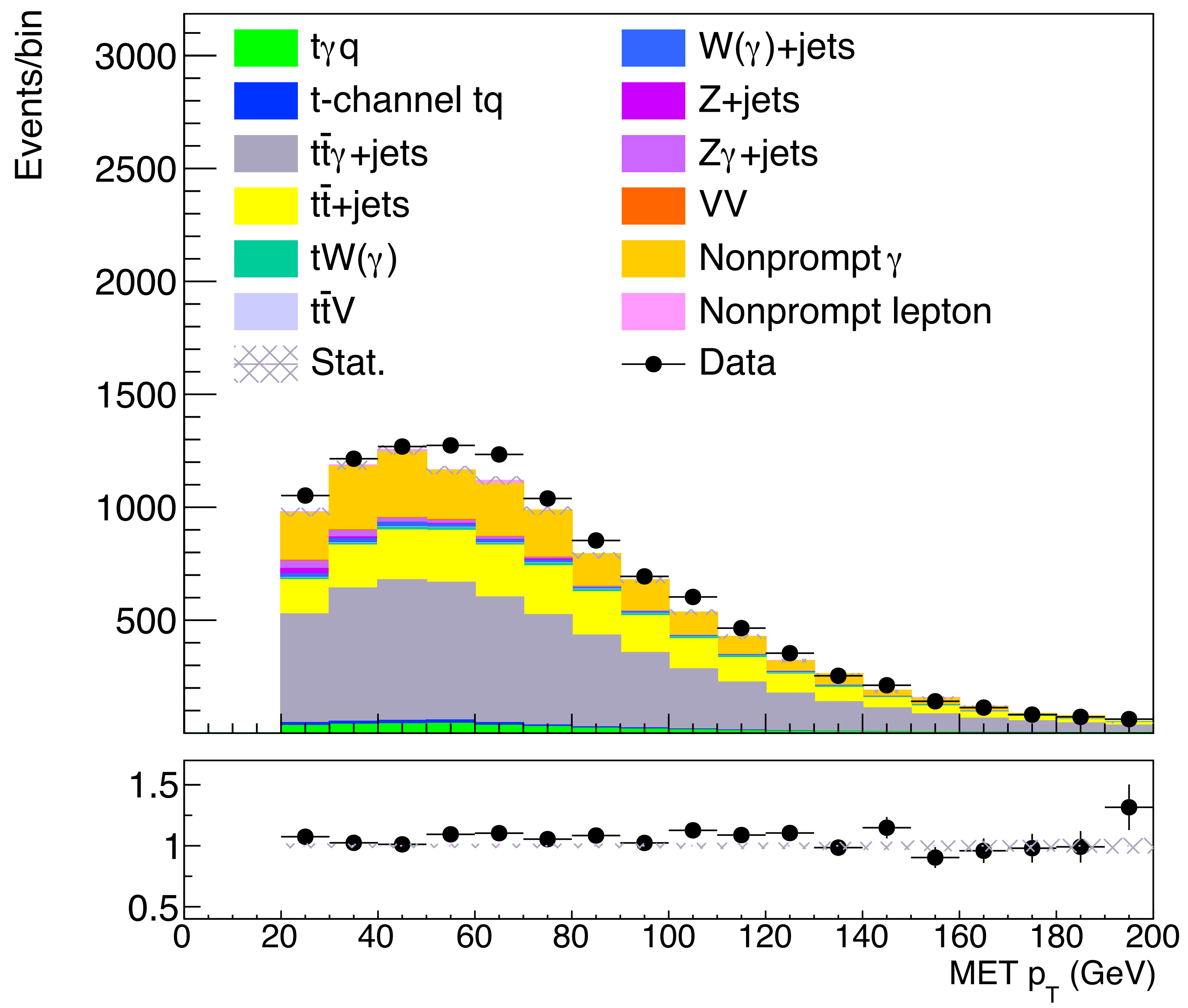










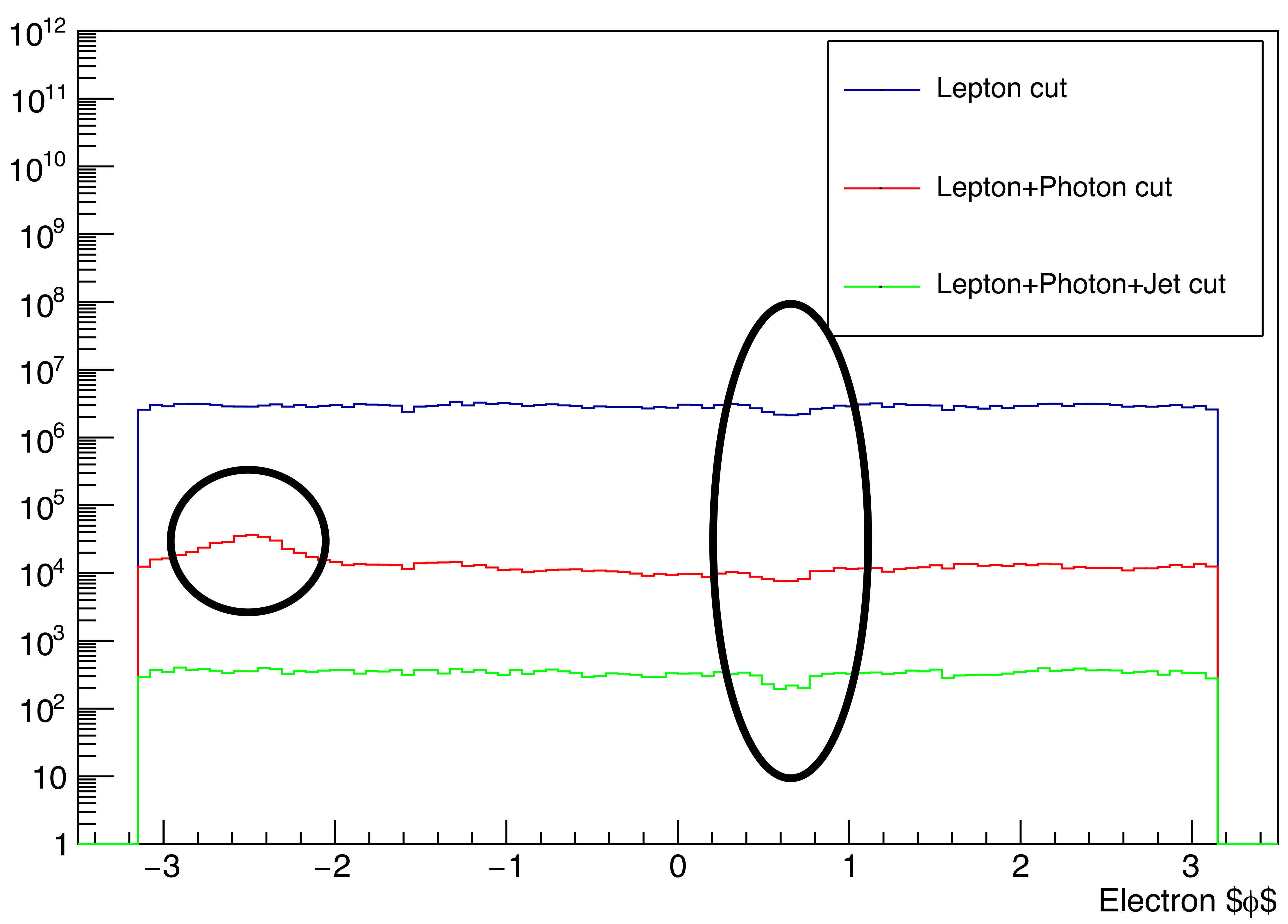


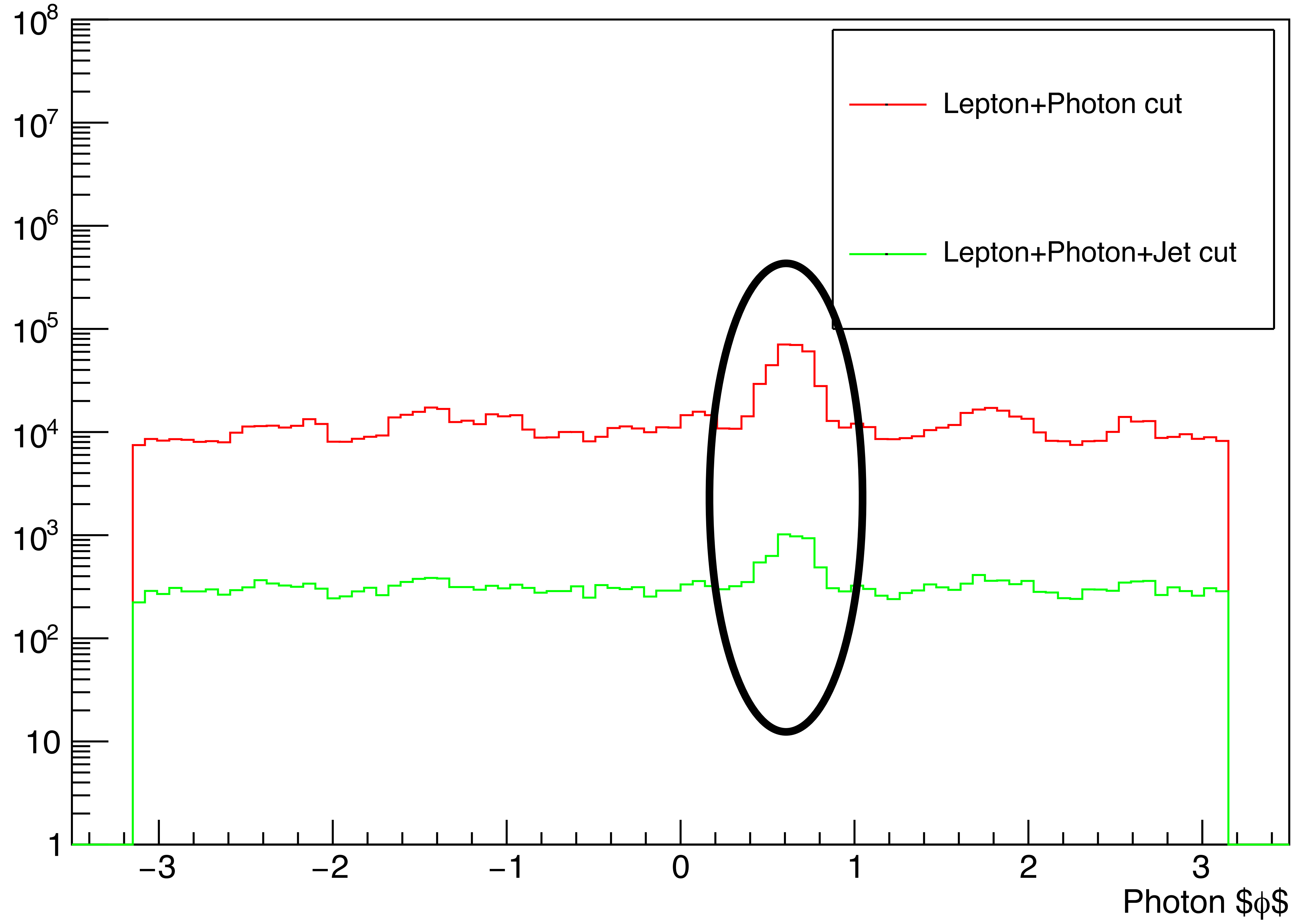
Use $t\bar{t}\gamma$ LO with k-factor 1.86

