Measurement of the electroweak production of Zγ and two jets at 13 TeV and constraints on EFTs

Ying An, Peking University

中国物理学会高能物理分会

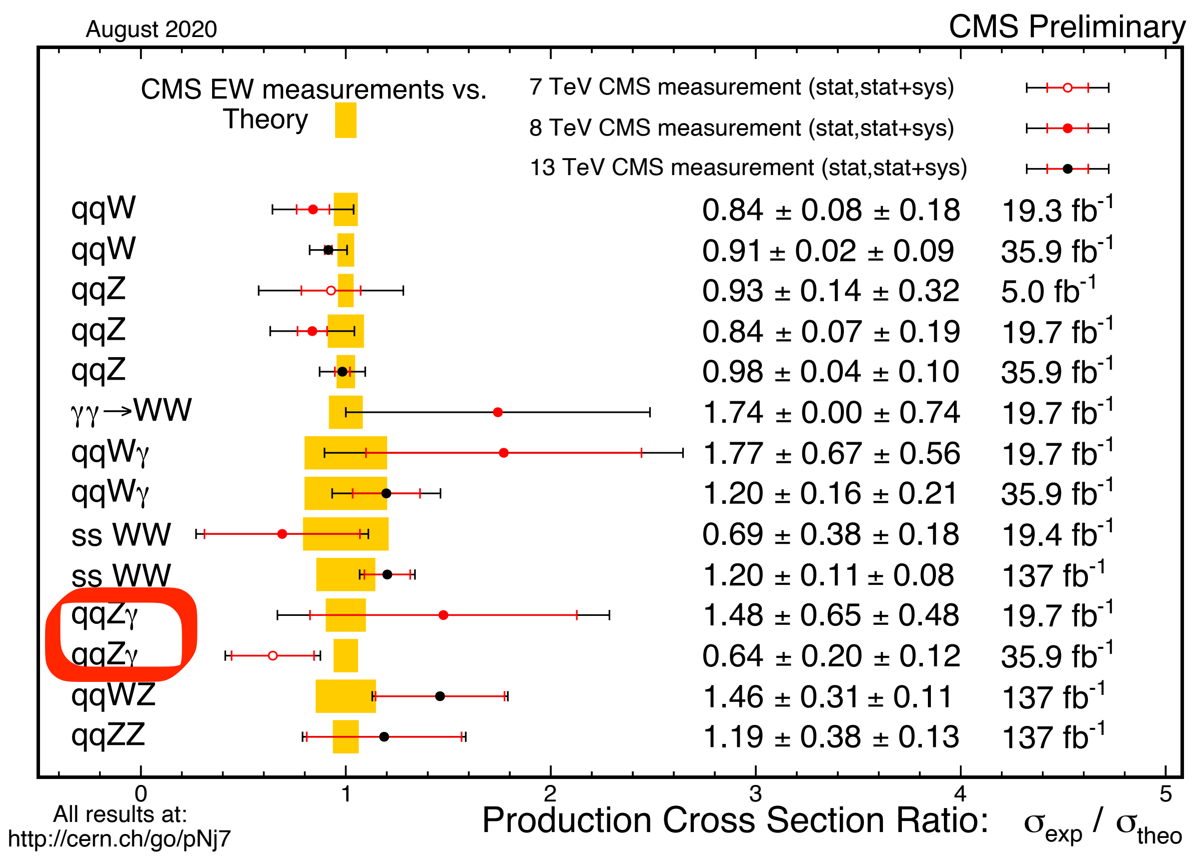
第十一届全国会员代表大会暨学术年会

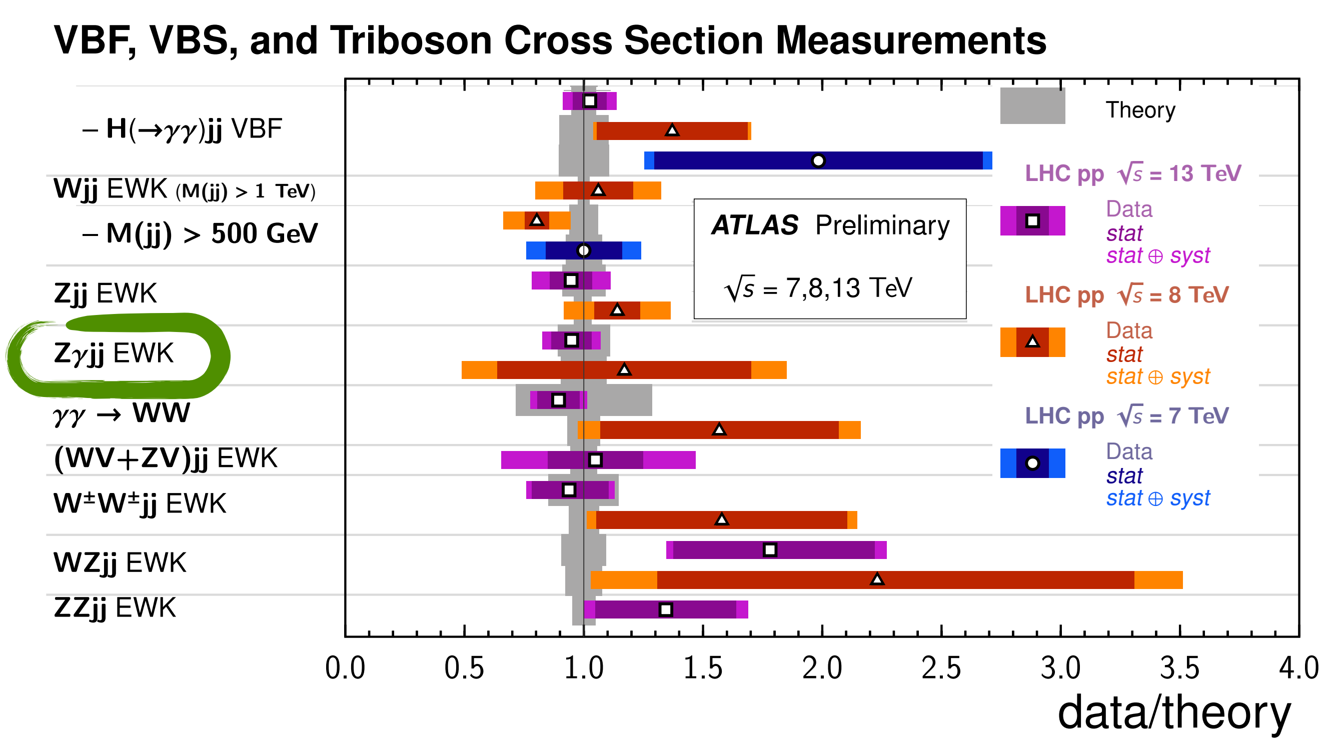
Aug 10 2022

Outline

- Introduction and physics motivation
- Object definitions and samples
- Background estimation and event selection
- Systematic uncertainty
- Results
    - Signal significance
    - Inclusive and differential cross sections
    - Anomalous couplings

Physics motivation





CMS:

13 TeV (35.9 fb^{-1}):

- Observed (expected) significance is 3.9σ (5.2σ)
- Observed (expected) significance is 4.7σ (5.5σ), if combine data in 8 TeV

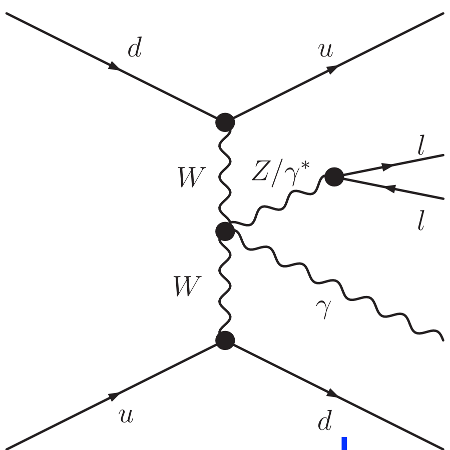
ATLAS:

13 TeV (36 fb^{-1}):

- Observed (expected) significance is 4.1σ (4.1σ)

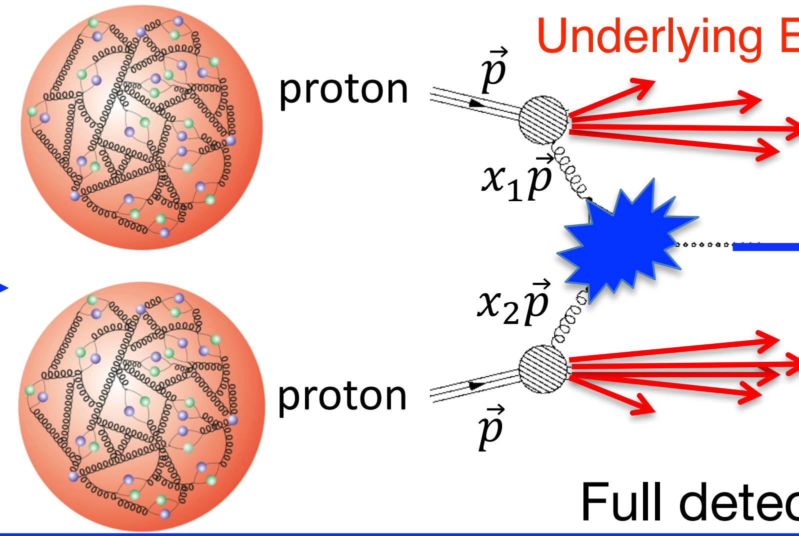
Expect to give the first observation and give more accurate cross section results

