Recent V+jets measurements in CMS

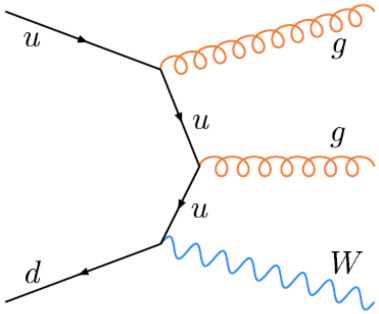
Duong Nguyen
University at Buffalo

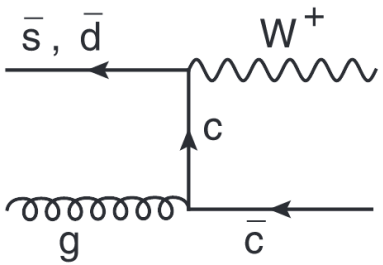
On behalf of the CMS collaboration

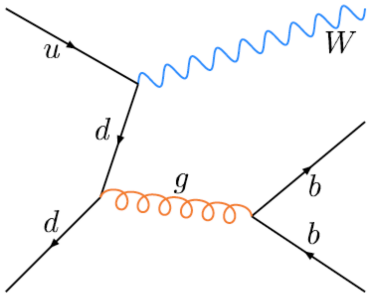
CERN-LHC seminar
March 16, 2021

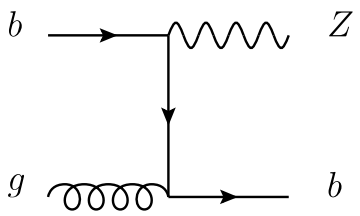
Motivation to measure V+jets

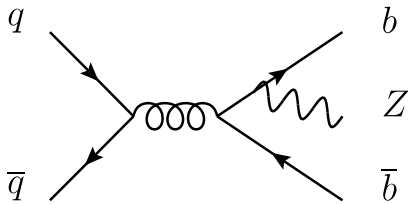
- Studies of V+jets processes provides important tests of pQCD calculations and validation of MC generators
  - Precise experimental results are obtained mostly from the leptonic channels
  - Various (N)NLO predictions including NLO ME+PS are available for both V+jets and V+heavy flavor (c, b) jets (V+HF jets)
- V+HF jets are sensitive to s, c, and b parton distribution functions