back

Glance at electron data

2018 EGamma data

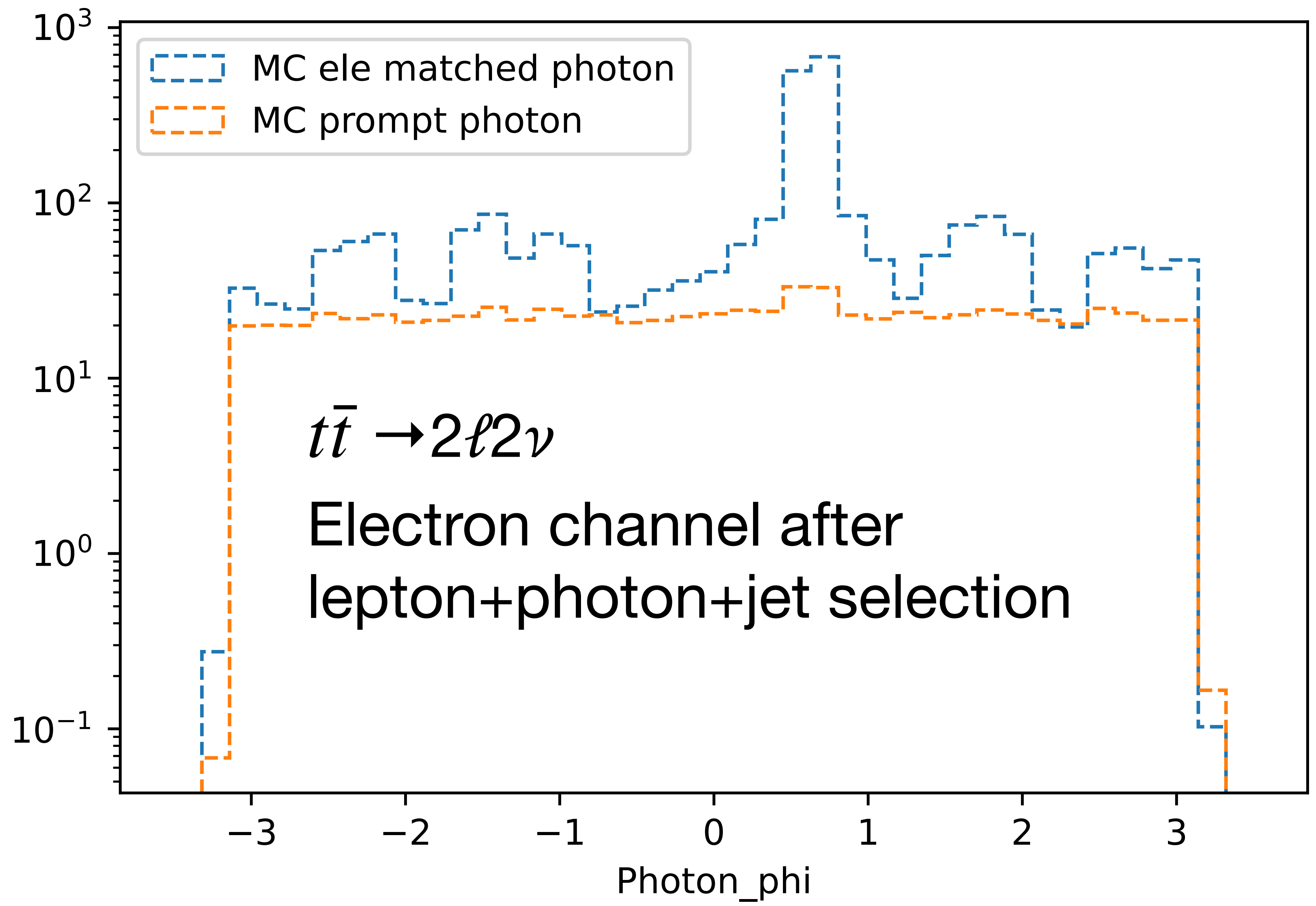


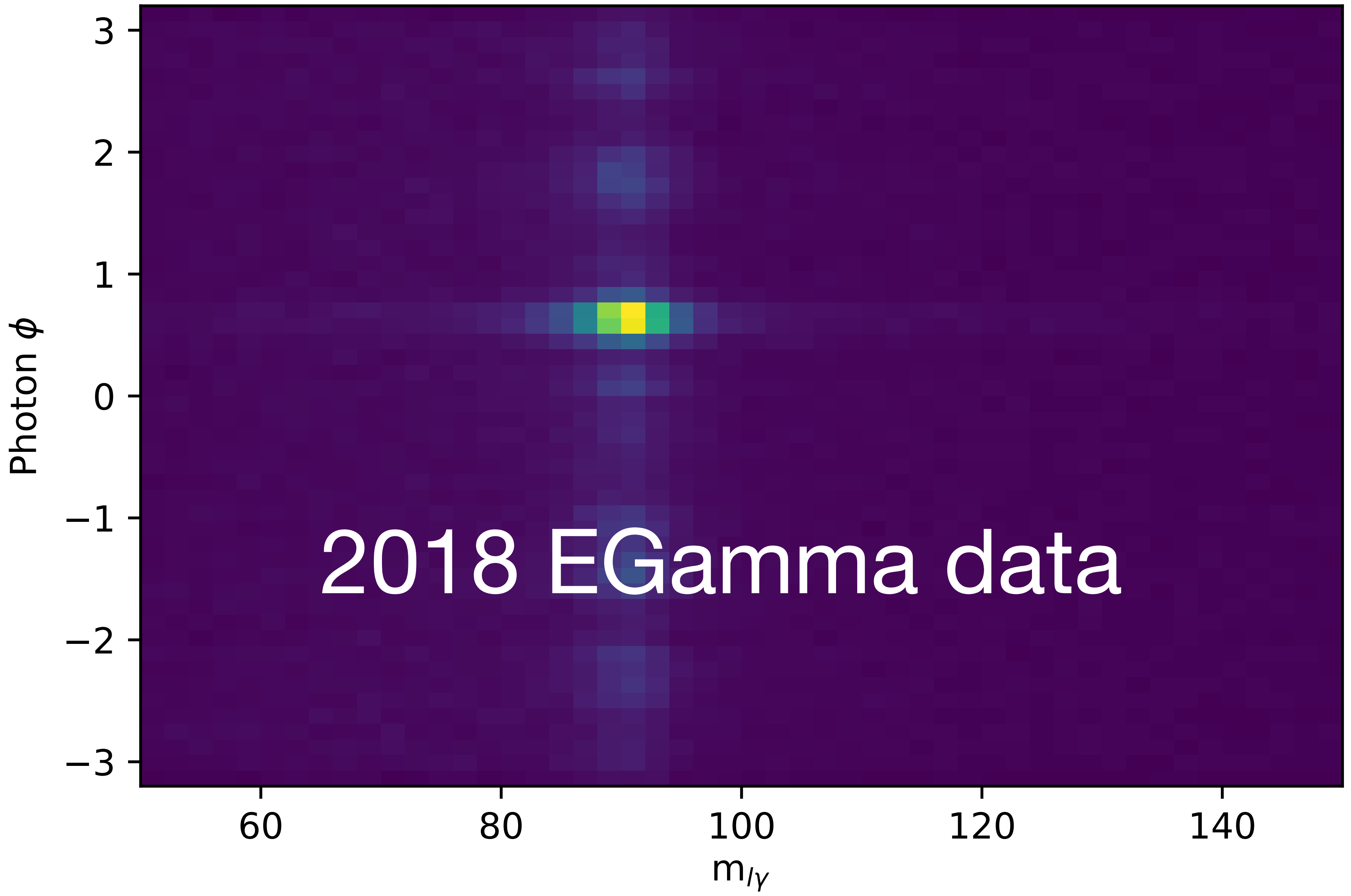


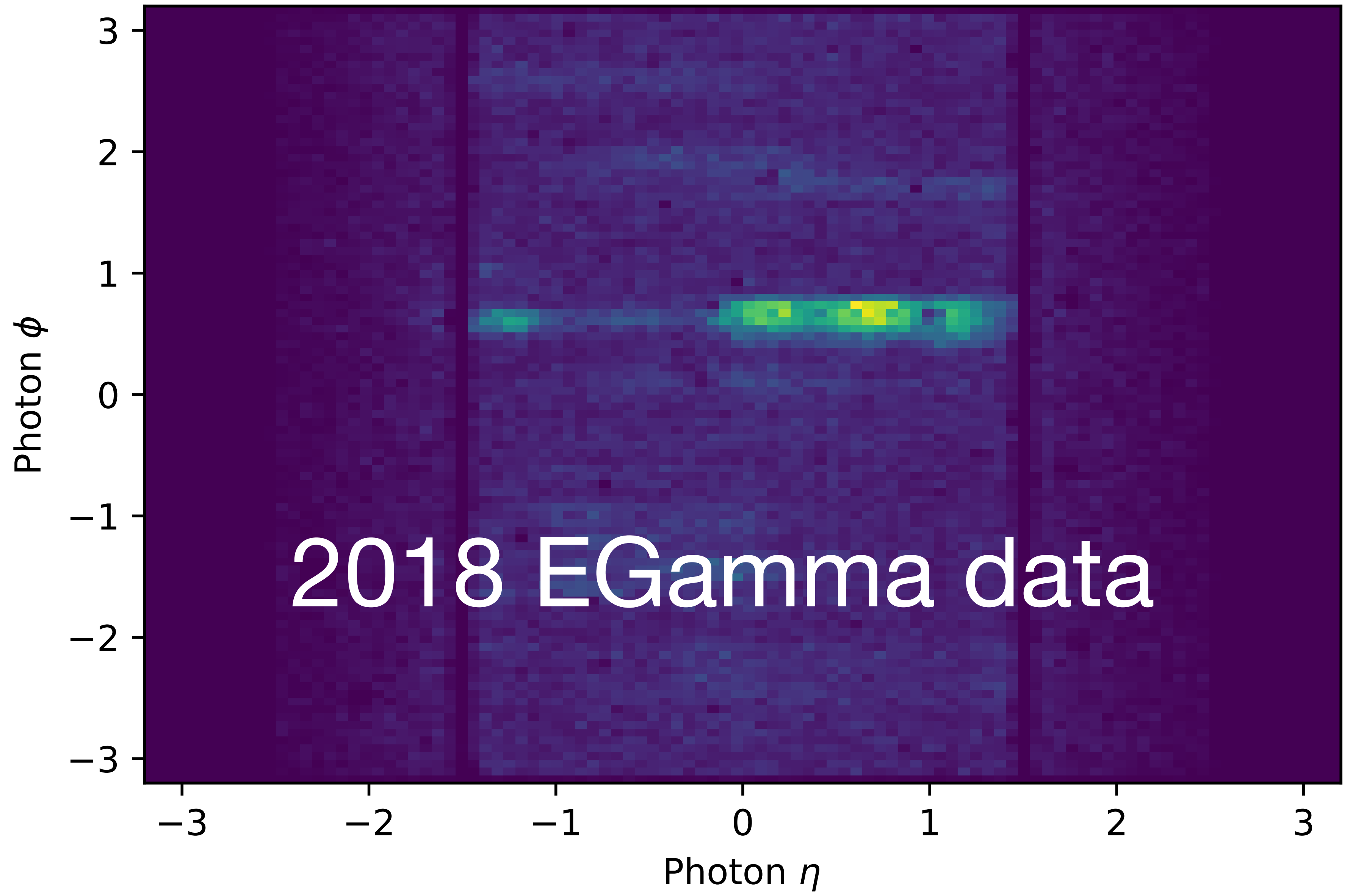
Pixel failure in phi around 0.5 to 0.7

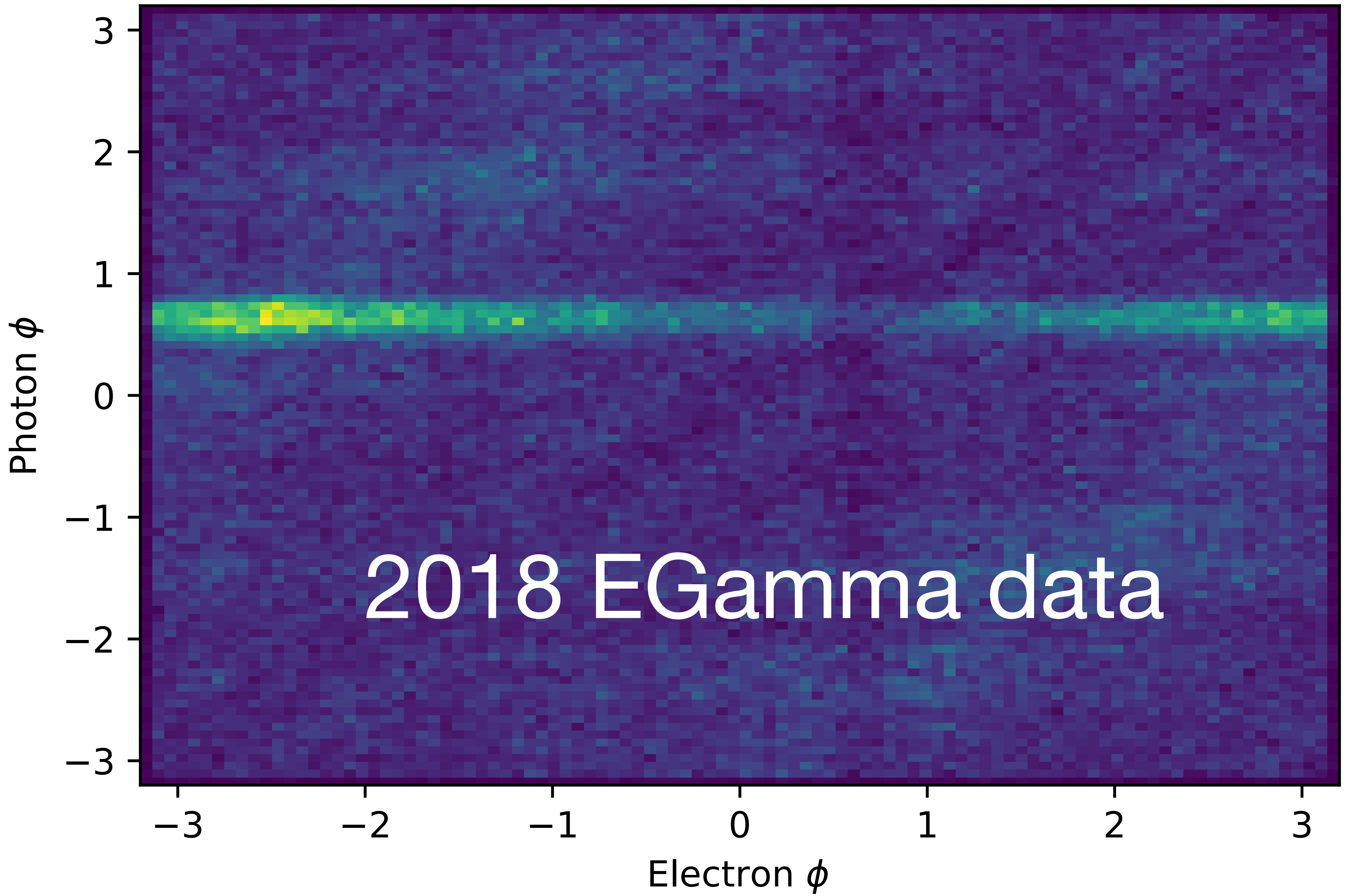