Object and Sample

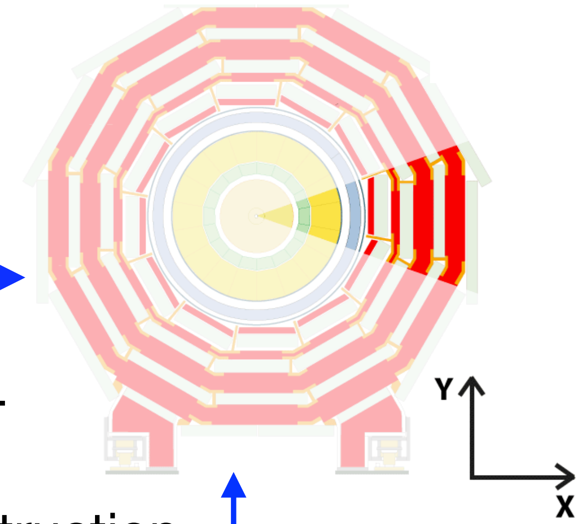


- Parton distribution
- Parton showering/ hadronization
- Underlying Events

MadGraph
Pythia8
Generator-level



Geant4



Full detector reconstruction

Final states	Generator (MadGraph_aMC@NLO)	Lepton identification	Photon identification	Jet identification
eeγjj	process: pp > ℓℓγjj QCD=0 $m_{\ell\ell}$ > 50 GeV, m_{jj} > 120 GeV	cut-based-medium (≈80%)	cut-based-medium (≈80%)	anti-k_T AK4CHS 98%-99%
μμγjj		cut-based-tight (≥95%)		

Vector boson scattering (VBS) signature:
large dijet mass and large η separation between the jets

Main results:

- ✓ Signal significance
- ✓ Inclusive and differential cross sections
- ✓ Limits on anomalous couplings

Background estimation

- Diboson processes with ≥ 2ℓ → decreased by the loose lepton ID
- TTγ and tW where the top decays to b and W
- QCD Zγ with total same background →irreducible background
- Interference sample as a part of QCD Zγ
- Nonprompt photon from hadronic jet

estimated by normalizing MC to data (first four items)

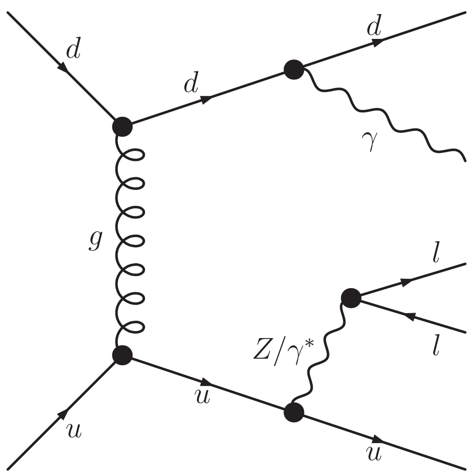
Nonprompt photon from hadronic jet → estimated with data-driven method

QCD Zγ estimation

- Diboson processes with ≥ 2ℓ → decreased by the loose lepton ID
- TTγ and tW where the top decays to b and W
- QCD Zγ with total same background →irreducible background
- Interference sample as a part of QCD Zγ
- Nonprompt photon from hadronic jet

estimated by normalizing MC to data

estimated with data-driven method



QCD background

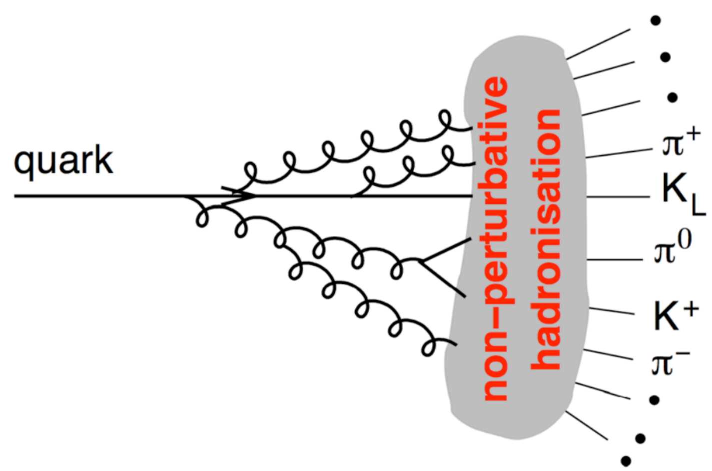
Normalization is constrained by the simultaneous fit of the signal region and low m_{jj} control region

Nonprompt γ estimation

- Diboson processes with ≥ 2ℓ → decreased by the loose lepton ID
- TTγ and tW where the top decays to b and W
- QCD Zγ with total same background →irreducible background
- Interference sample as a part of QCD Zγ

estimated by normalizing MC to data

- Nonprompt photon from hadronic jet → estimated with data-driven method



Nonprompt γ background

From Data

Fake photon enriched sample by **inverting the isolation variables** in the photon ID in data

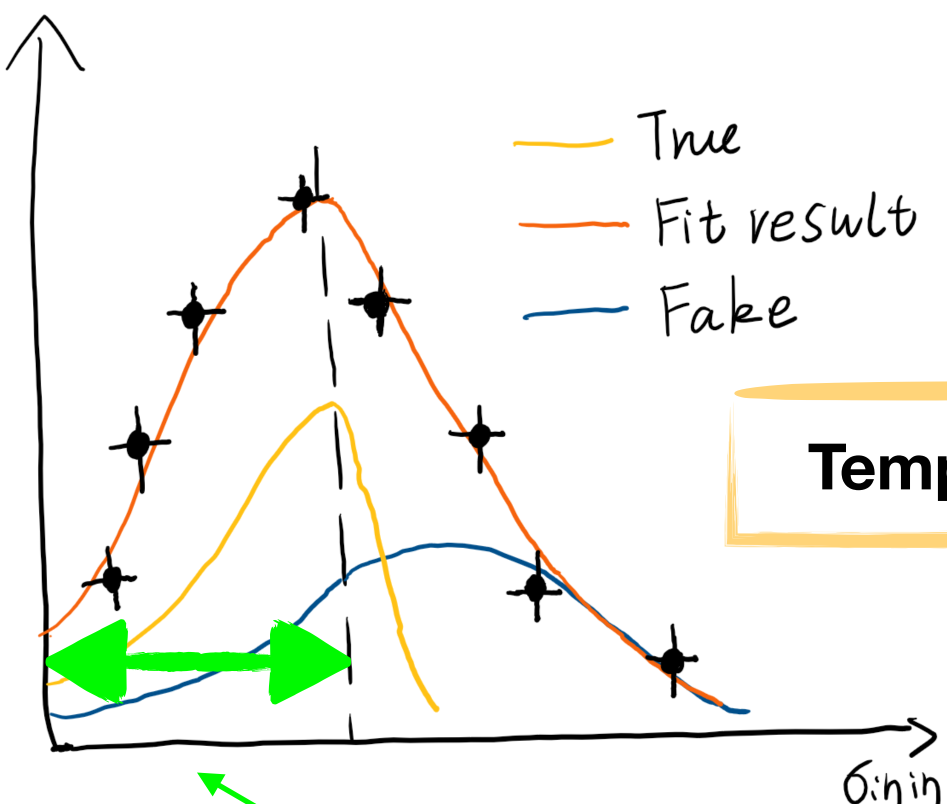
$$n^{predicted}_{fake-in-SR} = n_{tot} \times \epsilon_{\text{fake-fraction}} = N^{unweighted}_{fake-in-CR} \times weights$$

Nonprompt γ estimation

- Diboson processes that produce ≥ 2 leptons
- TTγ and tW where the top decays to b and W
- QCD Zγ with total same background →irreducible background
- Interference sample as a part of QCD Zγ

estimated by normalizing MC to data

- Nonprompt photon from hadronic jet → estimated with data-driven method

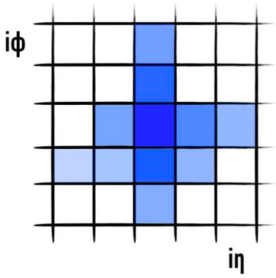


Template fit

Based on the Z+jets events, two good leptons from Z, $70 < m_{\ell\ell} < 110$ GeV

Data	Remove $\sigma_{i\eta i\eta}$ cut
True Template	Remove $\sigma_{i\eta i\eta}$ cut $\Delta R(\gamma^{reco},\gamma^{gen}) < 0.3$ Get shape from simulation
Fake template	Remove $\sigma_{i\eta i\eta}$ cut **Invert the charged isolation variable** Get shape from data

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_i^{5\times5} w_i(\eta_i - \bar{\eta}_{5\times5})^2}{\sum_i^{5\times5} w_i}}$$



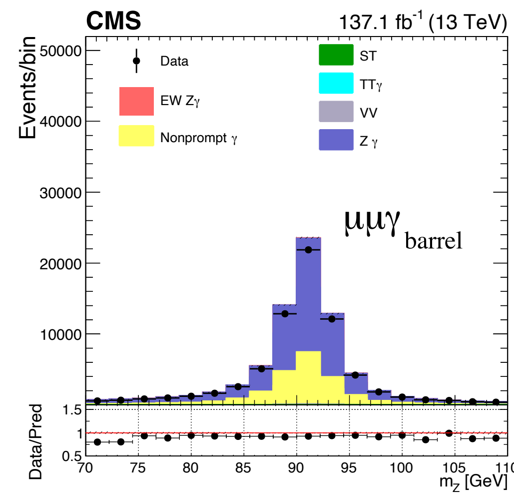
$w_i \neq 0$, if $E_i > 0.9\%$ of E_{5x5}

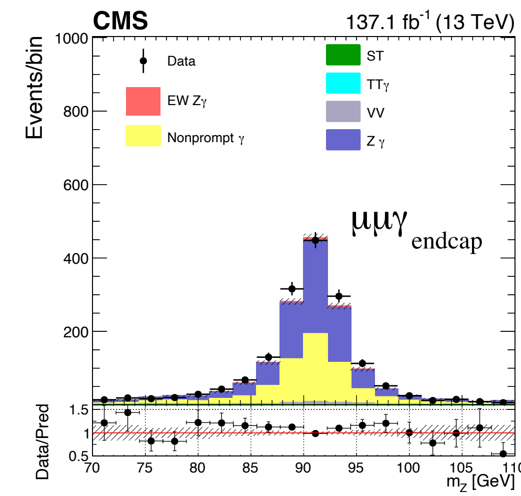
$$n^{predicted}_{fake-in-SR} = n_{tot} \times \epsilon_{fake-fraction} = N^{unweighted}_{fake-in-CR} \times weights$$

Event selection

Basic event selection

- $p_T^{\mu/e} > 20/25$ GeV (double lepton HLT)
- $70 < m_{\ell\ell} < 110$ GeV
- $p_T^{\gamma} > 20$ GeV in barrel/endcap
- Two jets with $p_T^{j} > 30$ GeV, $|\eta| < 4.7$
- $m_{\ell\ell\gamma} > 100$ GeV (suppress FSR)