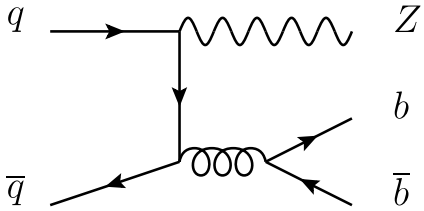








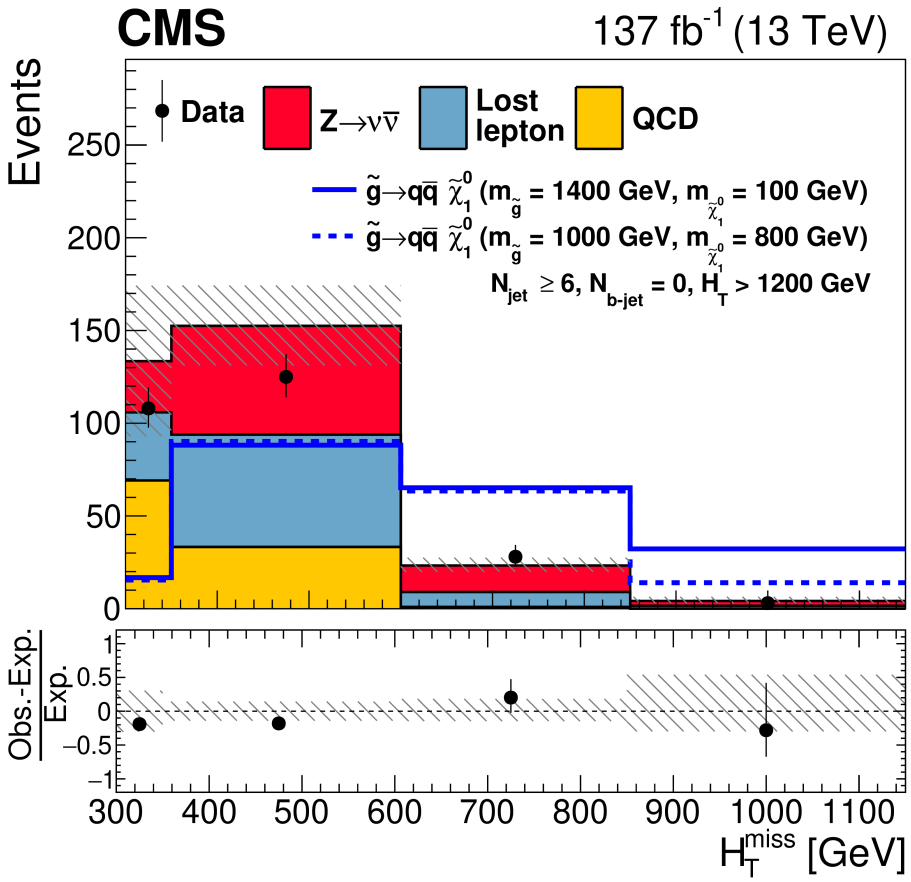




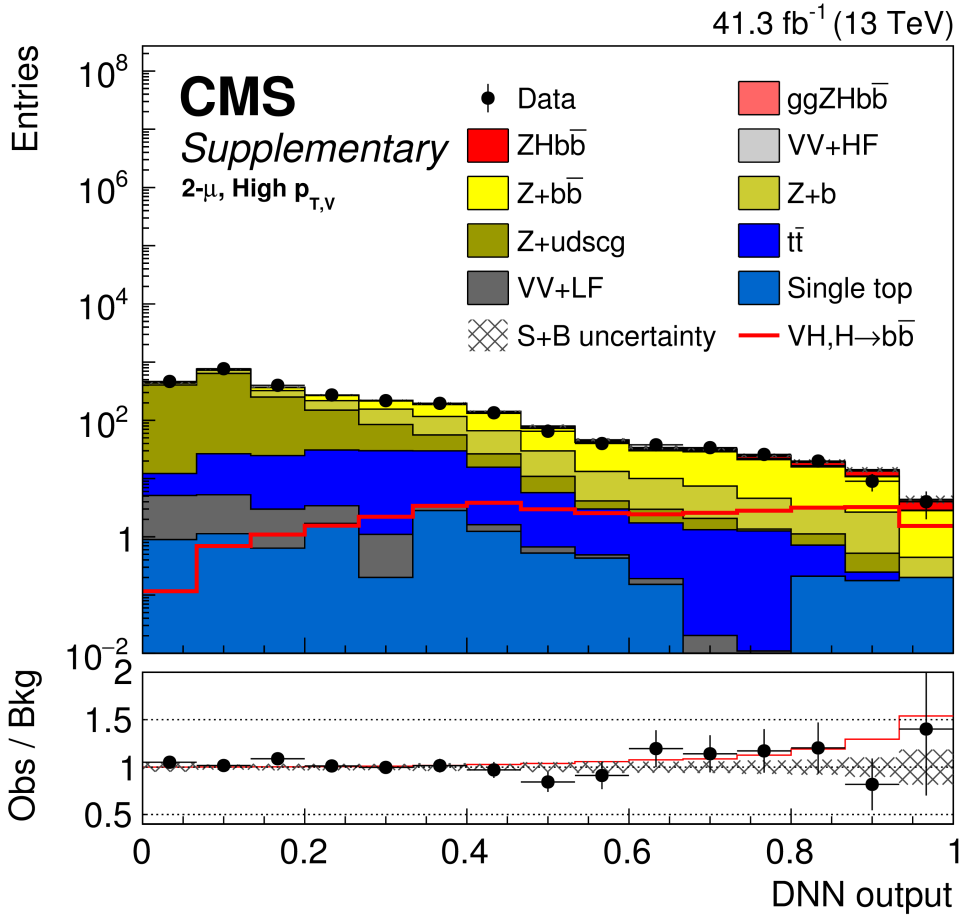
Motivation (cont.)