back

Photon ϕ spike in electron



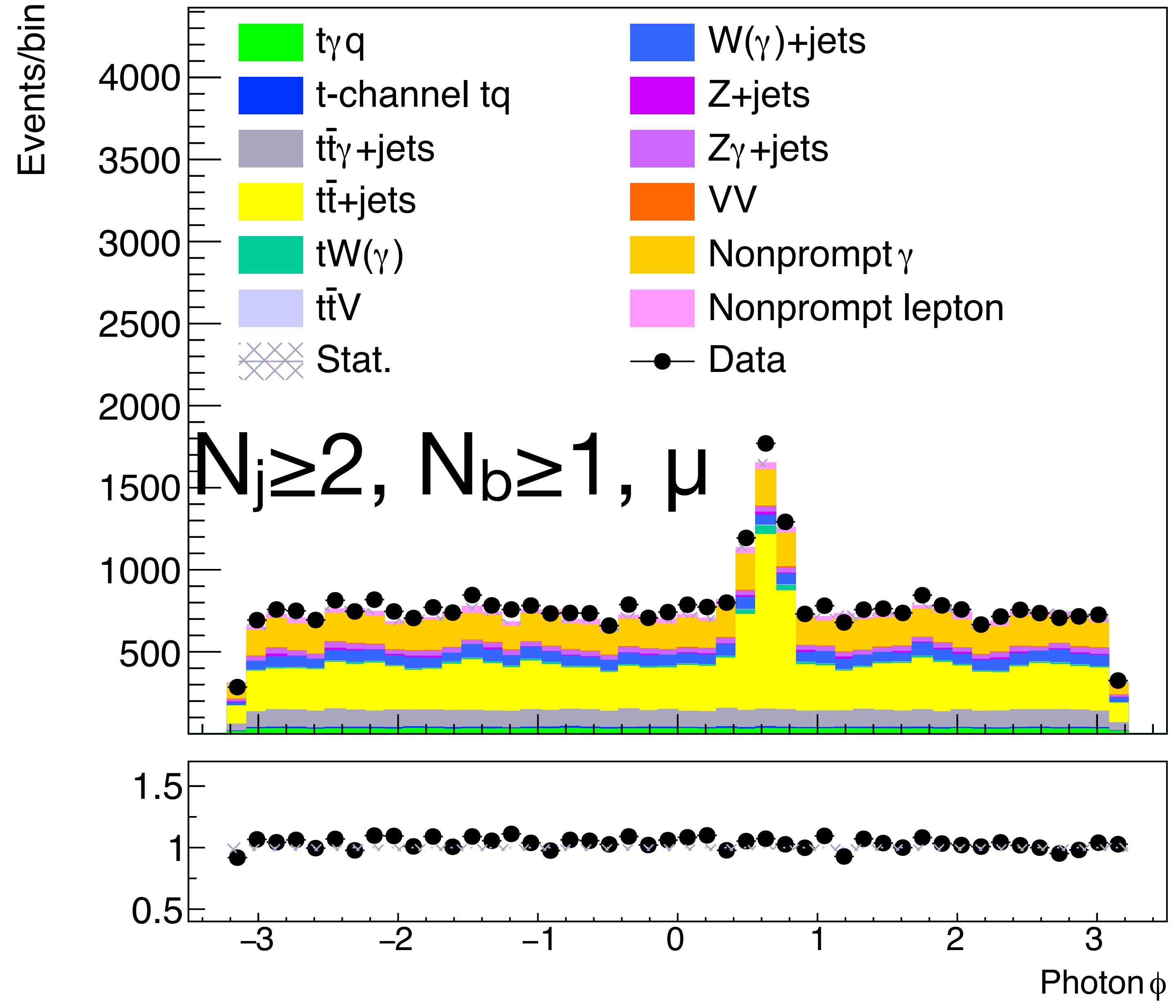


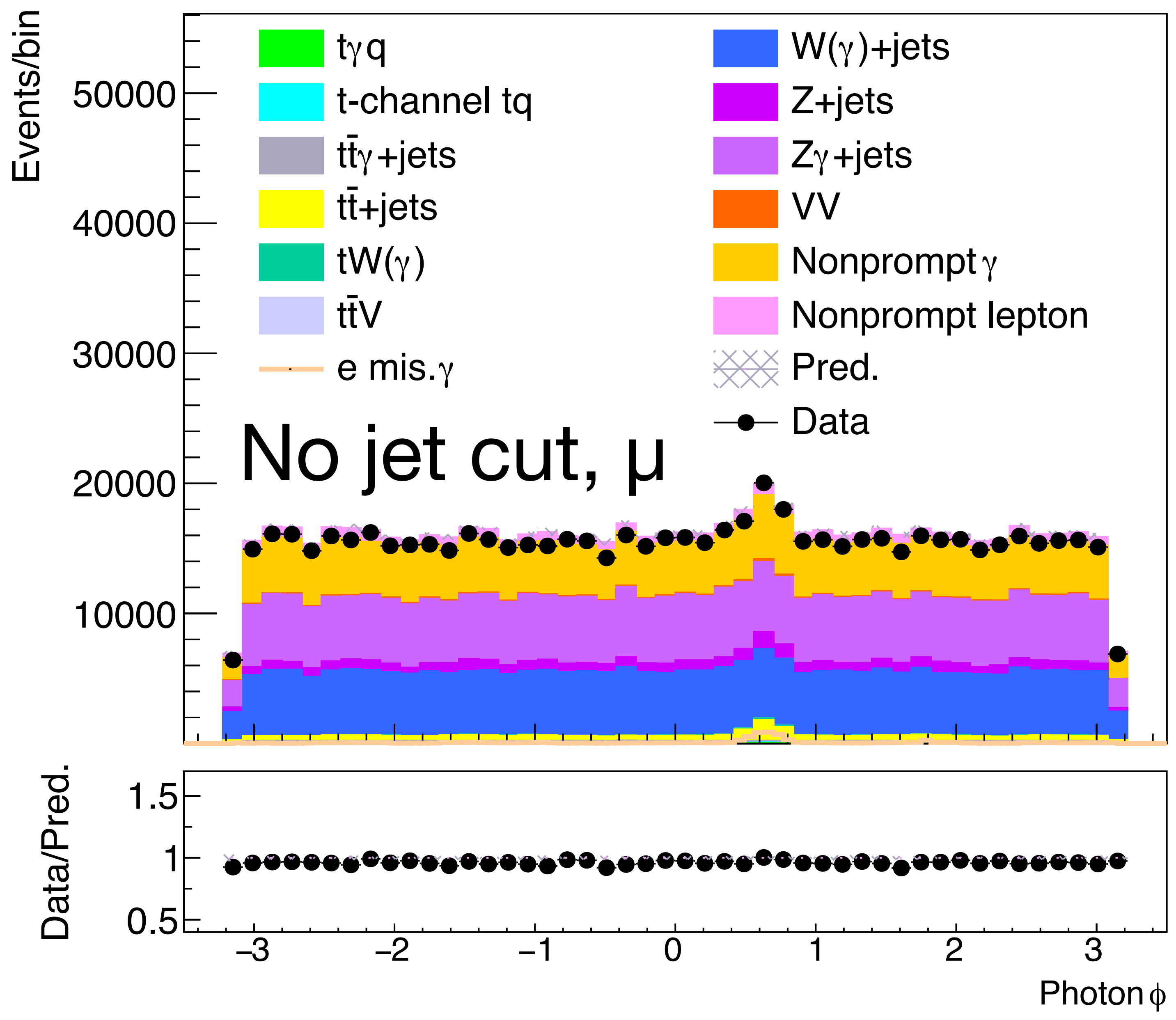


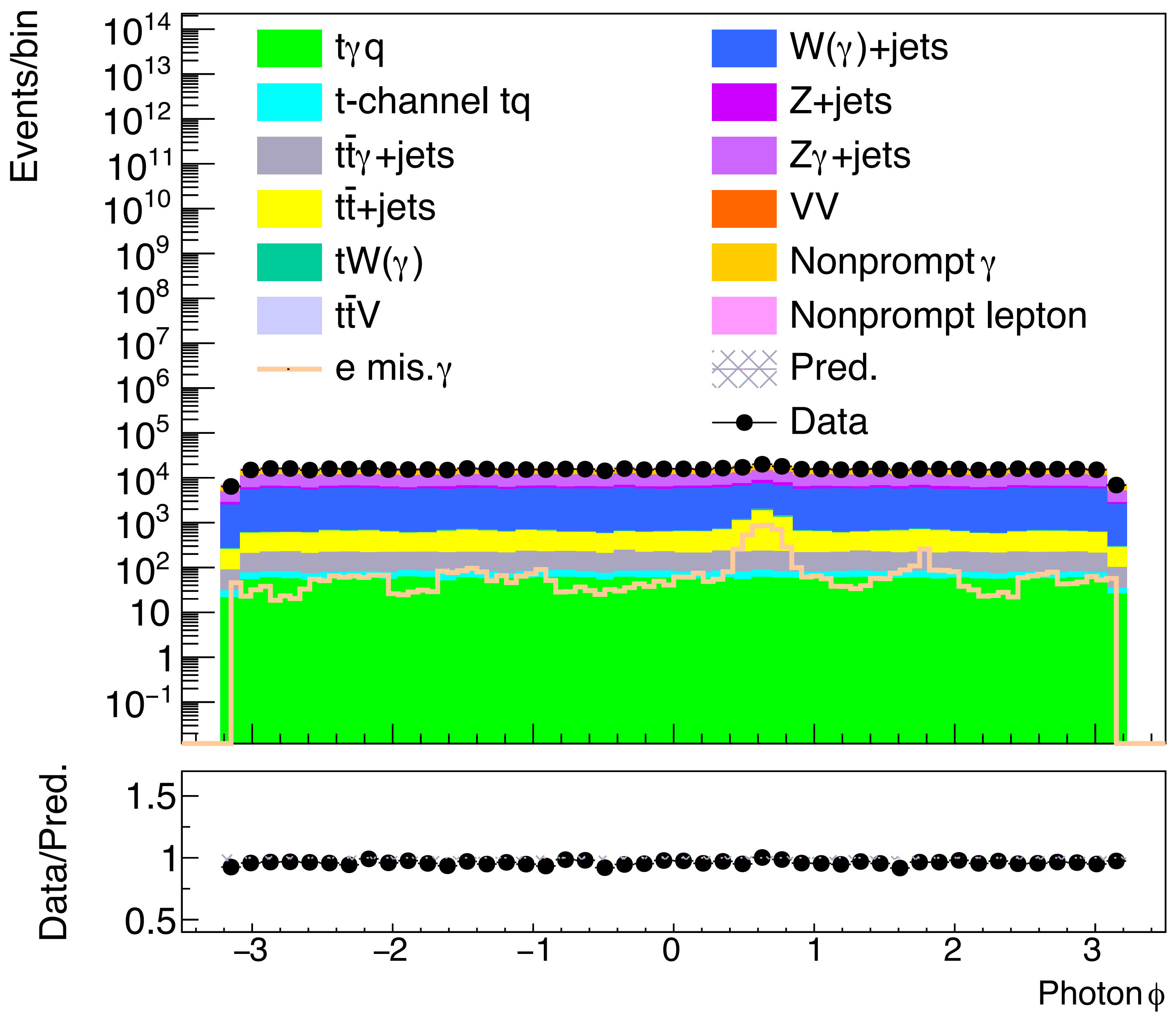


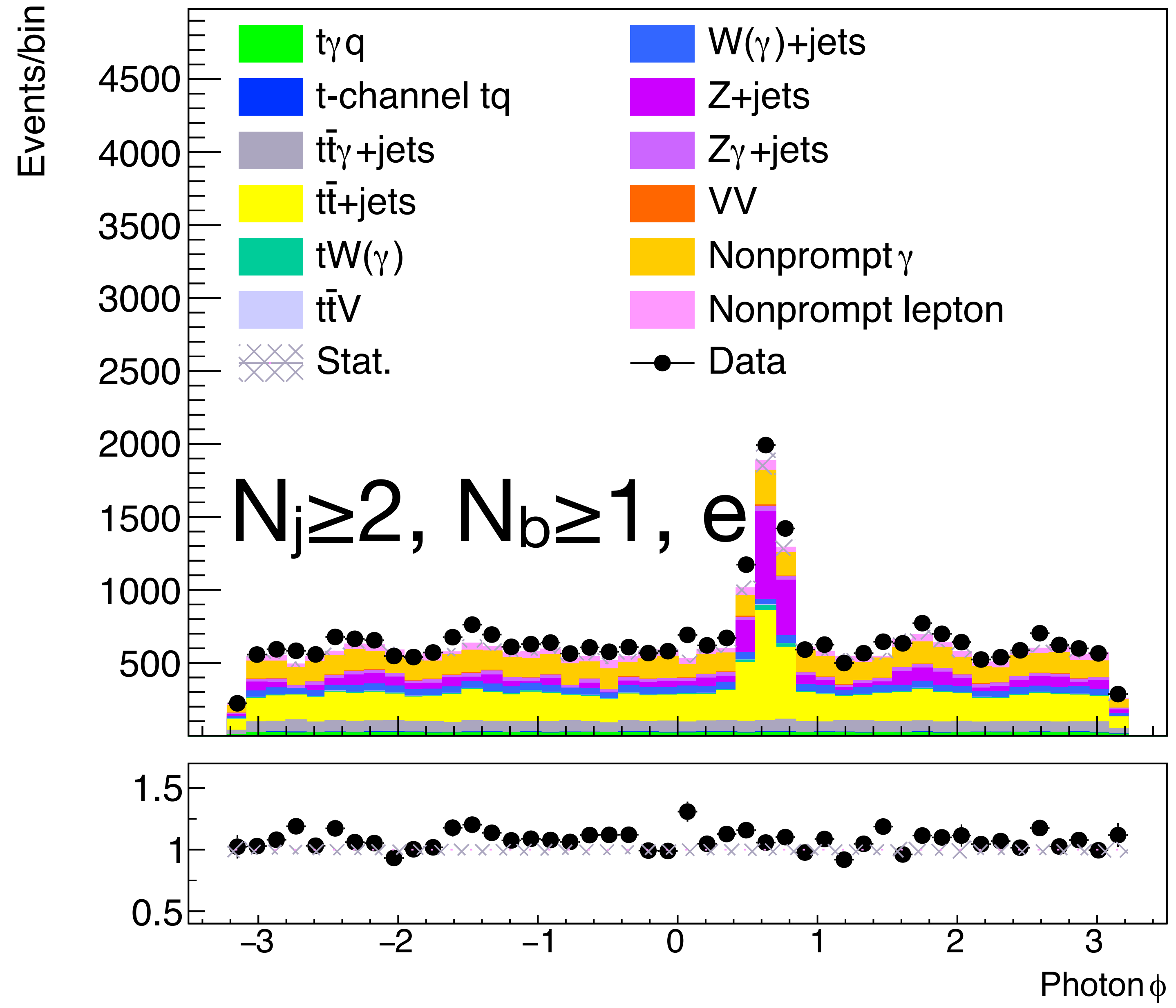
back

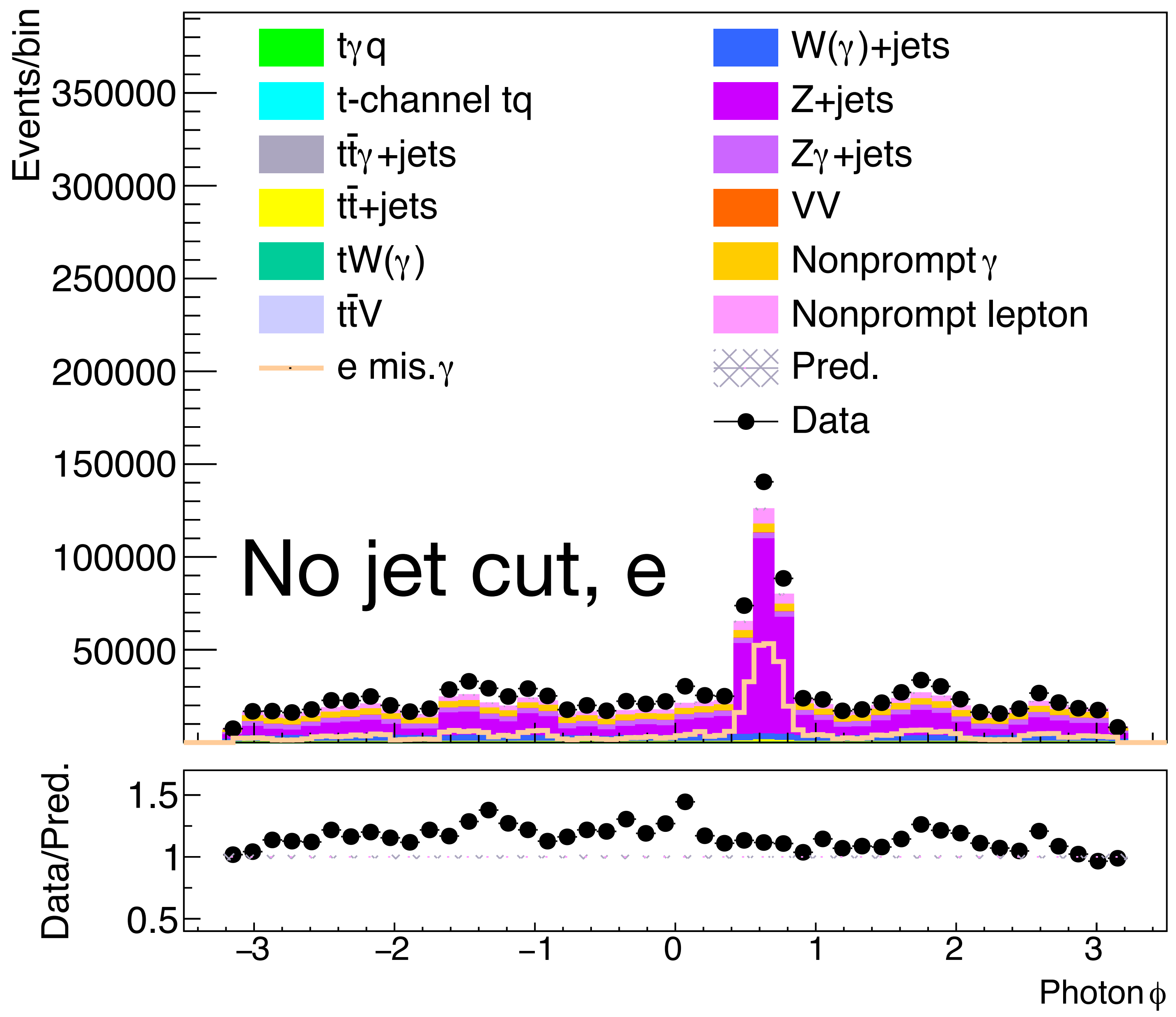