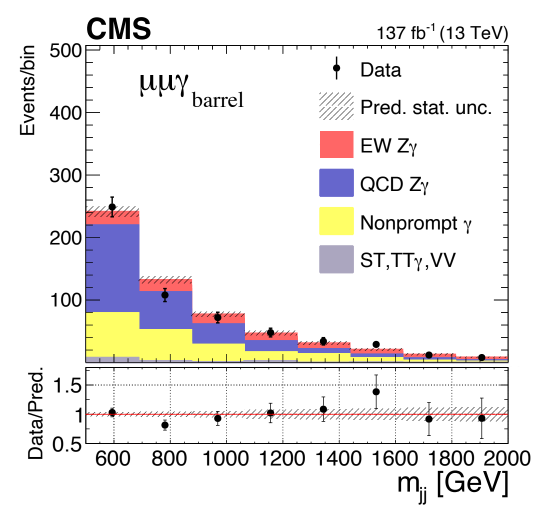
Validate the background estimation

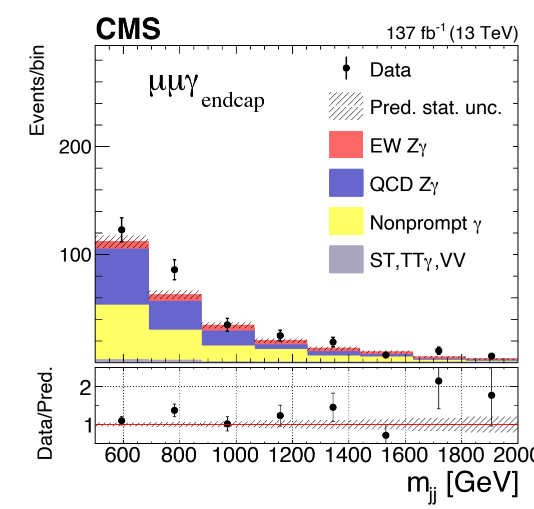
Low m_{jj} control region

- $150 < m_{jj} < 500$ GeV

VBS signal region

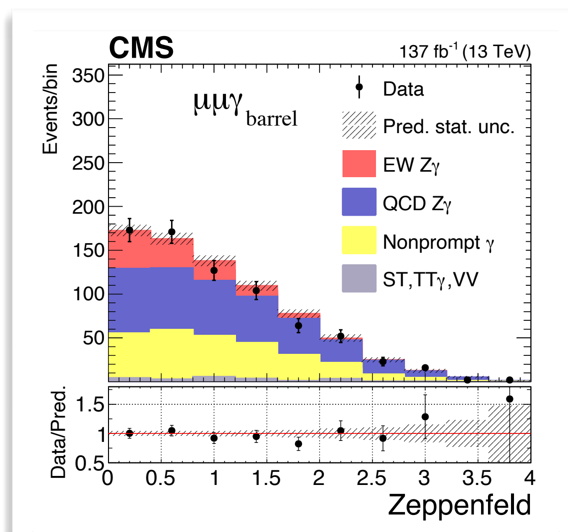
- $m_{jj} > 500$ GeV
- $\Delta\eta_{jj} > 2.5$

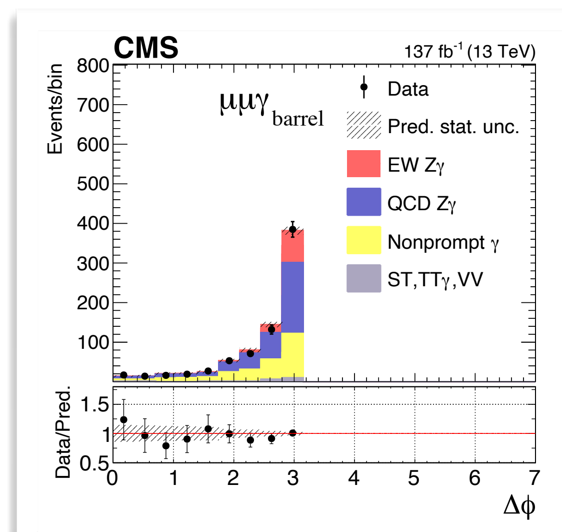




EW signal extraction

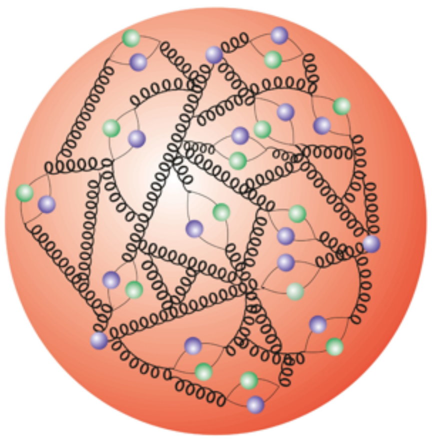
- Zeppenfeld $= |\eta_{Z\gamma} - (\eta_{j1} + \eta_{j2})/2| < 2.4$
- $\Delta\phi = |\phi_{Z\gamma} - (\phi_{j1} + \phi_{j2})| > 1.9$





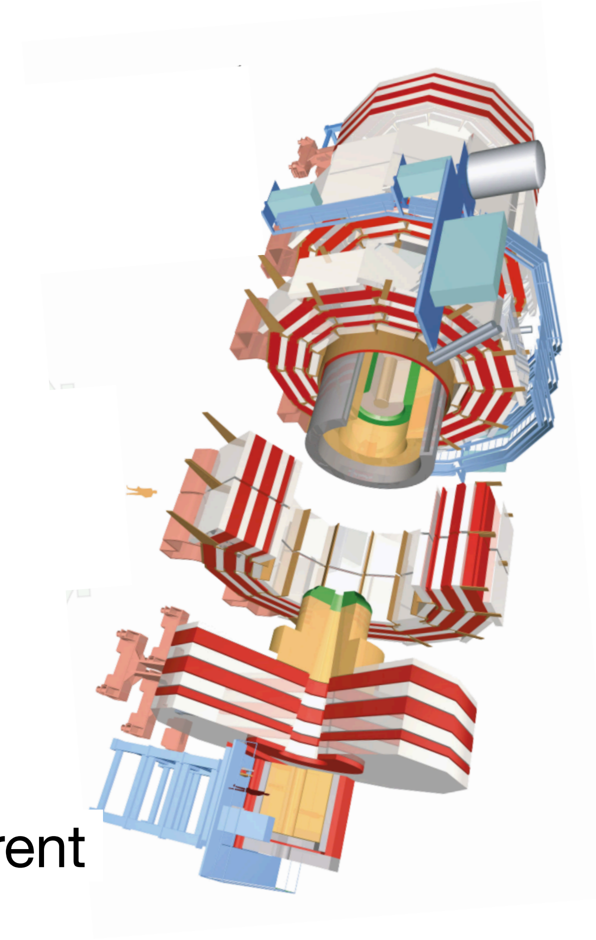
Systematic uncertainty

Theoretical uncertainties:



- **Factorization and renormalization scale uncertainty**
  - Calculated bin-by-bin, correlated between bins, channels, and years
- PDF uncertainty
  - Standard deviation of the around 100 NNPDF PDF set variations
  - Calculated bin-by-bin, correlated between bins, channels, and years

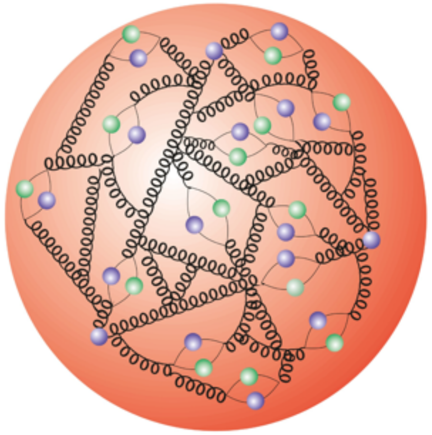
Experimental uncertainties:



- JER and JES uncertainty
  - Largest deviation of up/down to the central
  - Calculated bin-by-bin, correlated between bins and channels, uncorrelated in different years
- Efficiencies, pileup, L1prefiring and etc
- Fake photon uncertainty
  - Three components (Closure test + Sideband choice + True template choice)
  - Calculated bin-by-bin, correlated between bins, uncorrelated in different channels and years

Systematic uncertainty

$$\sigma_{tot}^{pp\to X}(\mu_F,\mu_R) = \sum_{i,j}\int dx_1 dx_2 f_{i,p}(x_1,\mu_F) f_{j,p}(x_2,\mu_F)\hat{\sigma}_{ij}^{ij\to X}(x_1 x_2 S,\mu_F,\mu_R)$$



- Constrained QCD $Z\gamma$ normalization by data in a low m_{jj} control region.