- V+jets are irreducible backgrounds for SM measurements and BSM searches

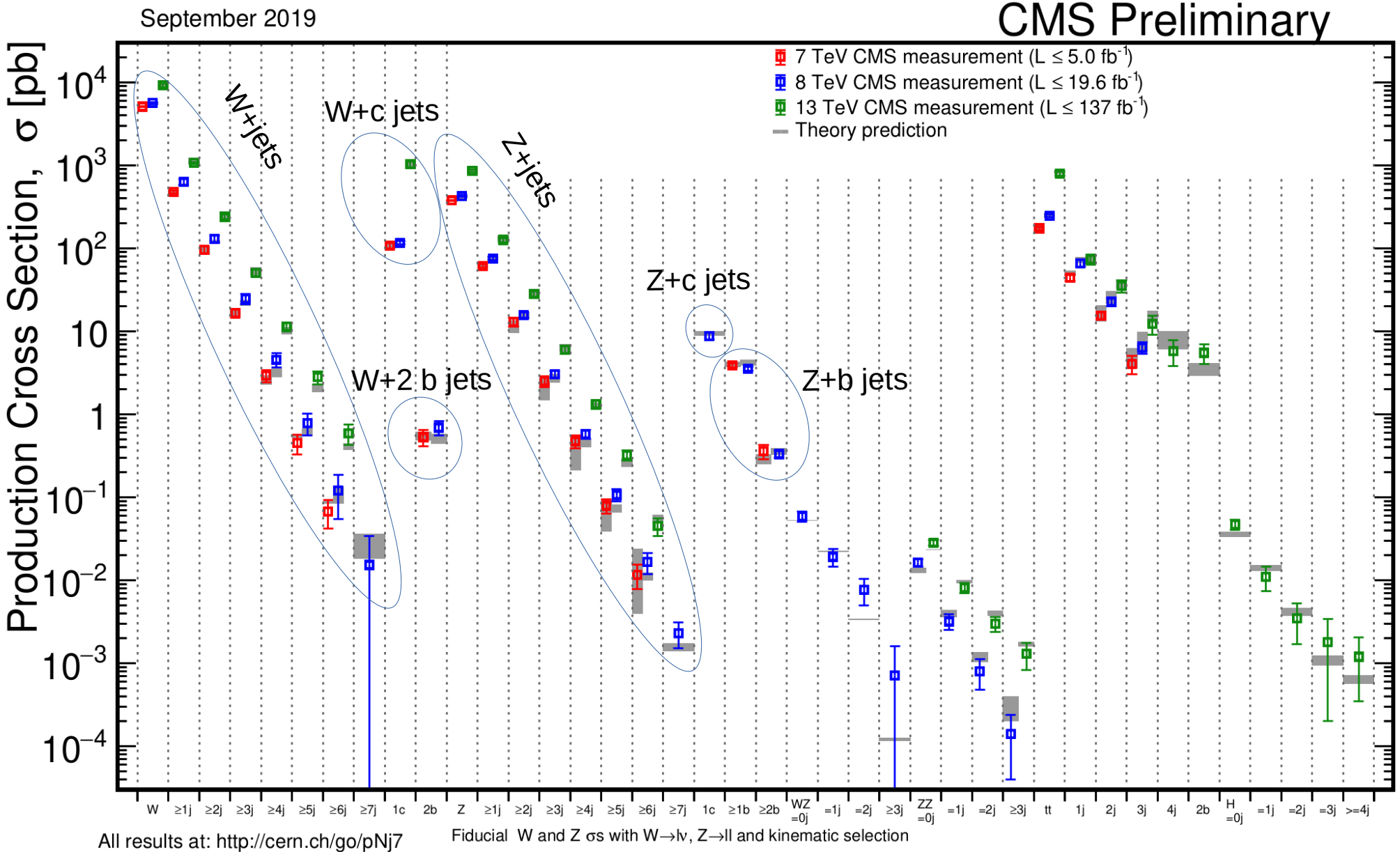
JHEP 10 (2019) 244
Search for supersymmetry with jets and missing transverse energy



Phys. Rev. Lett. 121 (2018) 121801
VH(→bb) observation



Overview of CMS V+jets measurements



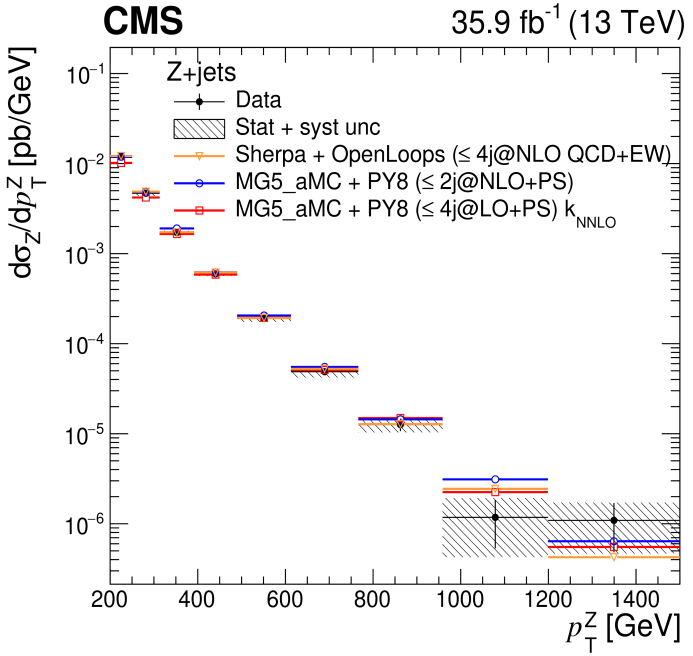
- Various V+X cross section results spanning over ~6 orders of magnitude
- In general, within uncertainties, predictions agree with data