Photon ϕ spike

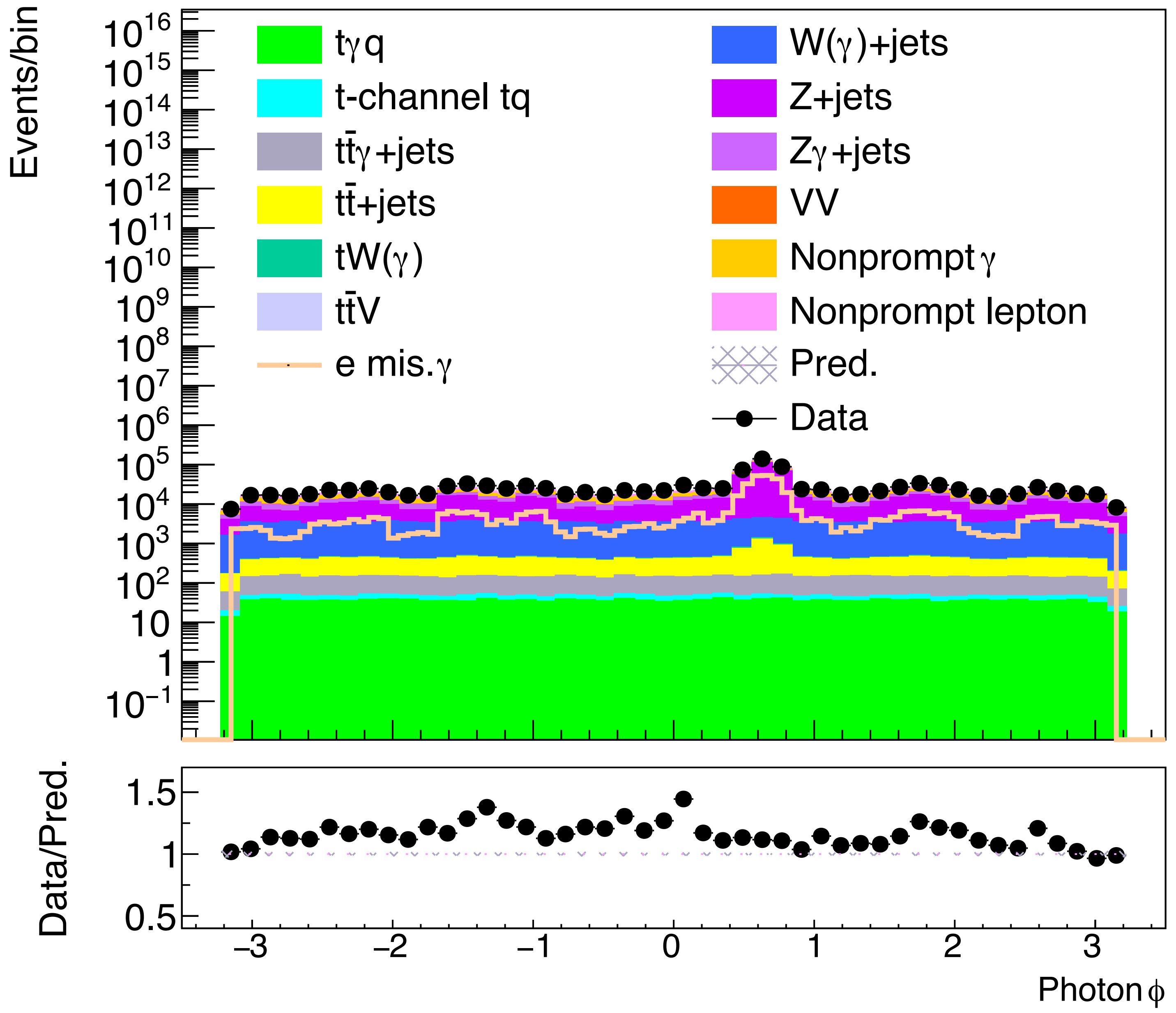








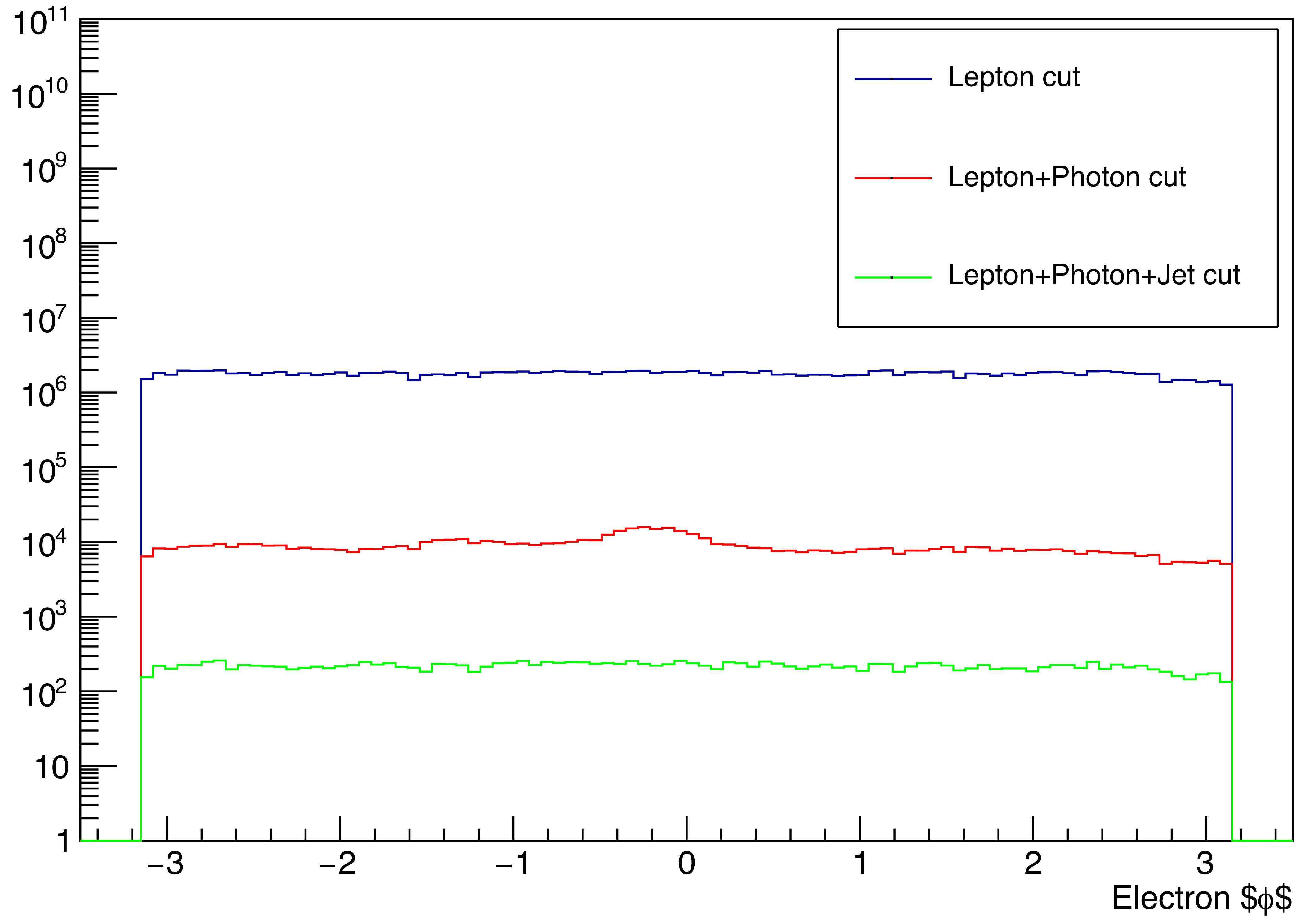


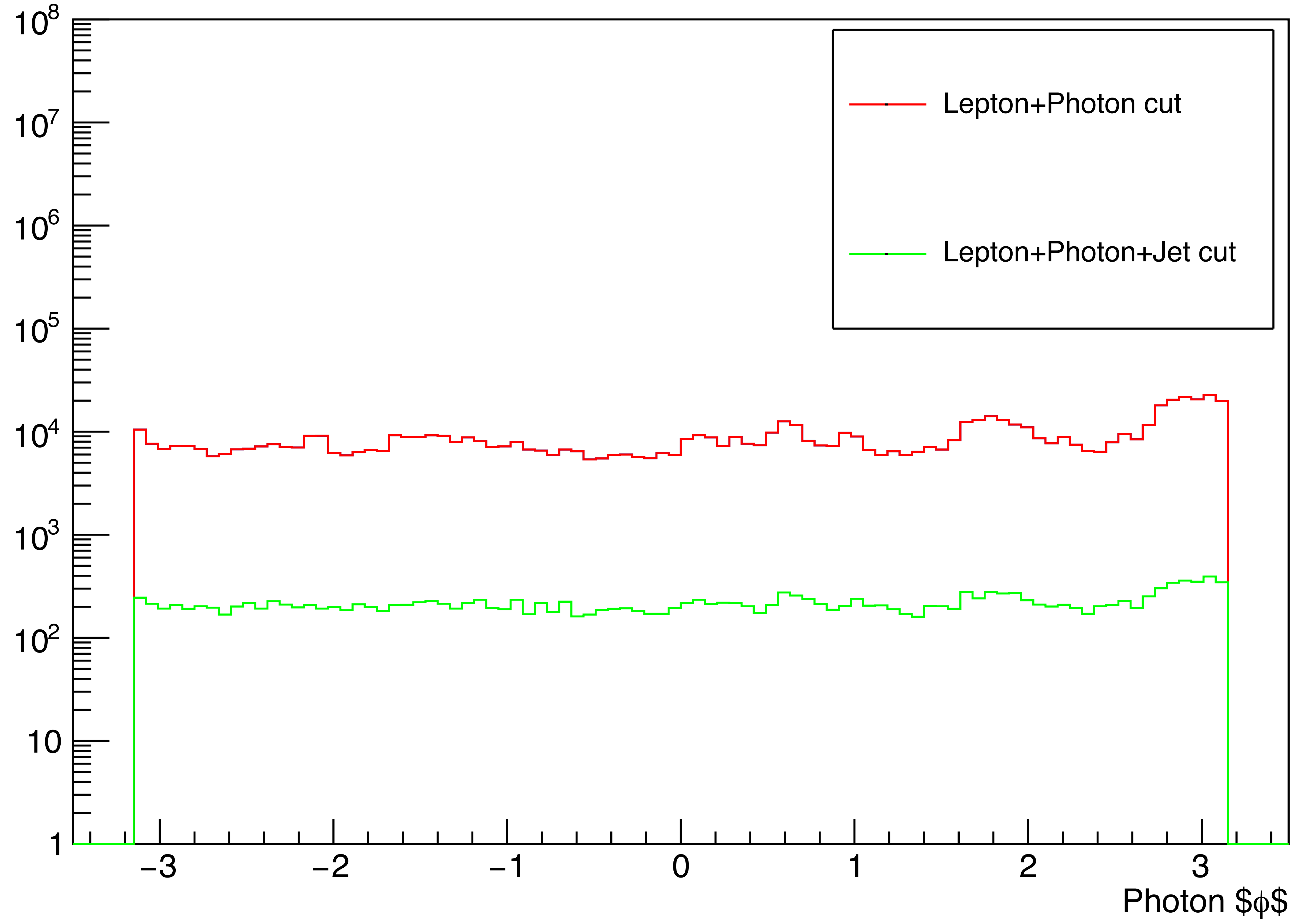


back

Photon ϕ spike

2017 SingleElectron data





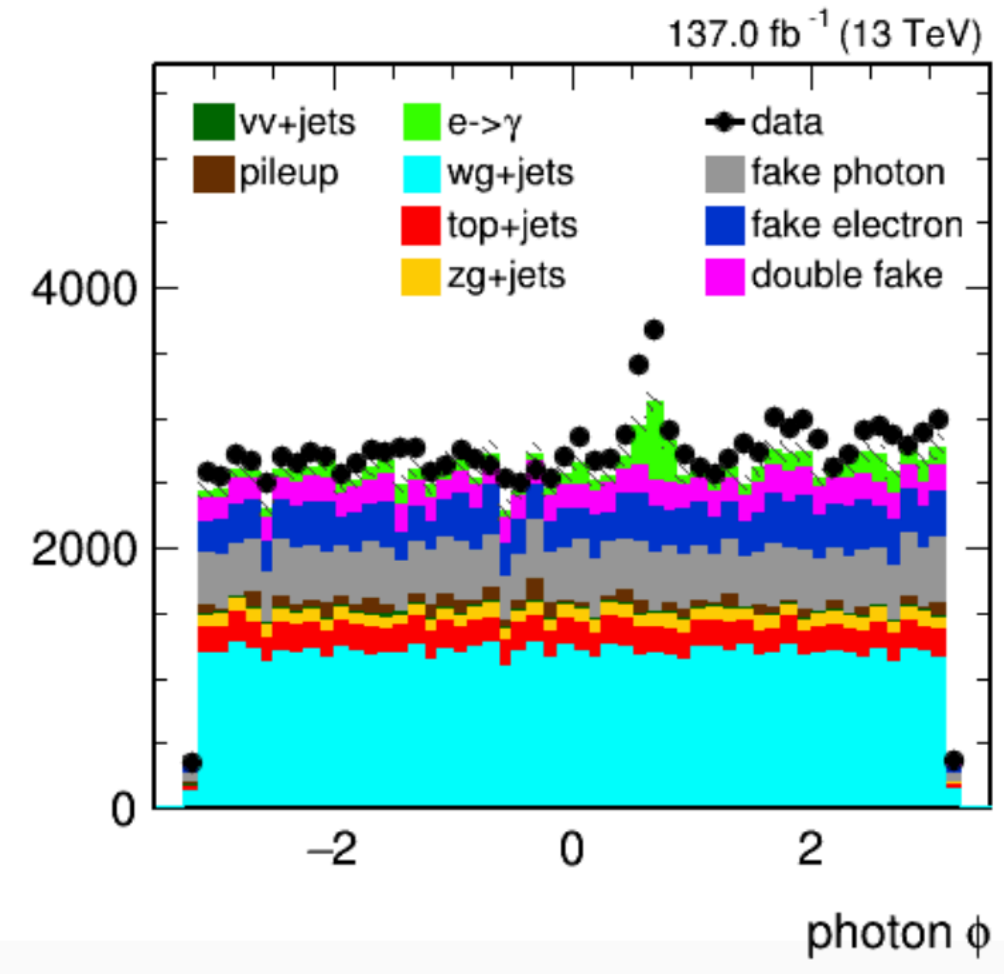