SR: m_{jj} $[500, 800, 1200, \infty)$
CR: m_{jj} $[150, 300, 400, 500]$

Very different kinematic distribution

- From the combinations of μ_F and μ_R, if μ_0 is the nominal scale, we have:
  - ☑ Nuisance parameter 1: μ_F only $(2\mu_0,\mu_0)$ and $(0.5\mu_0,\mu_0)$
  - ☑ Nuisance parameter 2: μ_R only $(\mu_0,2\mu_0)$ and $(\mu_0,0.5\mu_0)$
  - ☑ Nuisance parameter 3: $\mu_R + \mu_F$ fully correlated $(2\mu_0,2\mu_0)$ and $(0.5\mu_0,0.5\mu_0)$

~~$(2\mu_0,0.5\mu_0)$, $(0.5\mu_0,2\mu_0)$~~

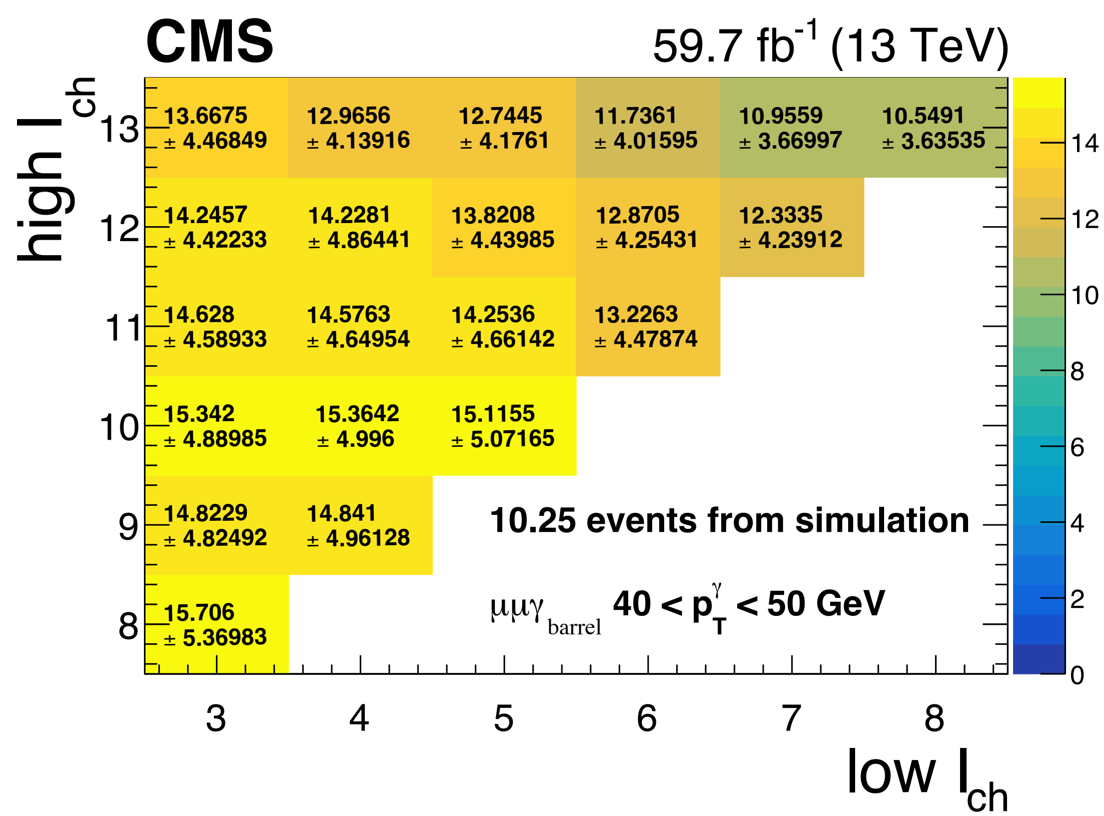
Make the theoretical uncertainty more conservative and reasonable

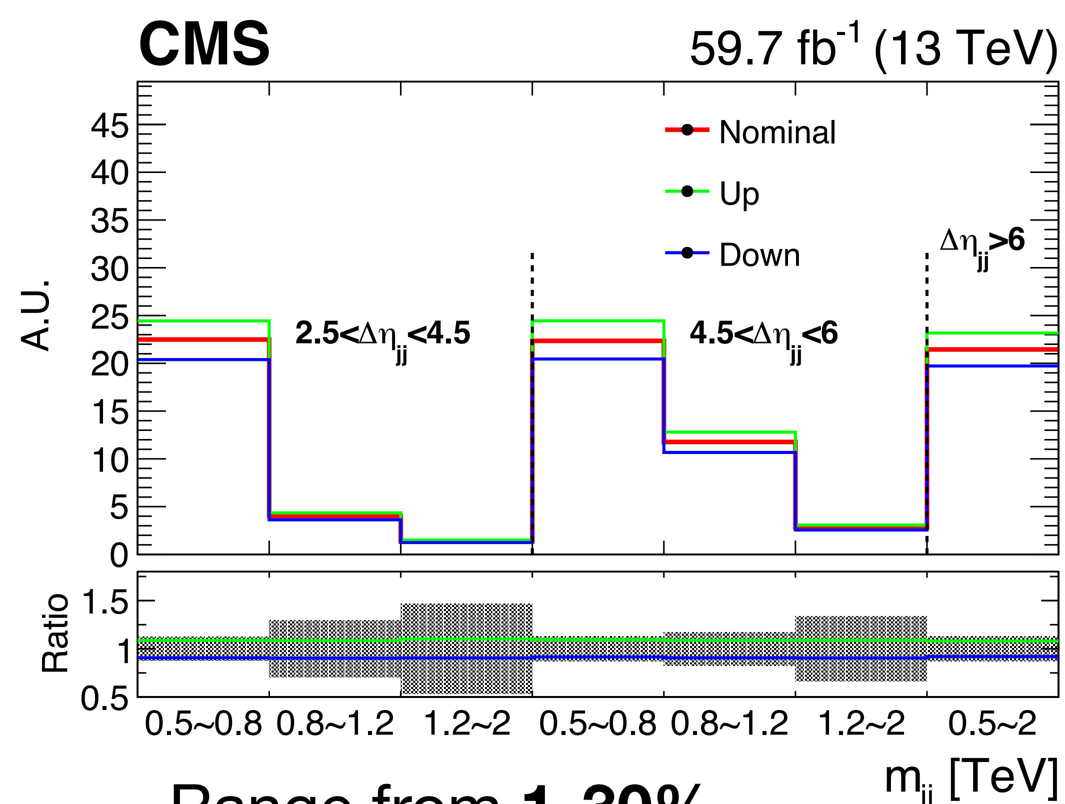
Nonprompt γ uncertainty

Three sources uncertainty contribute to the nonprompt γ fraction added in quadrature

1. Different MC sample (EWK, QCD) used for true template
2. Charge isolation sideband choice
3. Difference of fit and MC truth(closure test)

Main part





Range from **1-30%**

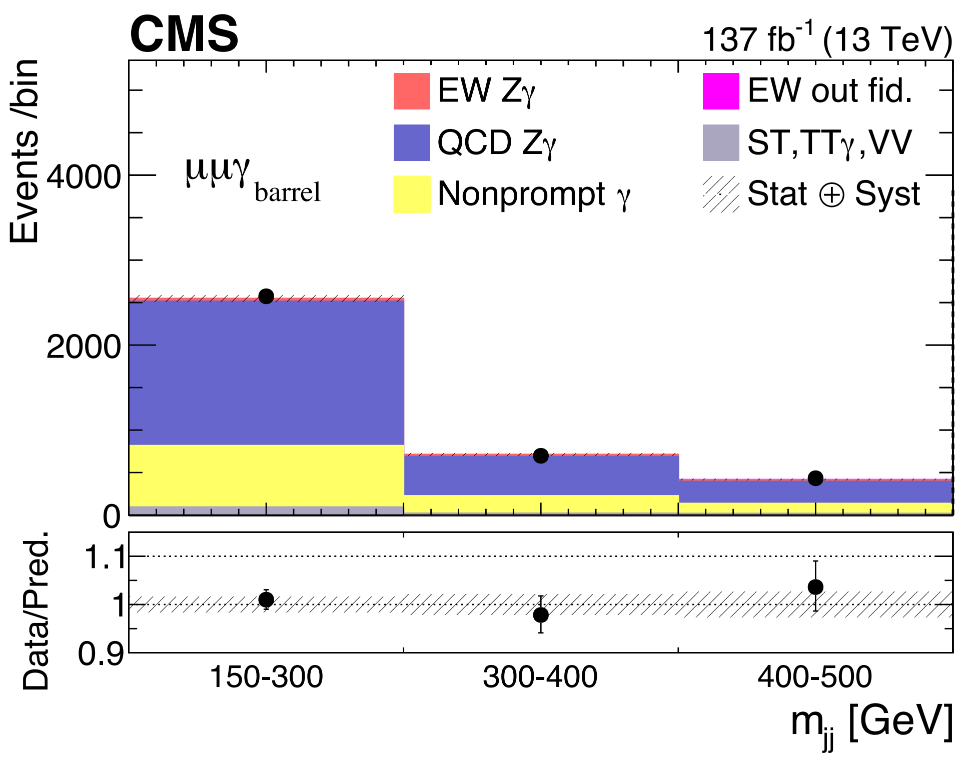
Control **the nonprompt uncertainty from data-driven estimation** ***well***

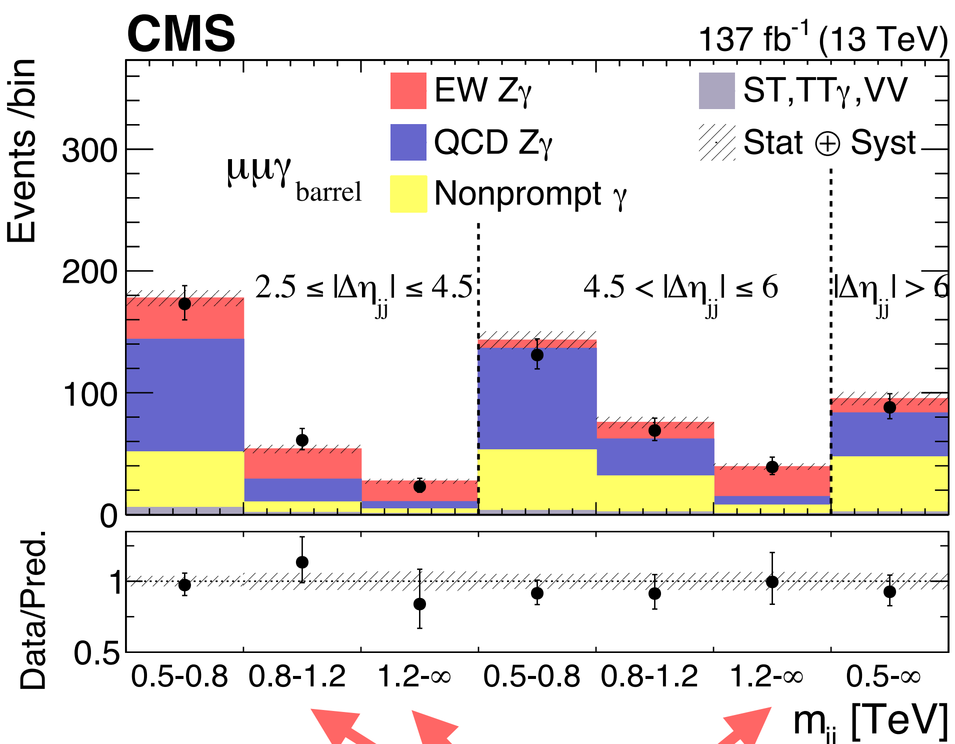
Signal significance

- Perform the simultaneous fit in the signal region and the low m_{jj} control region
- Using the 2D variable (m_{jj} and $\Delta\eta_{jj}$) in the signal region
- 1D variable m_{jj} in the control region

$$\mathscr{L}(\vec{\mu};\vec{\theta}) = \prod_j \textbf{Poisson}(n_j \,|\, \mu \cdot s_j(\vec{\theta}) + b_j(\vec{\theta})) \cdot p(\tilde{\vec{\theta}} \,|\, \vec{\theta})$$ → systematic error