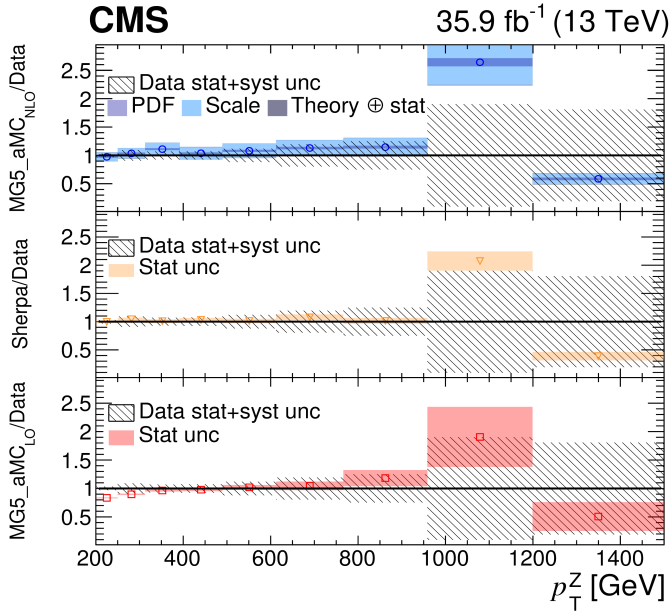
arXiv:2102.02238

Measurements of the differential cross sections of the production of Z + jets and γ + jets and of Z boson emission collinear with a jet in pp collisions at $\sqrt{s} = 13\,\text{TeV}$