back

Photon ϕ spike

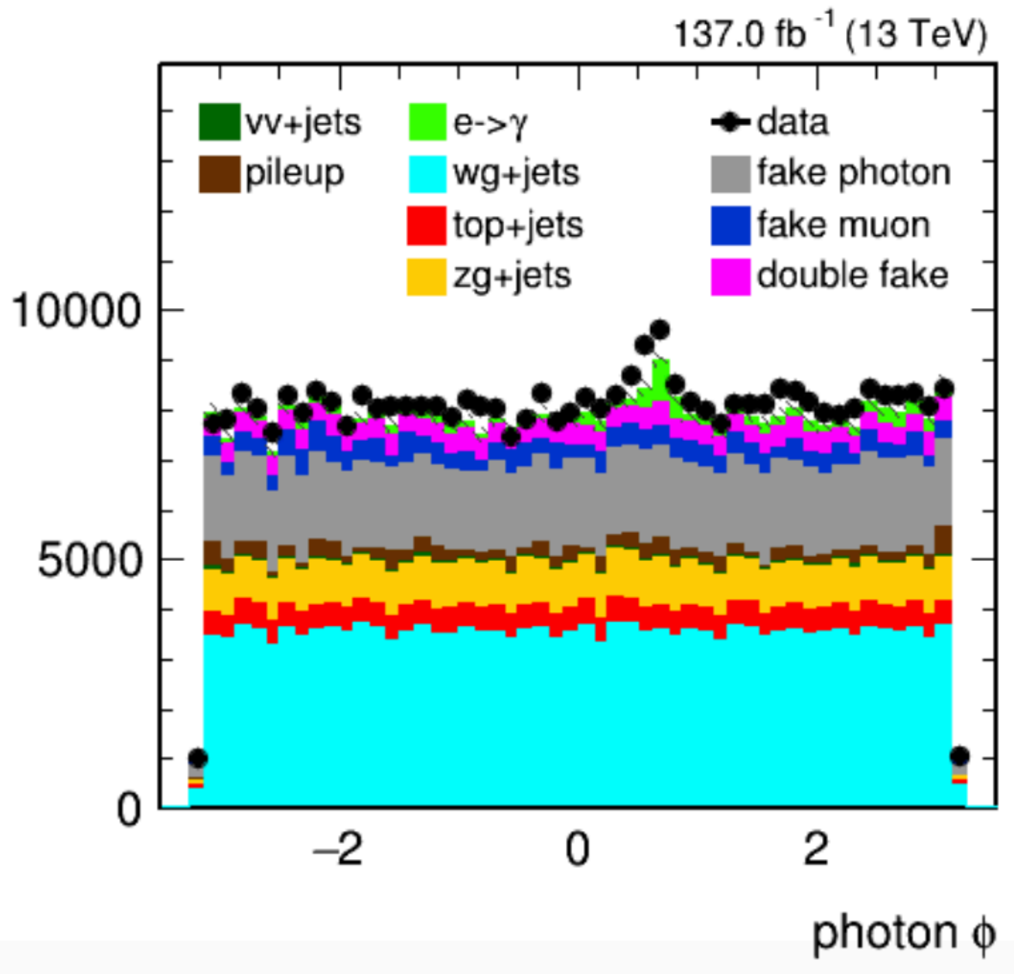
SMP-19-002: Wgamma cross section

- Signal region
  - The photon ϕ region $0.55 < \phi < 0.7$ is removed for the 2018 samples in order to avoid an obvious detector issue spikes