SMP-21-016, PhysRevD.104.072001



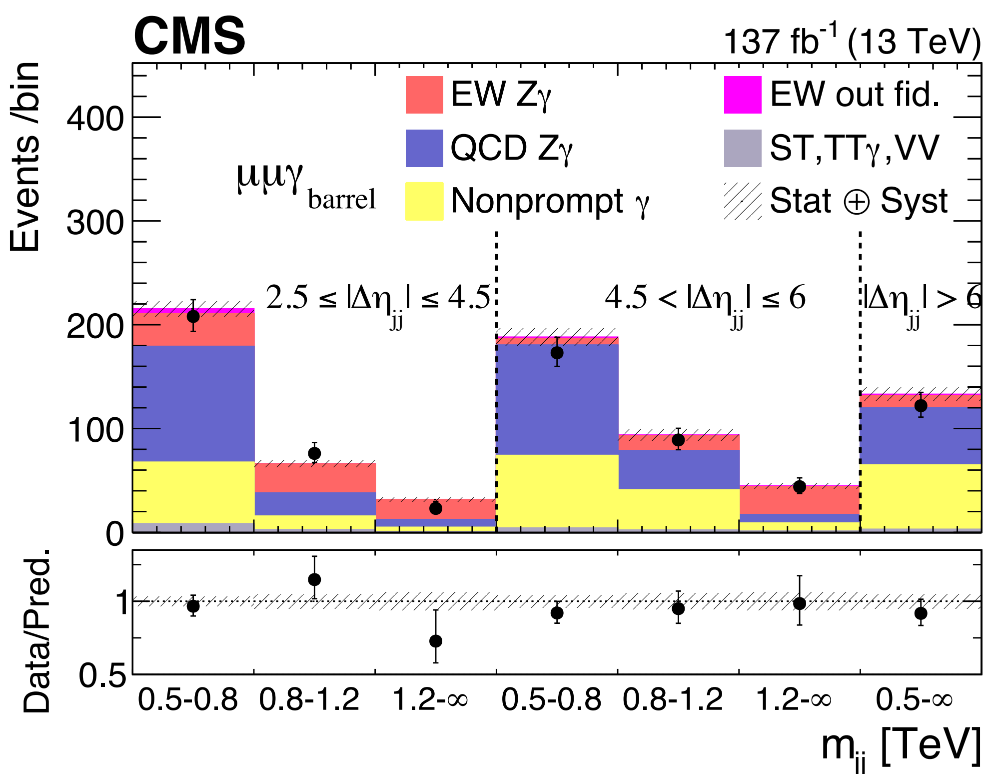


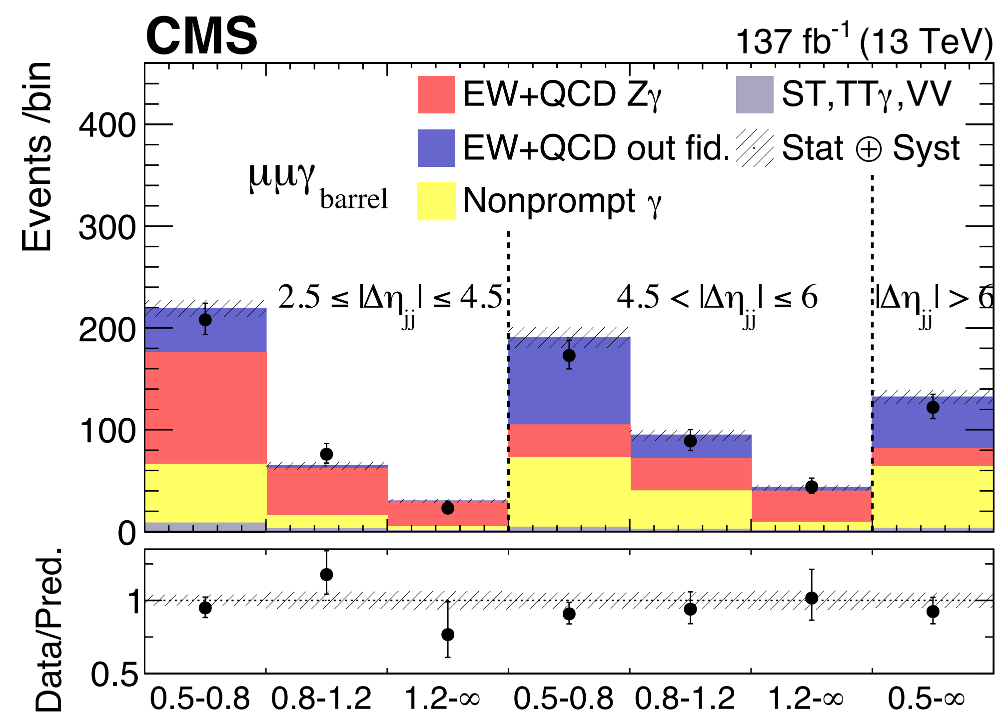
- The observed (expected) significance is 9.4 σ (8.5 σ)

Signal events concentrate on the high m_{jj} and $\Delta\eta_{jj}$ bins

Measurement of cross section

SMP-21-016, PhysRevD.104.072001





- Remove the normalization part of the theoretical uncertainties
- Divide the electroweak Zγ+2j yield into ***in-fiducial*** and ***out-fiducial*** contributions

$$\sigma_{\text{theory}} = 13.3 \pm 1.72\,(\text{scale}) \pm 0.10\,(\text{PDF})$$

$$\mu = 1.11^{+0.06}_{-0.06}\,(\text{stat})\,^{+0.10}_{-0.09}\,(\text{syst}) = 1.11^{+0.12}_{-0.11}$$

$$\sigma_{\text{measured}} = 14.7 \pm 0.80\,(\text{stat}) \pm 1.26\,(\text{syst}) = 14.7 \pm 1.53\text{ fb}$$

- $\sigma_{\text{EW}}^{\text{SM pred.}} = 4.34 \pm 0.26\,(\text{scale}) \pm 0.06\,(\text{PDF})$ fb
- $\mu = 1.20^{+0.12}_{-0.12}\,(\text{stat})\,^{+0.14}_{-0.12}\,(\text{syst}) = 1.20^{+0.18}_{-0.17}$
- $\sigma_{\text{measured}} = 5.21 \pm 0.52\,(\text{stat}) \pm 0.56\,(\text{syst})$ fb $= 5.21 \pm 0.76$ fb

The systematic uncertainty is comparable with the statistical uncertainty

Results — EW cross section

Systematic uncertainty	Impact [%]	
Jet energy correction	+7.9	−6.7
Theoretical uncertainties	+5.5	−4.7
MC statistical uncertainties	+4.7	−4.5
PU	+4.7	−4.1
Related to e, γ	+4.5	−3.6
PU jet ID	+3.7	−3.4
ECAL timing shift at L1	+3.5	−2.8
Nonprompt-γ bkg. estimate	+2.0	−1.6
Related to μ	+1.7	−1.4
Integrated luminosity	+0.8	−0.6
Total systematic uncertainty	+14	−12

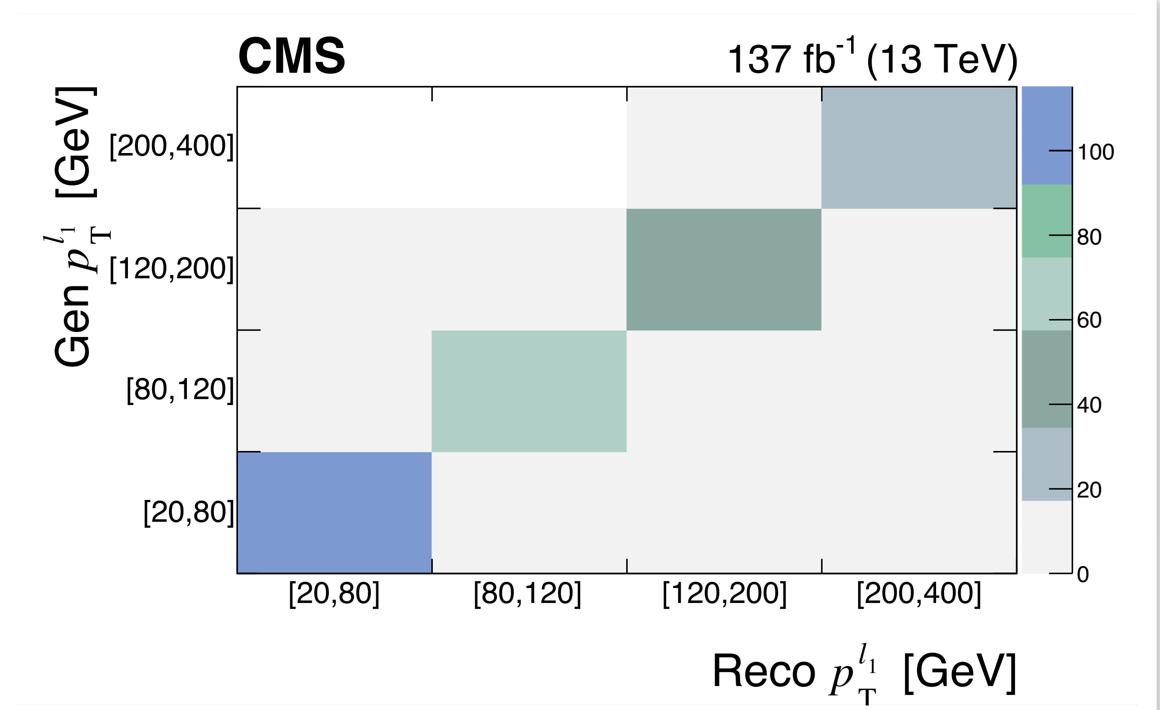
Unfolded differential cross sections

Similar with the fiducial XS measurement, we perform 'unfolding' to revert the 'detector smearing' on the data to get the 'True' distribution