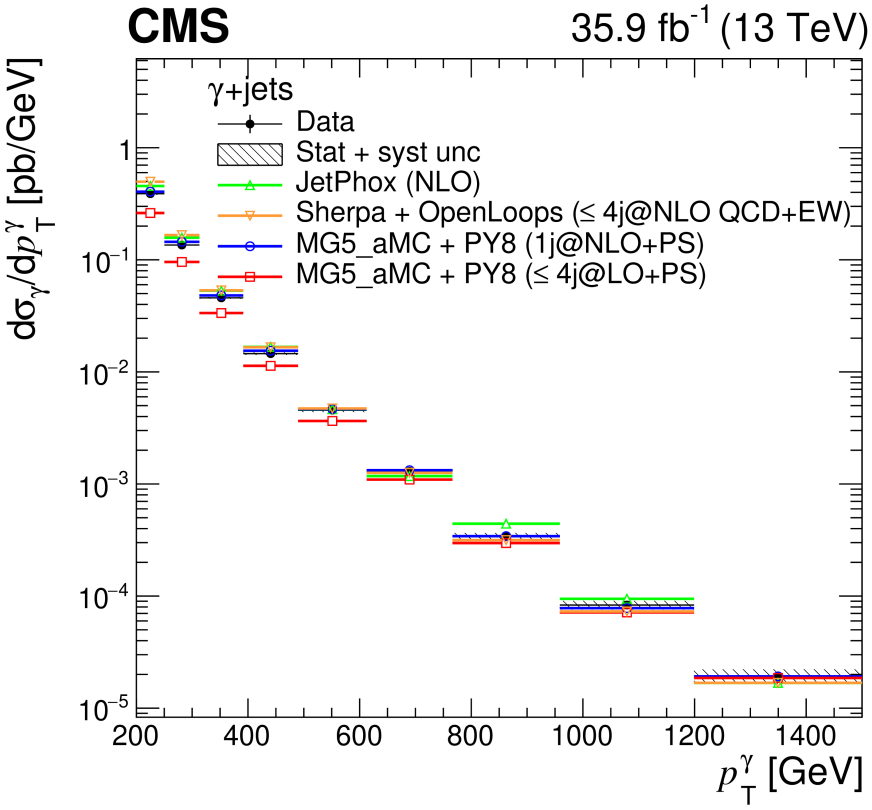
Z+jets cross sections

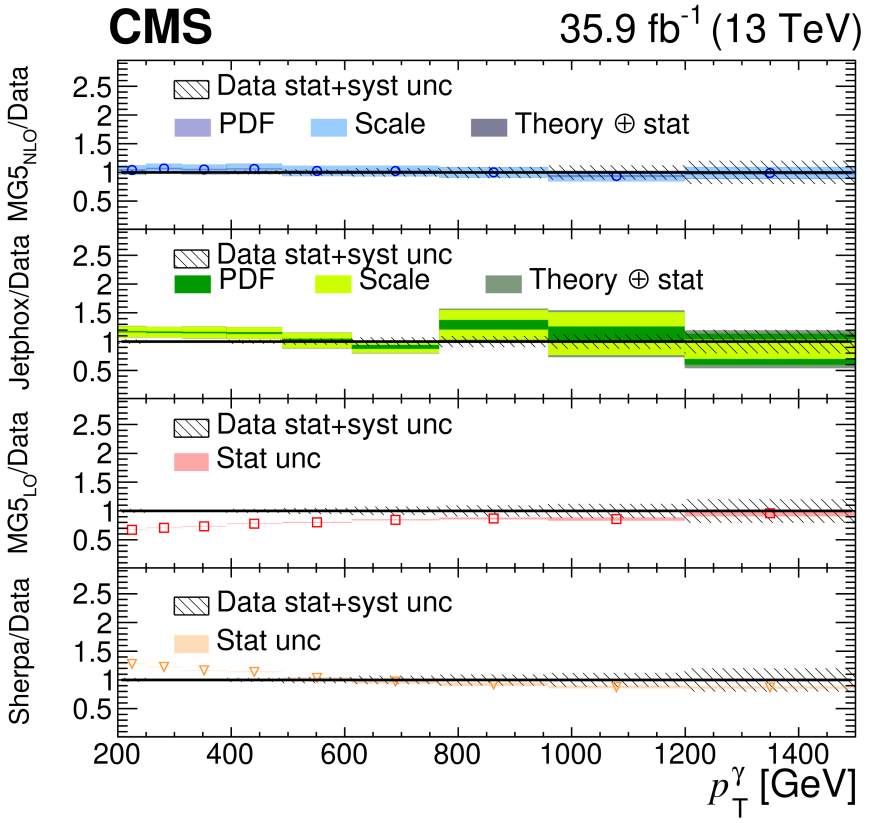
- The ratio between Z+jets/γ+jets is sensitive to high order QCD and EW corrections at high p_T range
- It is also useful for estimation of $Z \to \nu\nu$ backgrounds from γ+jets for new physics searches and invisible Higgs measurements
- $Z(\to \ell\ell)$ +jets selections
  - Z p_T > 200 GeV, |y| < 1.4
  - At least one jet with p_T > 100 GeV, |η| < 2.4
- MG5_aMC+PY8 (Madgraph5_aMC@NLO interfaced with Pythia8) LO prediction is normalized to a NNLO cross section (FEWZ)
- MG5_aMC+PY8 (≤ 2 jets@NLO+PS) and Sherpa+OpenLoops (≤ 4 jets@NLO QCD+EW) predictions agree with data
- MG5_aMC+PY8 (≤ 4 jets@LO+PS) underestimate data below 300 GeV