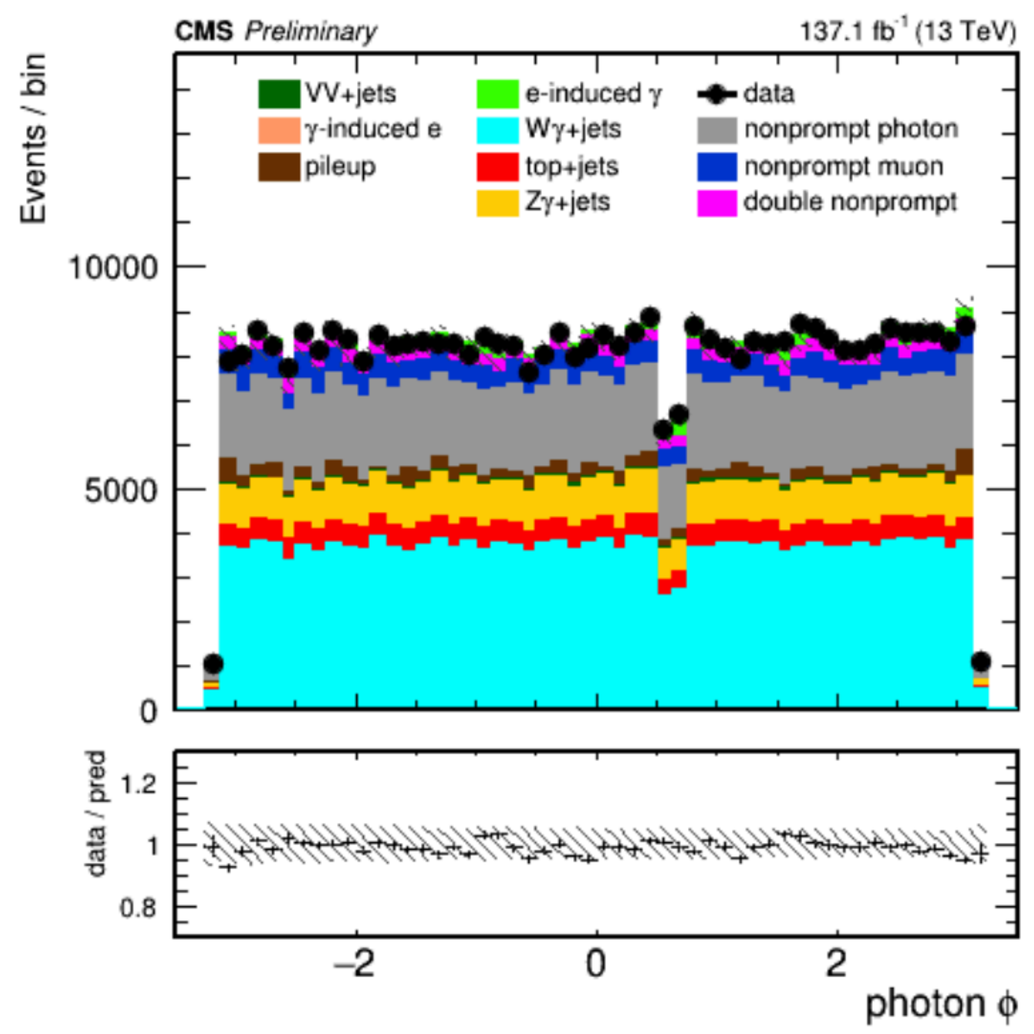
Electron channel, photon ϕ



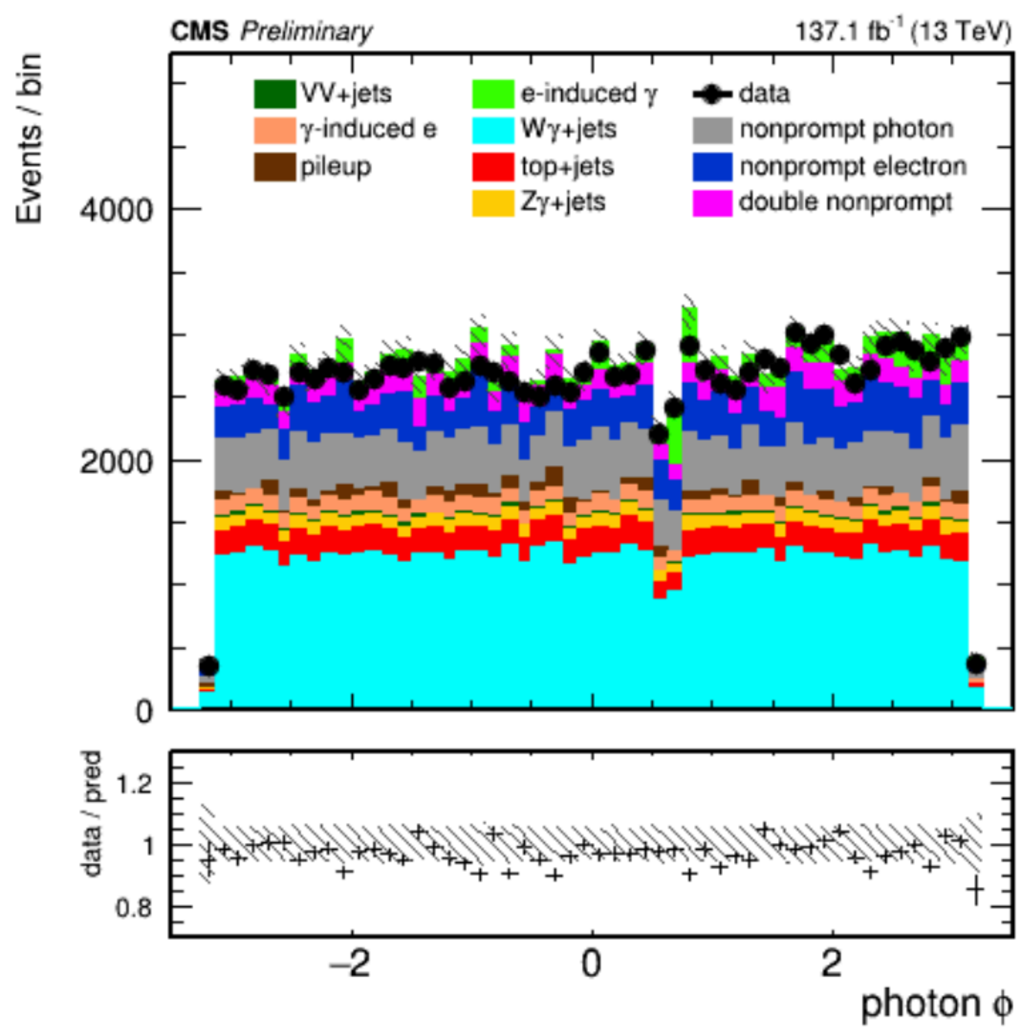
Muon channel, photon ϕ



Muon channel | **Photon ϕ**



Electron channel | **Photon ϕ**

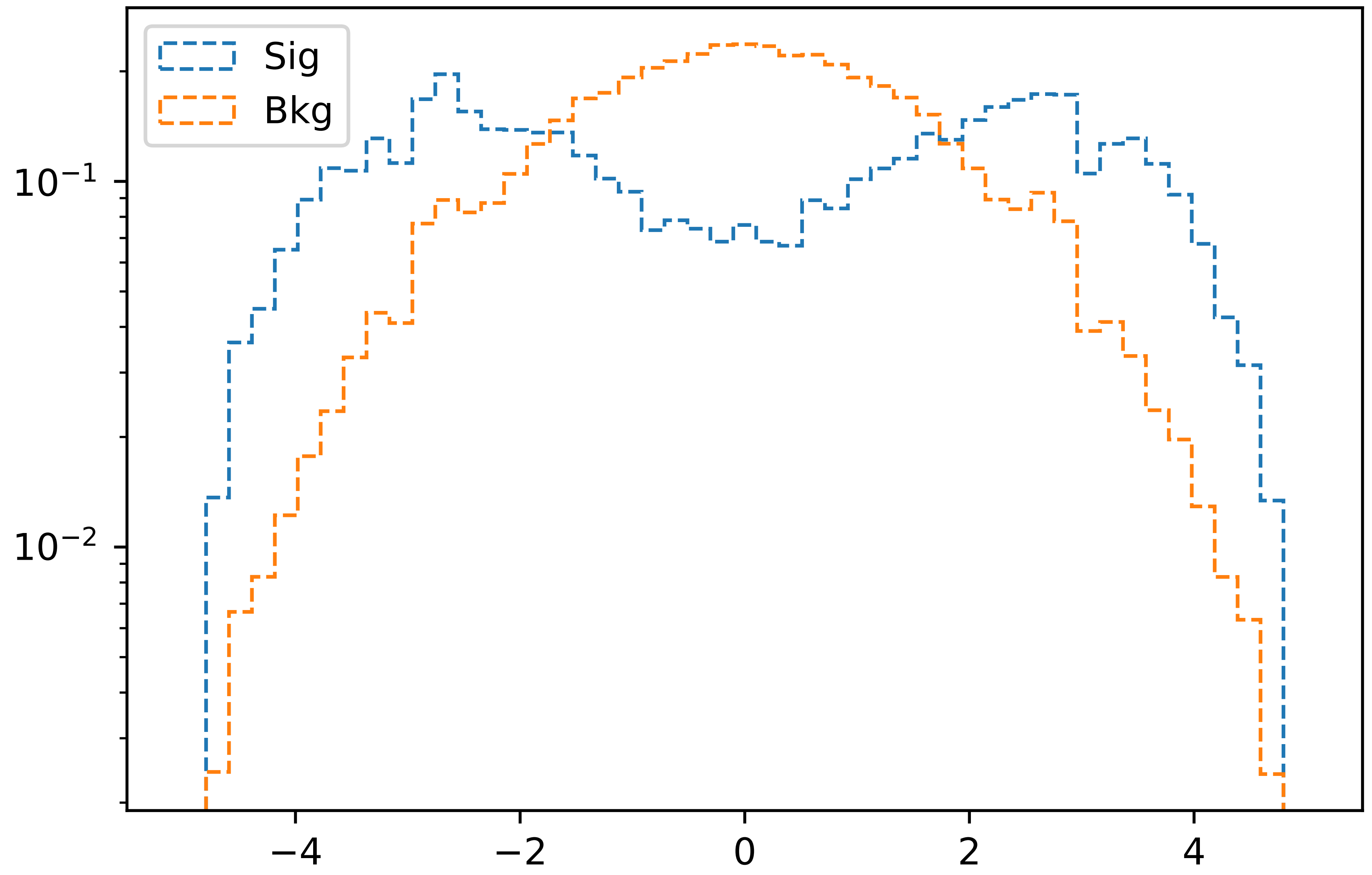


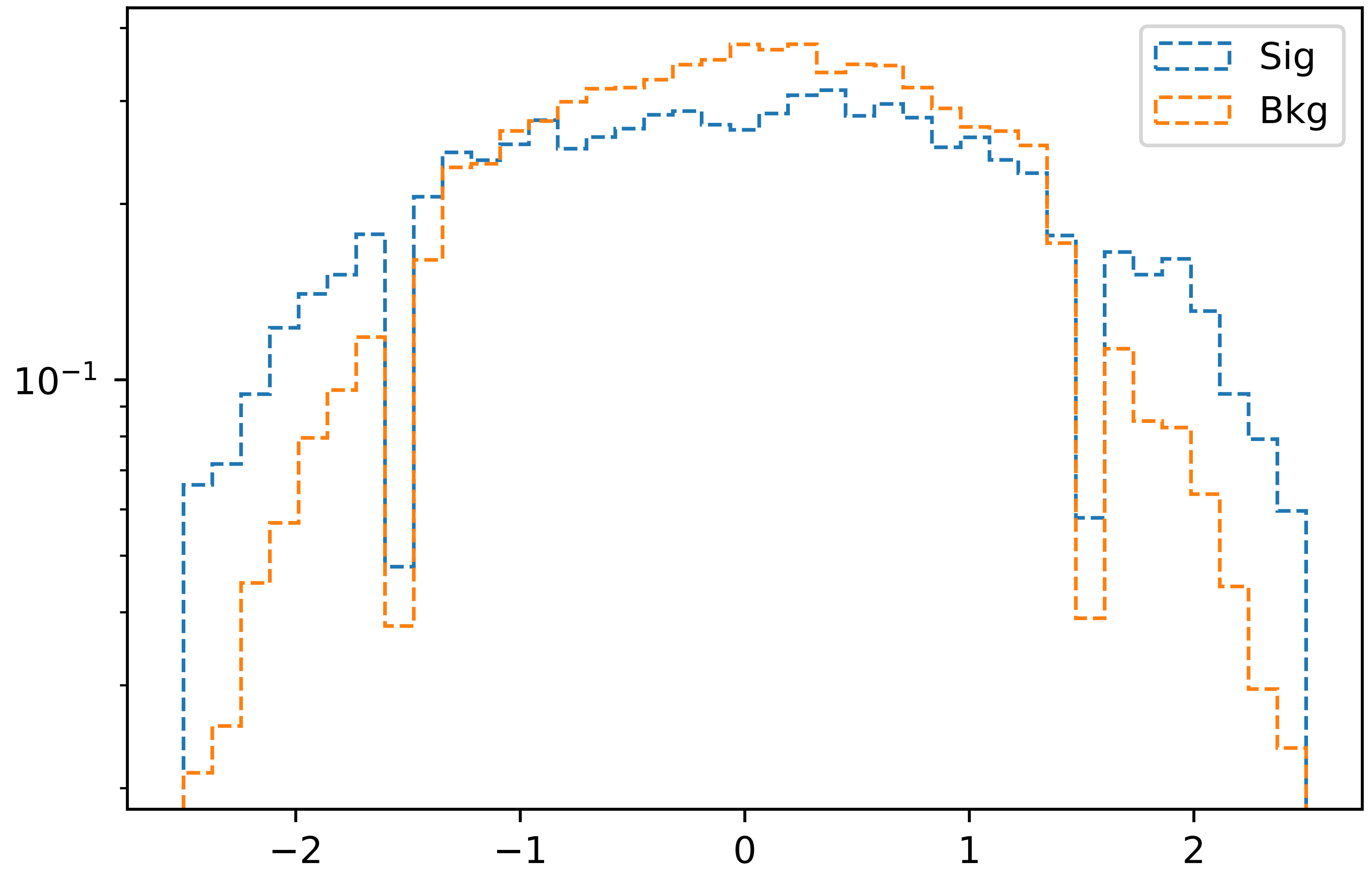
Pre-approval

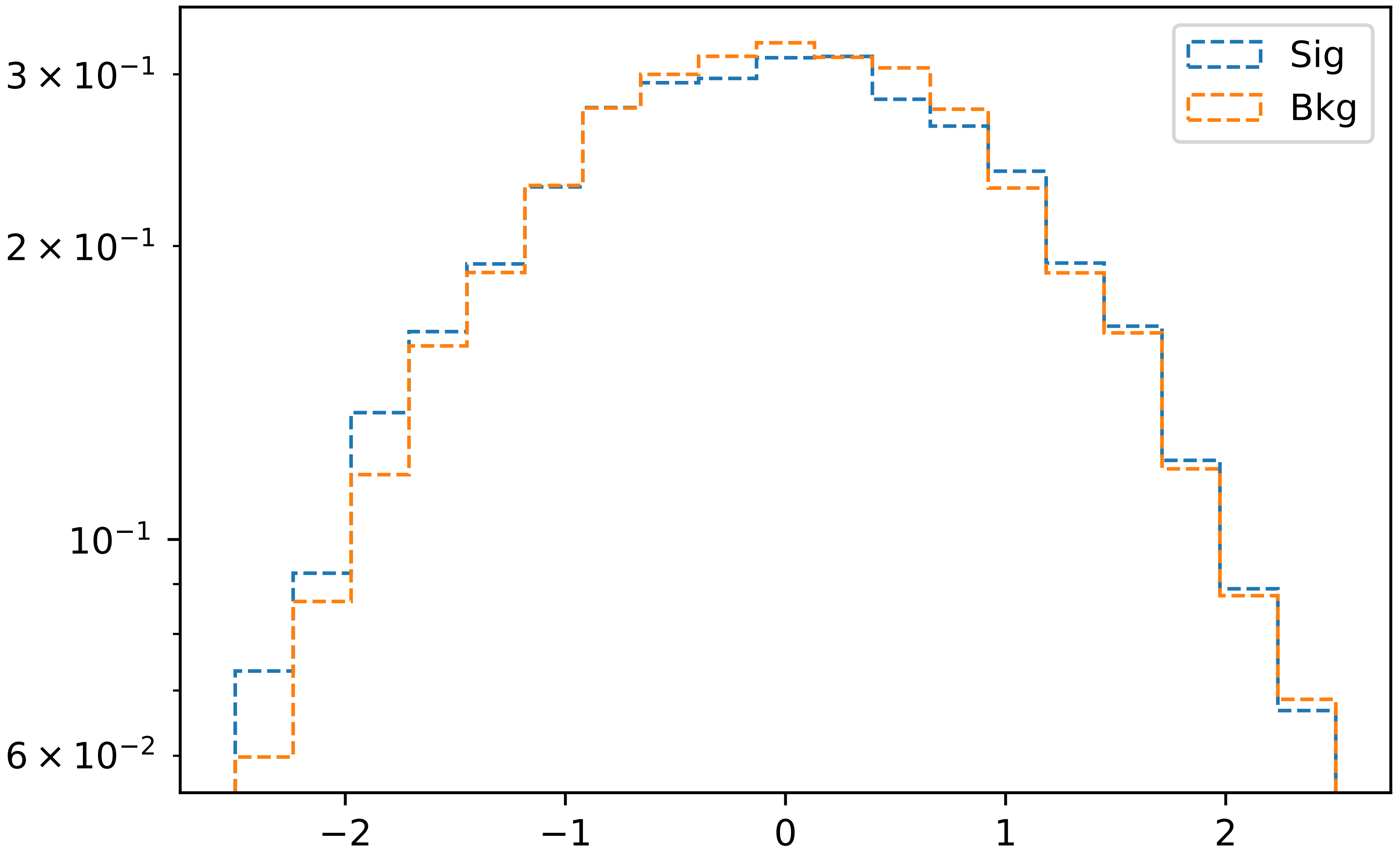
Approval