$$\mathscr{L}(\vec{\mu};\vec{\theta}) = \prod_i \mathrm{Poisson}(y_i \mid \sum_j R_{ij}(\vec{\theta}) u_j s_j(\vec{\theta}) + b_i(\vec{\theta})) \cdot p(\tilde{\vec{\theta}} \mid \vec{\theta})$$

$$y_i^{\mathrm{reco}} = \sum_j R_{ij} \cdot x_j^{\mathrm{gen}} + b_i,$$

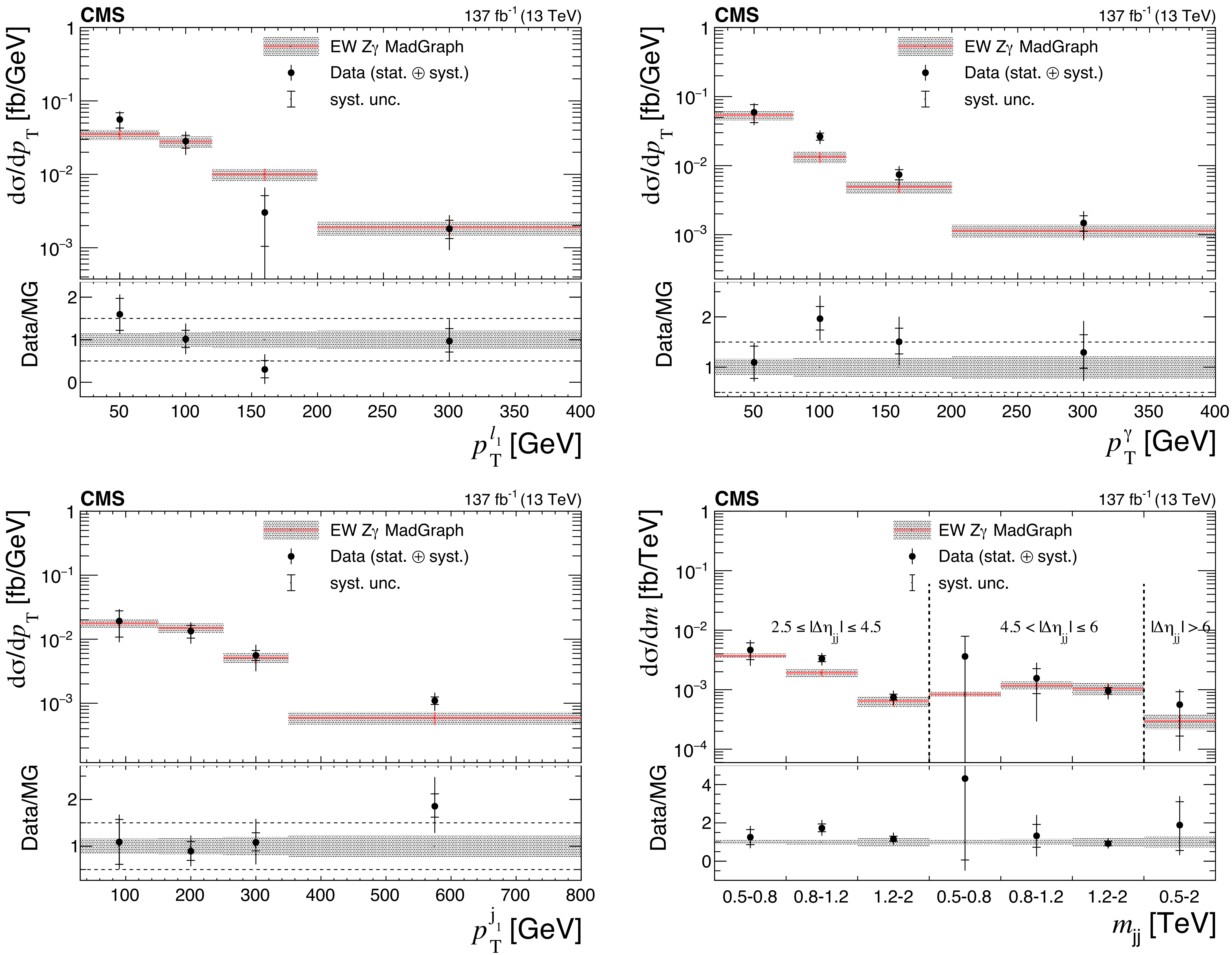
$$R_{ij} = P(\text{observed in bin}_\mathrm{i} \mid \text{generated in bin}_\mathrm{j})$$



- Each reconstructed bin (j) describes the contribution from each truth bin (i) - this is the R_{ji} (response matrix)
  - ☑ Condition number of the R is smaller than about 10, so the regularization is not needed
- Same uncertainties with significance measurement
- Differential cross sections are measured for:
  - Three 1D variables: p_T^{γ}, $p_T^{\ell_1}$, $p_T^{j_1}$,
  - One 2D variable: m_{jj}-$\Delta\eta_{jj}$

Nonsingular, Condition number < 10

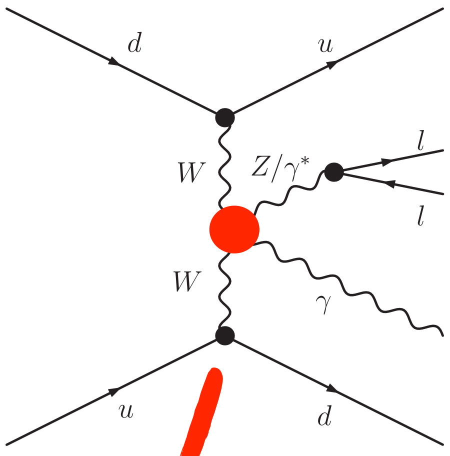
Unfolded differential cross sections



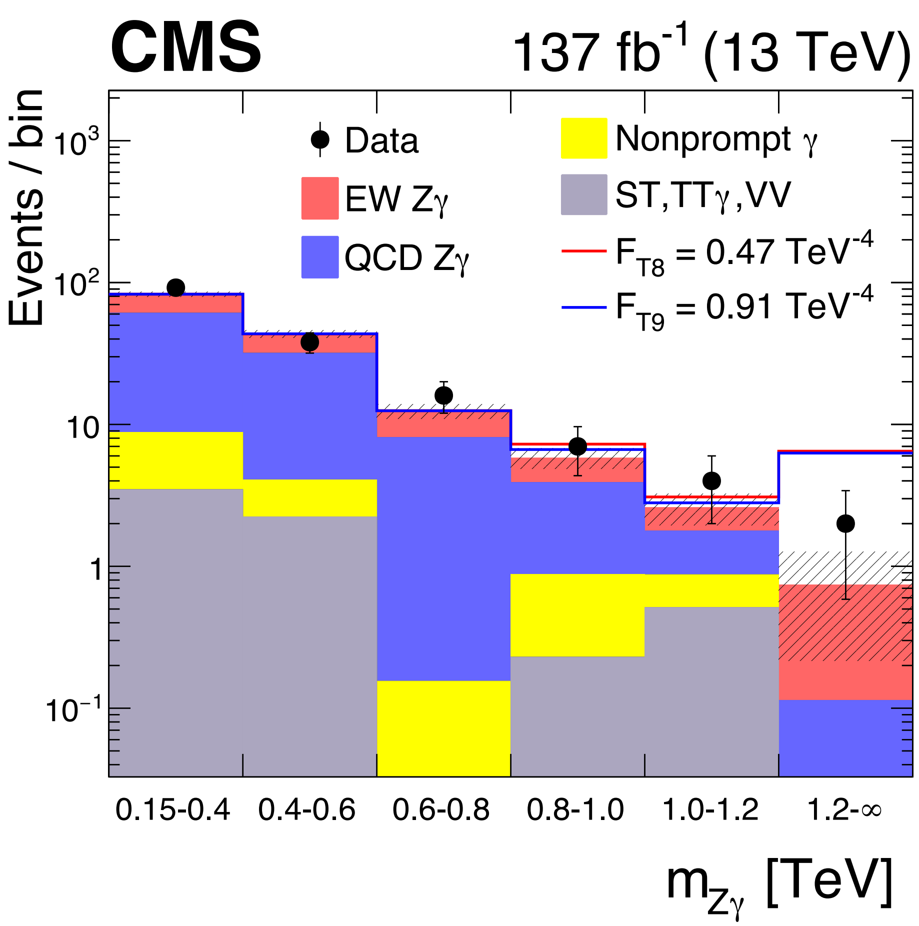
Within the uncertainties, the measurements agree with the predictions.