γ+jets cross sections





- γ+jet selections
  - A prompt, isolated photon

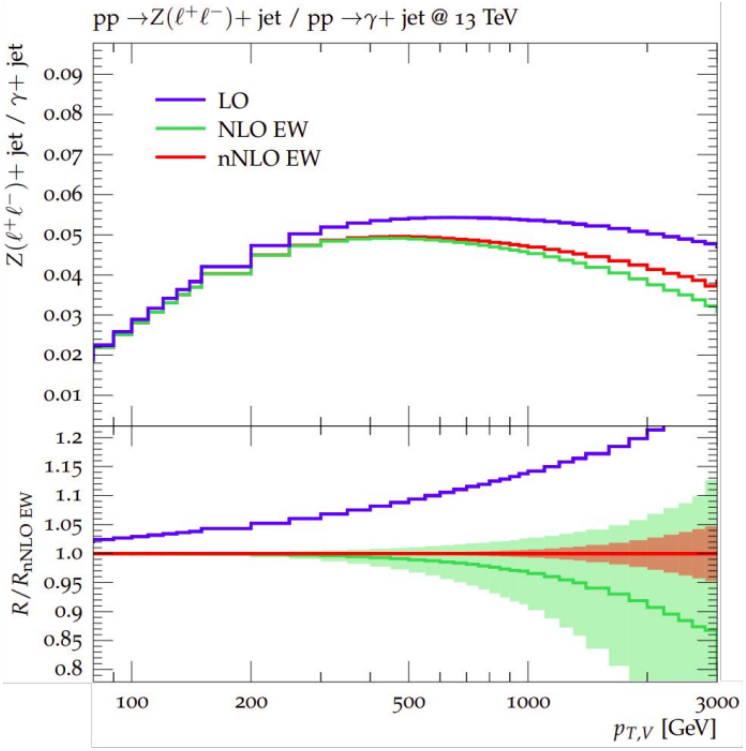
    p_T > 200 GeV and |η| < 1.4
  - At least one jet with

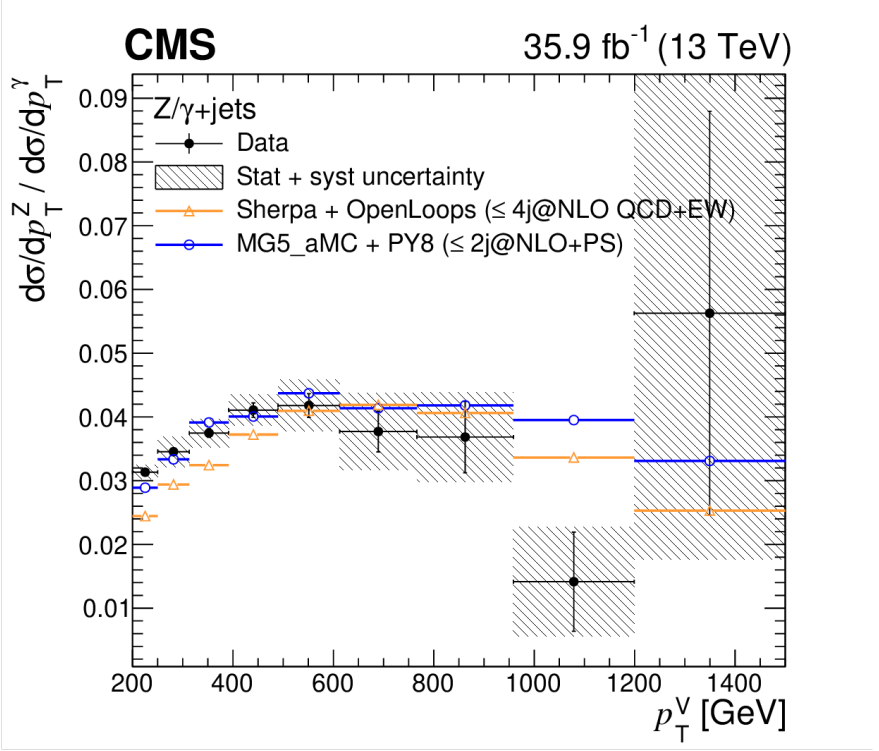
    p_T > 100 GeV, |η| < 2.4

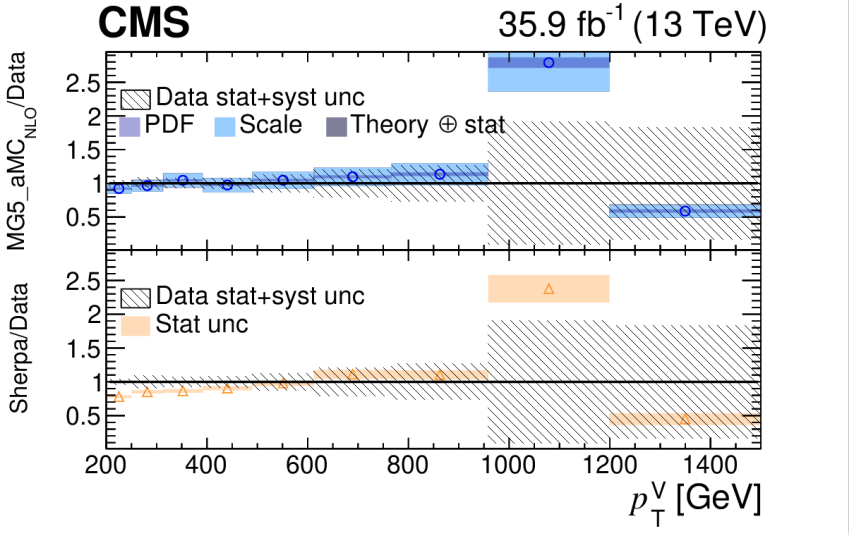
- MG5_aMC+PY8 (≤ 2 jets@NLO+PS) is consistent with data
- Both JETPHOX (NLO) and Sherpa+OpenLoops (≤ 4 jets@NLO QCD+EW) overestimate data by ~20% below p_T < 500 GeV
- MG5_aMC+PY8 (≤ 4 jets@LO+PS) shows significant disagreement with data