back

ML training







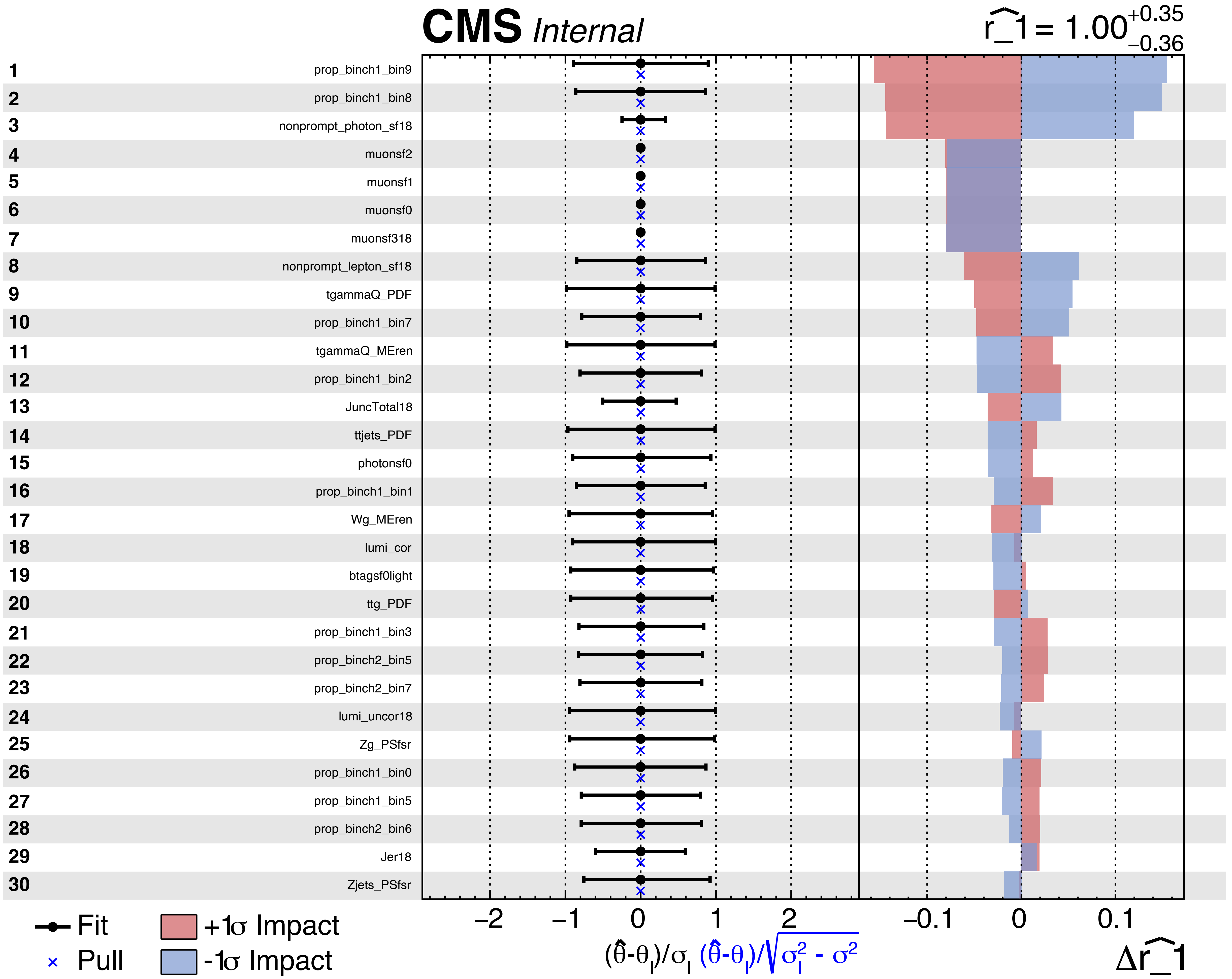
back

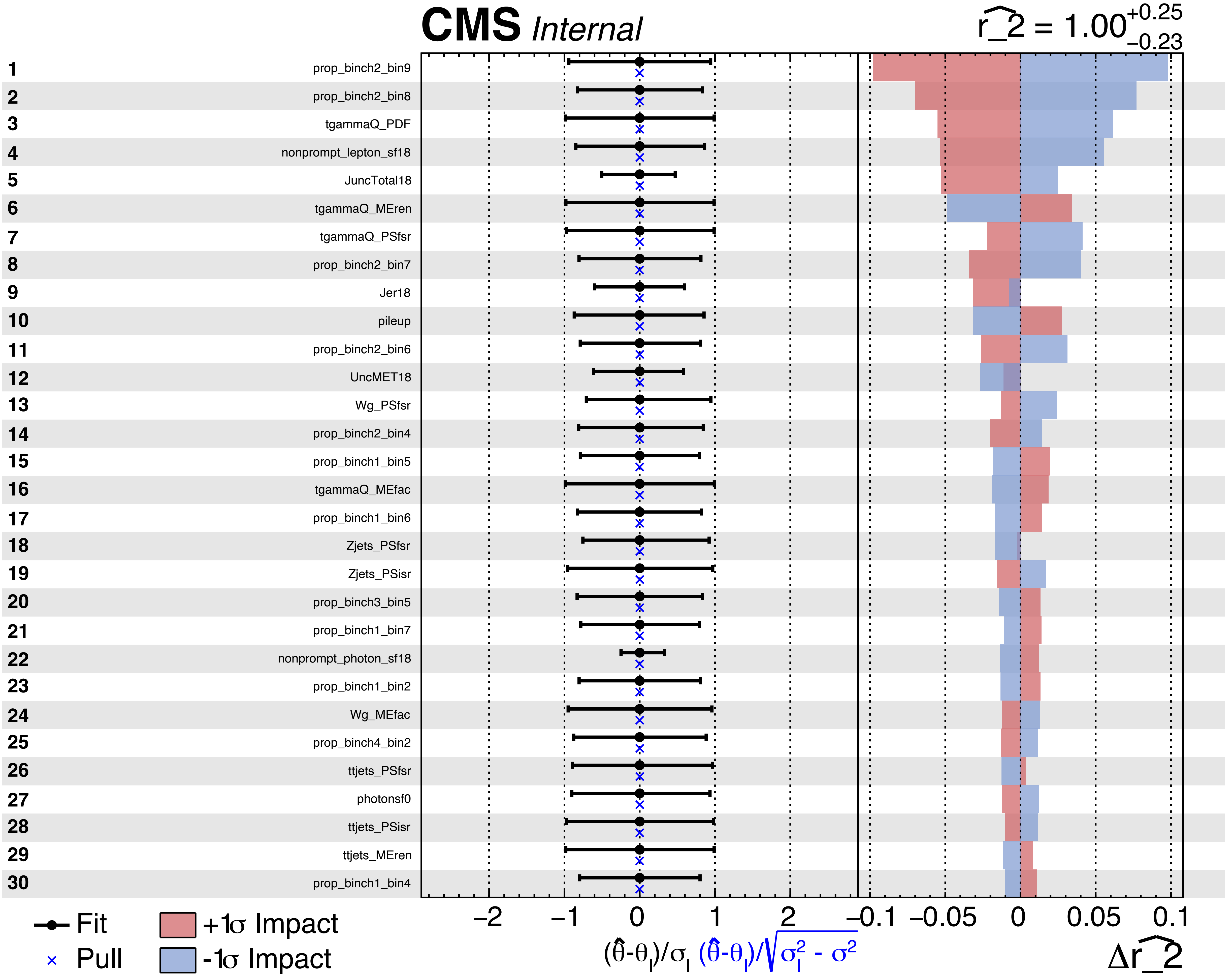
Fiducial region

Selection	gen-lepton	gen-photon	gen-Jet	gen-bJet
p_T/GeV	> 30	> 15	> 30	> 30
$\|\eta\|$	< 2.5	< 2.5	< 4.7	< 2.5
status	1	1	—	—
$\|$pdgID$\|$	13/11	22	—	—
Others	No meson mother	• No meson mother • Isolated • $\Delta