Results — aQGC

- Extended the SM Lagrangian with higher dimensional operators maintaining SU(2)×U(1) gauge symmetry:



$$\mathscr{L}_{EFT} = \mathscr{L}_{SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2}\mathscr{O}^{(6)} + \frac{c_i^{(8)}}{\Lambda^4}\mathscr{O}^{(8)} + \ldots .$$



Diboson system

	$\mathcal{O}_{S,0}$, $\mathcal{O}_{S,1}$, $\mathcal{O}_{S,2}$	$\mathcal{O}_{M,0}$, $\mathcal{O}_{M,1}$, $\mathcal{O}_{M,7}$	$\mathcal{O}_{M,2}$, $\mathcal{O}_{M,3}$, $\mathcal{O}_{M,4}$, $\mathcal{O}_{M,5}$	$\mathcal{O}_{T,0}$, $\mathcal{O}_{T,1}$, $\mathcal{O}_{T,2}$	$\mathcal{O}_{T,5}$, $\mathcal{O}_{T,6}$, $\mathcal{O}_{T,7}$	$\mathcal{O}_{T,8}$, $\mathcal{O}_{T,9}$
$WWWW$	X	X		X		
$WWZZ$	X	X	X	X	X	
$ZZZZ$	X	X	X	X	X	X
$WWZ\gamma$		X	X	X	X	
$WW\gamma\gamma$		X	X	X	X	
$ZZZ\gamma$		X	X	X	X	X
$ZZ\gamma\gamma$		X	X	X	X	X
$Z\gamma\gamma\gamma$				X	X	X
$\gamma\gamma\gamma\gamma$				X	X	X

The most stringent limit for operator T_9

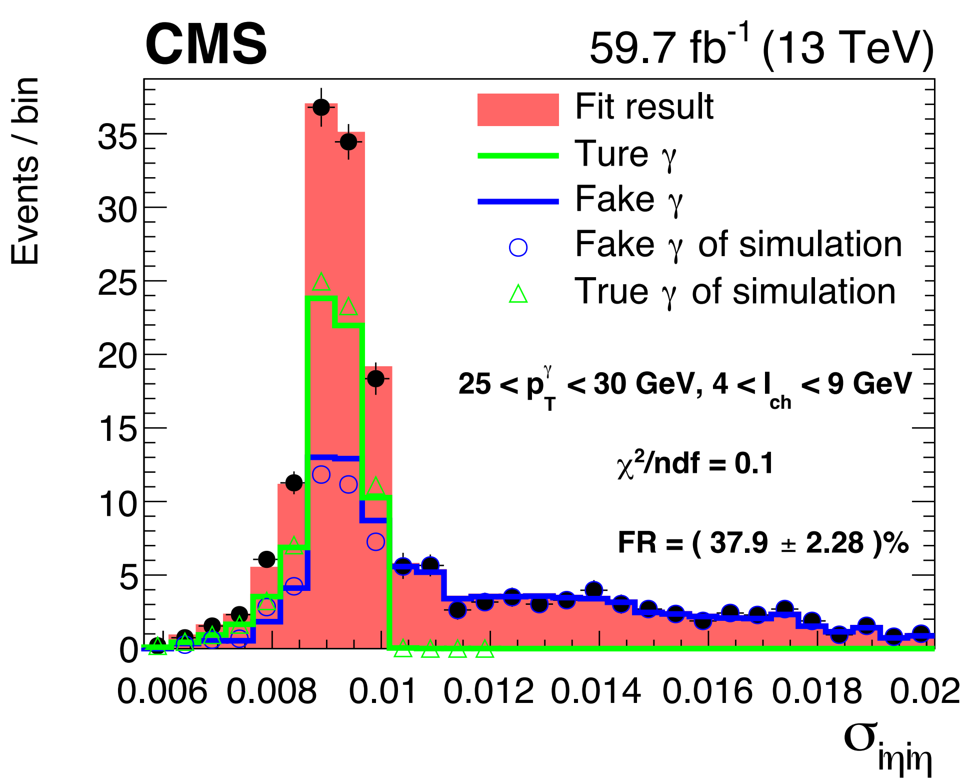
Summary

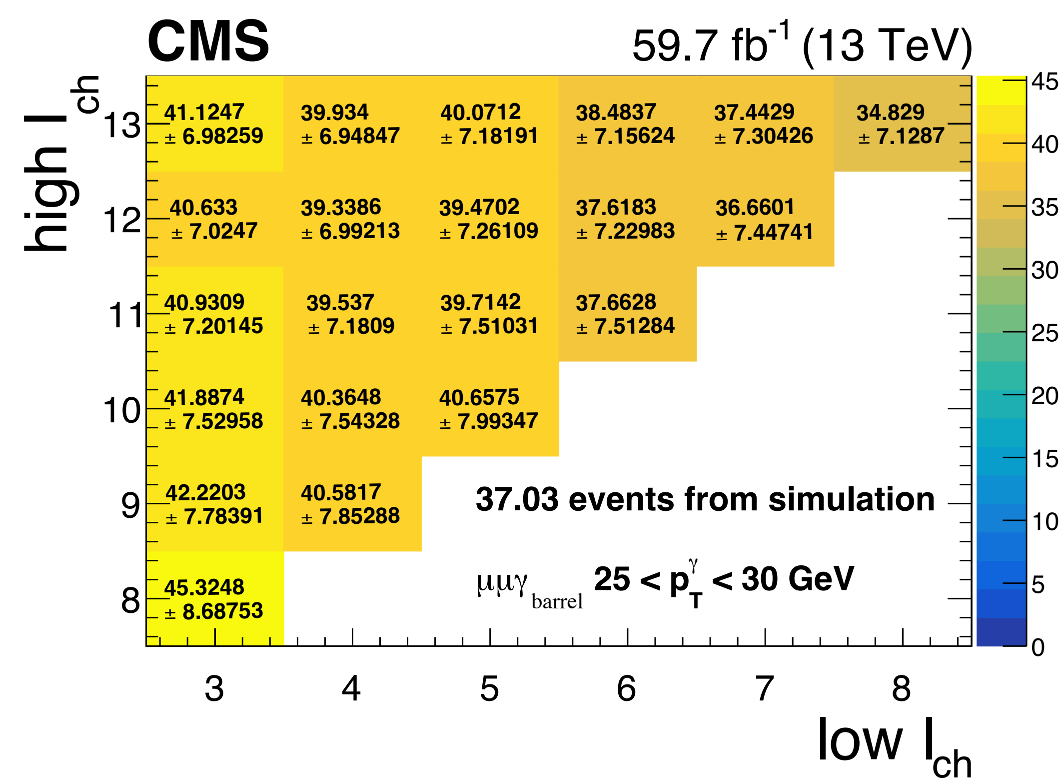
- ✓ First observation of VBS Zγ with significance far more 5σ
- ✓ Accurate fiducial cross section measurement
- ✓ First unfolded differential cross sections measurement
- ✓ AQGC limits for operator $M_{0-5,7}$, T_{0-2}, and T_{5-9}
  - ✓ Limit for T_9 is the most stringent limit to date

Backup

Nonprompt γ estimation

- Construct pseudo-data using all MC samples
- Perform the template fit on this pesudo-data with different choice of charged isolation sideband
- Compare the fit results with information from simulation directly





- The **charged isolation sideband is well tuned** in the nonprompt γ estimation
    - Give us more accurate nonprompt γ estimation
    - Control the corresponding uncertainty well

Backup

sigmaIEtaIEta is the log energy weighted RMS of the shower in units of crystals

- $\sigma_{i\eta i\eta} = \sqrt{\left(\frac{\Sigma_i^{5\times5} w_i (\eta_i - \bar{\eta}_{5\times5})^2}{\Sigma_i^{5\times5} w_i}\right)}$
- $w_i = 4.7 + \ln \frac{E_i}{E_{5\times5}}$
  - this is effectively a noise cut, each crystal needs to have > 0.9% of 5x5 energy
  - means that very low energy electrons are sensitive to noise as 0.9% of a small number brings it below noise threshold
- E_i = energy of crystal, $E_{5\times5}$ energy of 5x5
  - likewise for η
- η is in units of crystals, not absolute η
  - endcap uses $(ix^2 + iy^2)^{1/2}$ to get η in terms of crystals
- normalised to 0.01745 in barrel and 0.0447 in endcap
- cut effectively means that all the energy is within two crystals

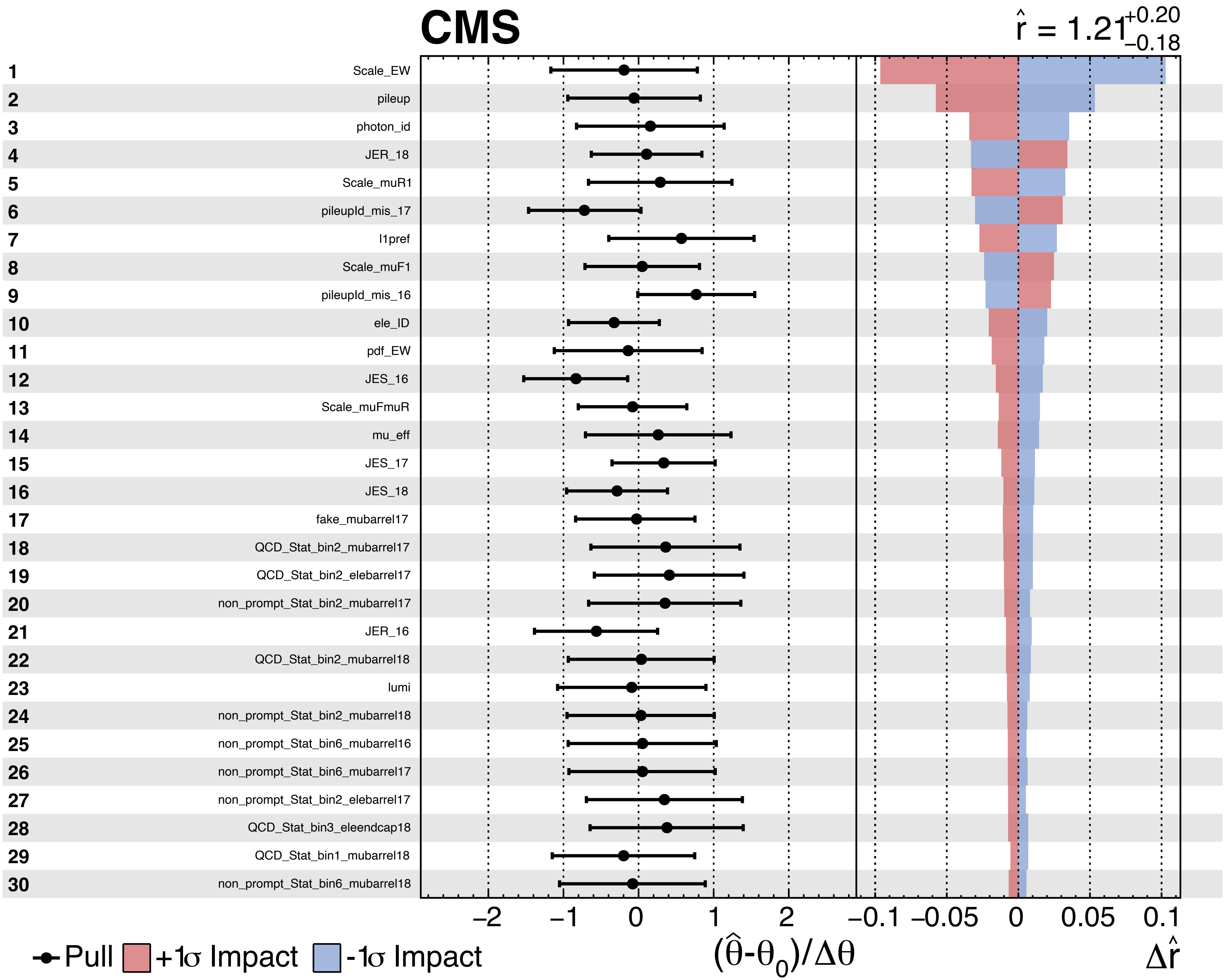
aQGC limits

As the sensitivity on the T_i operators of VBS Zγ, we show the comparison of the limits of T_i from recent public VBS results with the full Run2 data