Z+jets/γ+jets cross section ratio

Lindert, J.M. et al. arXiv:1705.04664



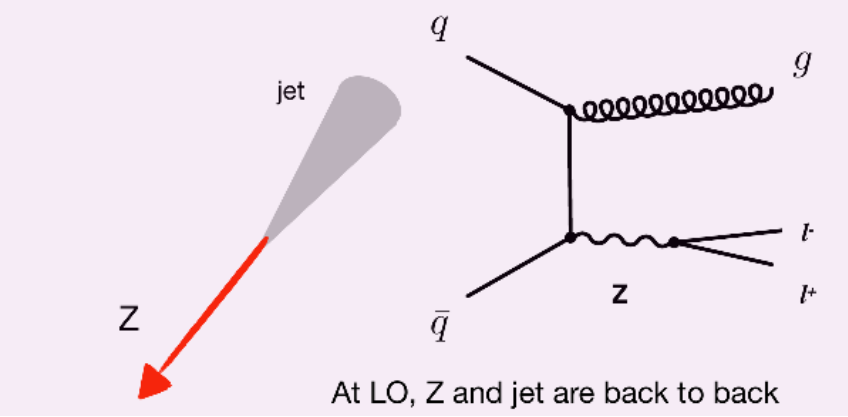


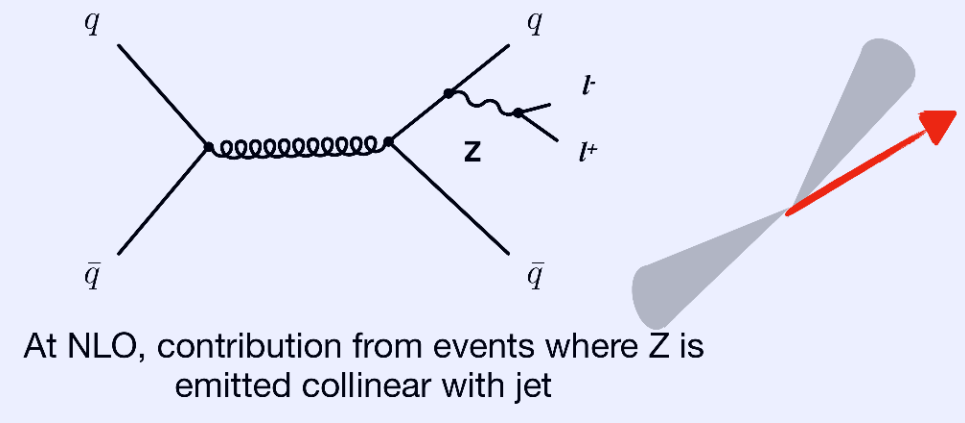


- The Z+jets/γ+jets cross section ratio is reduced at high p_T when including EW corrections
- Uncertainties in high p_T bins are dominated by energy resolution and statistic
- Good agreement with MG5_aMC+PY8 (≤ 2 jets@NLO+PS)
- Sherpa+OpenLoops (≤ 4 jets@NLO QCD+EW) is lower than data by 10-20% at low p_T but consistent with data within uncertainties above 400 GeV