R(\ell,\gamma) > 0.1$	• $\Delta R(\ell,j) > 0.4$ • $\Delta R(\ell,\gamma) > 0.1$	• $\|$partonFlavour$\|$ = 5 • $\Delta R(\ell,j) > 0.4$ • $\Delta R(\ell,\gamma) > 0.1$

back

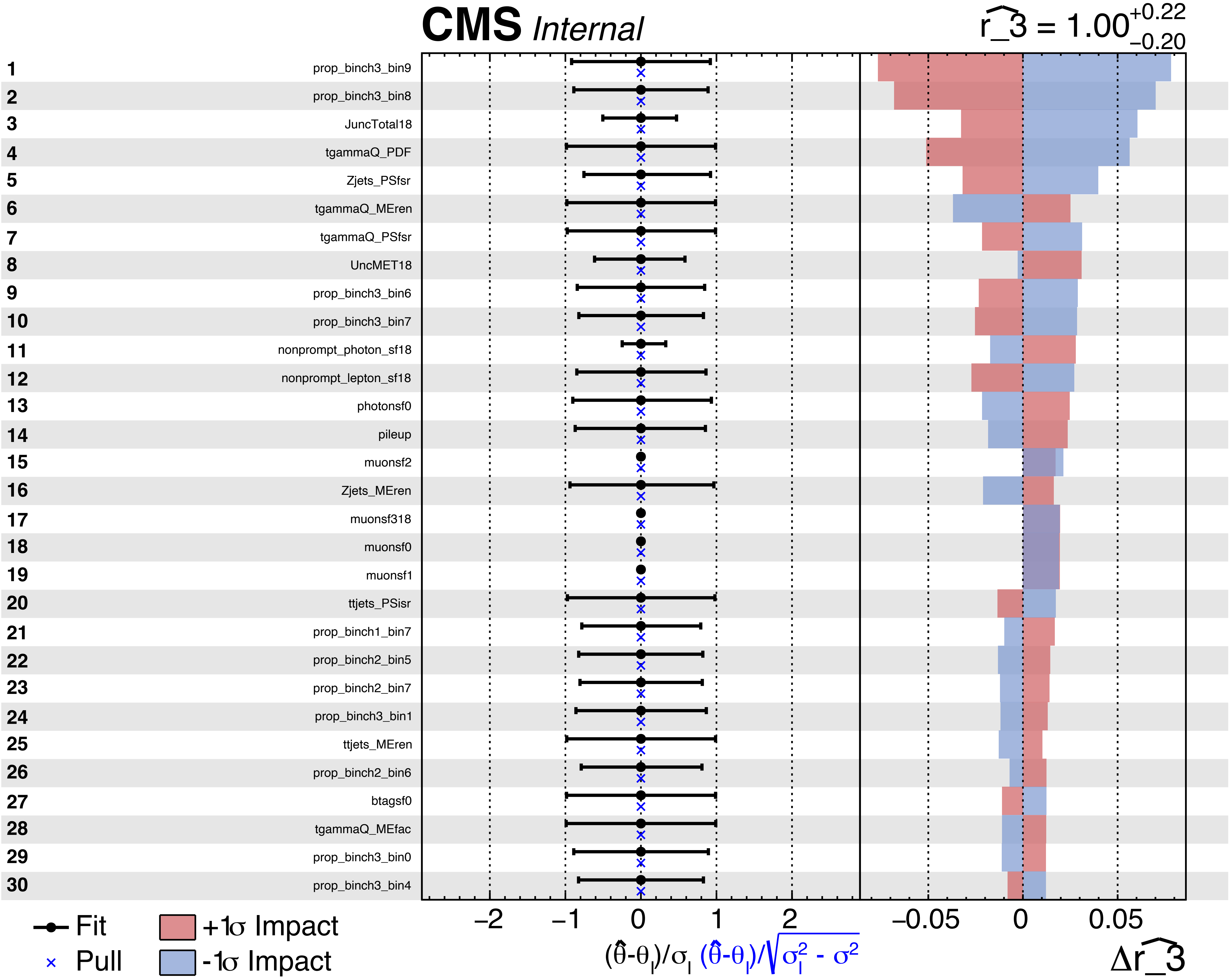
Differential fit of $t\gamma q$ — impact

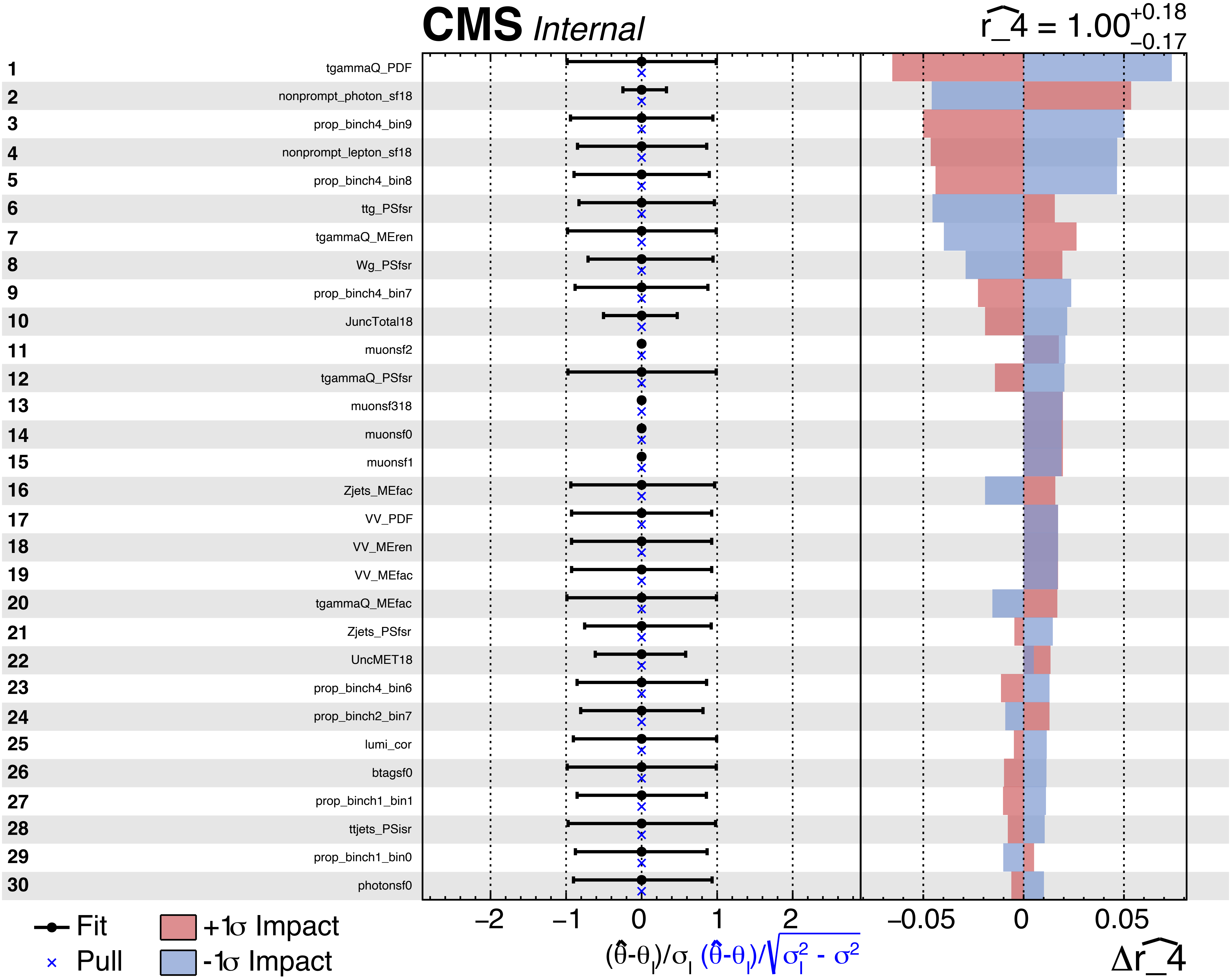




back

Differential fit of $t\gamma q$ — impact





back