Operator	SMP-20-016 VBS Zγ	SMP-20-001 VBS ZZ	SMP-19-012 VBS $W^{\pm}W^{\pm}$
f_{T0}	-0.64 , 0.57	-0.24 , 0.22	-0.28 , 0.31
f_{T1}	-0.81 , 0.90	-0.31 , 0.31	-0.12 , 0.15
f_{T2}	-1.68 , 1.54	-0.63 , 0.59	-0.38 , 0.50
f_{T5}	-0.58 , 0.64	—	—
f_{T6}	-1.30 , 1.33	—	—
f_{T7}	-2.15 , 2.43	—	—
f_{T8}	-0.47 , 0.47	-0.43 , 0.43	—
f_{T9}	-0.91 , 0.91	-0.92 , 0.92	—

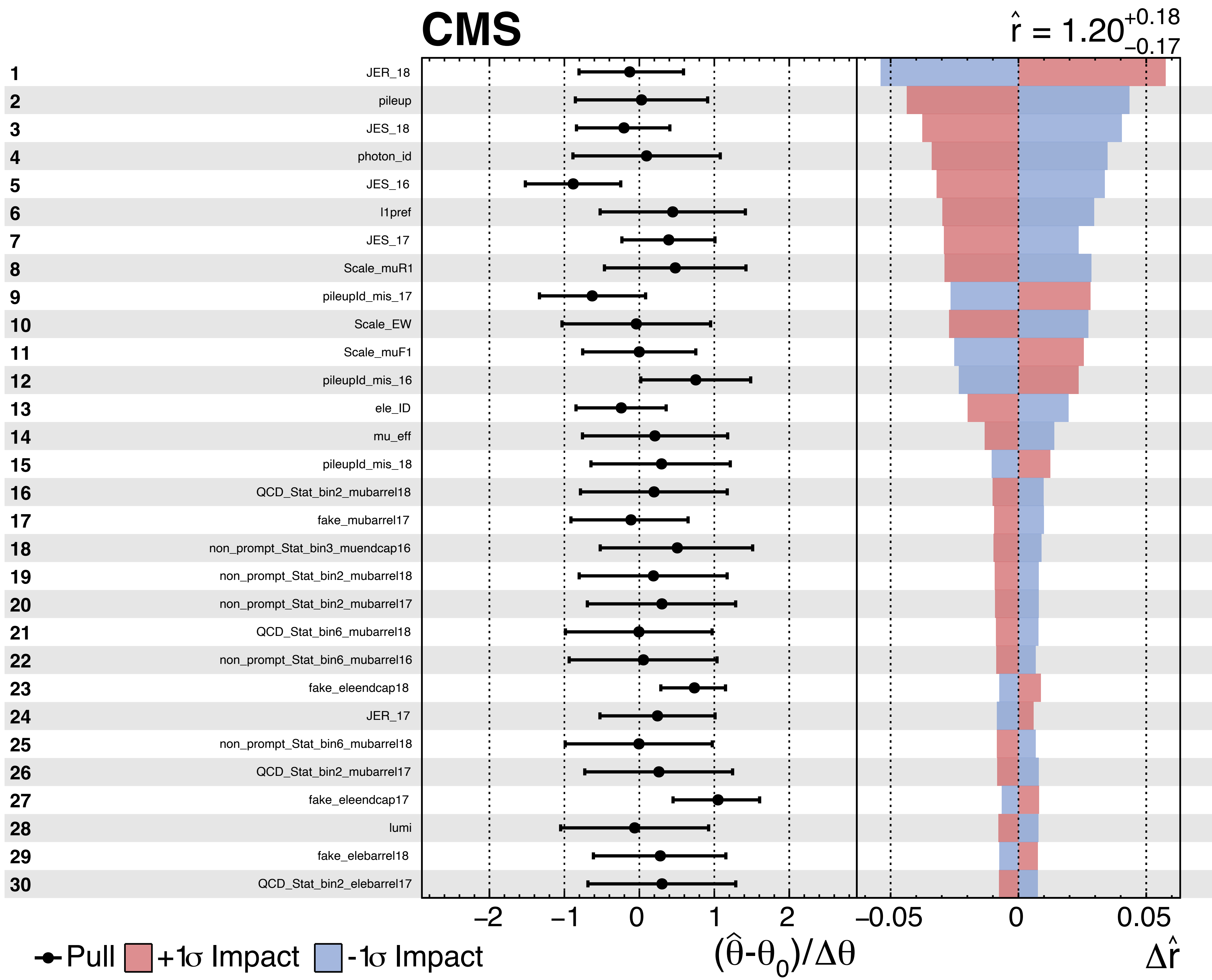
Similar sensitivity on T_8 and T_9 between VBS Zγ and VBS ZZ, which is expected, as the T_8 and T_9 give rise to QGCs only containing the neutral gauge bosons.

Backup



For signal significance

Backup



For signal strength

Backup

variables	2016	2017	2018
p_T^{γ}	1.08	1.12	1.21
$p_T^{j_1}$	1.35	1.41	1.44
$p_T^{l_1}$	1.09	1.09	1.11
m_{jj}-$\Delta\eta jj$	1.87	1.97	1.95

Condition Number of R for EW

variables	2016	2017	2018
p_T^{γ}	1.16	1.41	1.37
$p_T^{j_1}$	1.33	1.41	1.39
$p_T^{l_1}$	1.10	1.35	1.16
m_{jj}-$\Delta\eta jj$	1.93	2.32	2.09

Condition Number of R for EW+QCD

If the condition number is small (~10), then the problem is well-conditioned and can most likely be solved using the unregularized maximum likelihood estimate (MLE). This happens when the resolution effects are small and R is almost diagonal. If on the other hand, the condition number is large (~10^5) then the problem is ill-conditioned and the unfolded estimator needs to be regularized.

Backup

Building blocks:

- $D_\mu\Phi$: Higgs doublet field, affects the coupling of longitudinal modes of the gauge bosons.
- $\hat{W}_{\mu\nu}$, $\hat{B}_{\mu\nu}$: Field strength tensors

Dimension-8 operators (only field strength/mixed)

$$\mathcal{O}_{T,0} = \mathrm{Tr}\left[W_{\mu\nu}W^{\mu\nu}\right] \cdot \mathrm{Tr}\left[W_{\alpha\beta}W^{\alpha\beta}\right] ,$$
$$\mathcal{O}_{T,1} = \mathrm{Tr}\left[W_{\alpha\nu}W^{\mu\beta}\right] \cdot \mathrm{Tr}\left[W_{\mu\beta}W^{\alpha\nu}\right] ,$$
$$\mathcal{O}_{T,2} = \mathrm{Tr}\left[W_{\alpha\mu}W^{\mu\beta}\right] \cdot \mathrm{Tr}\left[W_{\beta\nu}W^{\nu\alpha}\right] ,$$
$$\mathcal{O}_{T,5} = \mathrm{Tr}\left[W_{\mu\nu}W^{\mu\nu}\right] \cdot B_{\alpha\beta}B^{\alpha\beta} ,$$
$$\mathcal{O}_{T,6} = \mathrm{Tr}\left[W_{\alpha\nu}W^{\mu\beta}\right] \cdot B_{\mu\beta}B^{\alpha\nu} ,$$
$$\mathcal{O}_{T,7} = \mathrm{Tr}\left[W_{\alpha\mu}W^{\mu\beta}\right] \cdot B_{\beta\nu}B^{\nu\alpha} ,$$
$$\mathcal{O}_{T,8} = B_{\mu\nu}B^{\mu\nu}B_{\alpha\beta}B^{\alpha\beta}$$
$$\mathcal{O}_{T,9} = B_{\alpha\mu}B^{\mu\beta}B_{\beta\nu}B^{\nu\alpha} .$$

$$\mathcal{O}_{M,0} = \mathrm{Tr}\left[W_{\mu\nu}W^{\mu\nu}\right] \cdot \left[(D_\beta\Phi)^\dagger D^\beta\Phi\right]$$
$$\mathcal{O}_{M,1} = \mathrm{Tr}\left[W_{\mu\nu}W^{\nu\beta}\right] \cdot \left[(D_\beta\Phi)^\dagger D^\mu\Phi\right]$$
$$\mathcal{O}_{M,2} = \left[B_{\mu\nu}B^{\mu\nu}\right] \cdot \left[(D_\beta\Phi)^\dagger D^\beta\Phi\right] ,$$
$$\mathcal{O}_{M,3} = \left[B_{\mu\nu}B^{\nu\beta}\right] \cdot \left[(D_\beta\Phi)^\dagger D^\mu\Phi\right] ,$$
$$\mathcal{O}_{M,4} = \left[(D_\mu\Phi)^\dagger W_{\beta\nu}D^\mu\Phi\right] \cdot B^{\beta\nu} ,$$
$$\mathcal{O}_{M,5} = \left[(D_\mu\Phi)^\dagger W_{\beta\nu}D^\nu\Phi\right] \cdot B^{\beta\mu} ,$$
$$\mathcal{O}_{M,6} = \left[(D_\mu\Phi)^\dagger W_{\beta\nu}W^{\beta\nu}D^\mu\Phi\right] ,$$
$$\mathcal{O}_{M,7} = \left[(D_\mu\Phi)^\dagger W_{\beta\nu}W^{\beta\mu}D^\nu\Phi\right] ,$$