Collinear Z emission

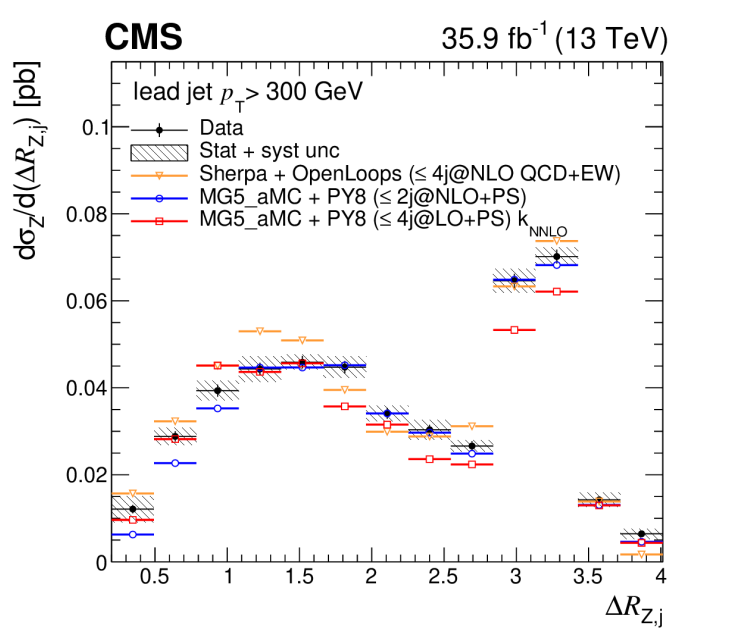
- Measure differential cross section as a function of ΔR(Z, closest jet) to probe collinear emission of the Z boson. Collinear Z emission effect is enhanced by requiring a high p_T jet (p_T > 300 or 500 GeV)
- Sensitive to higher order QCD and EWK corrections

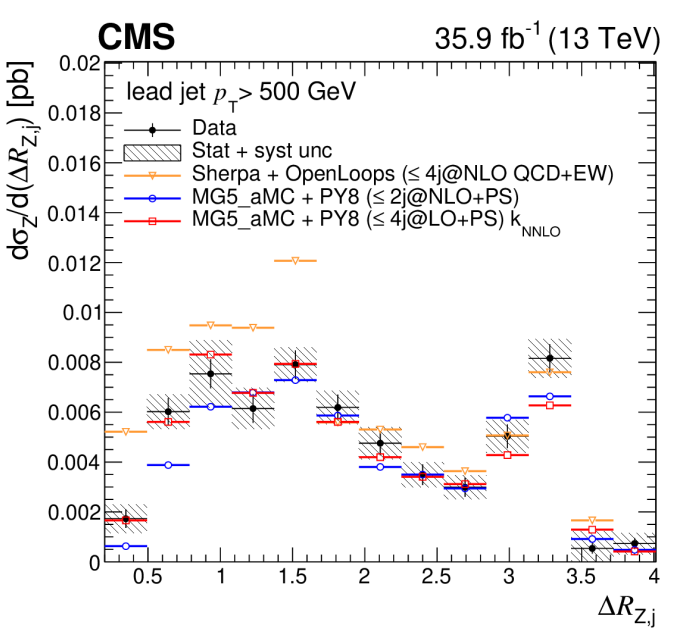


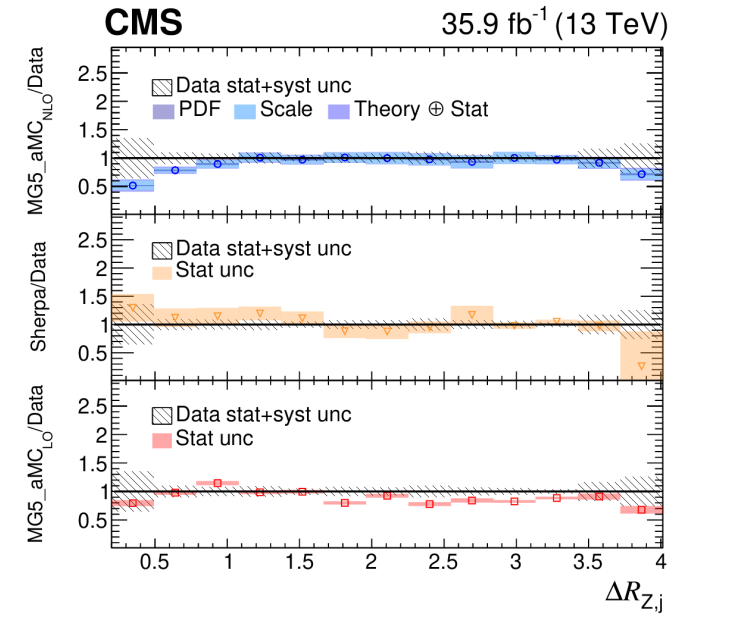


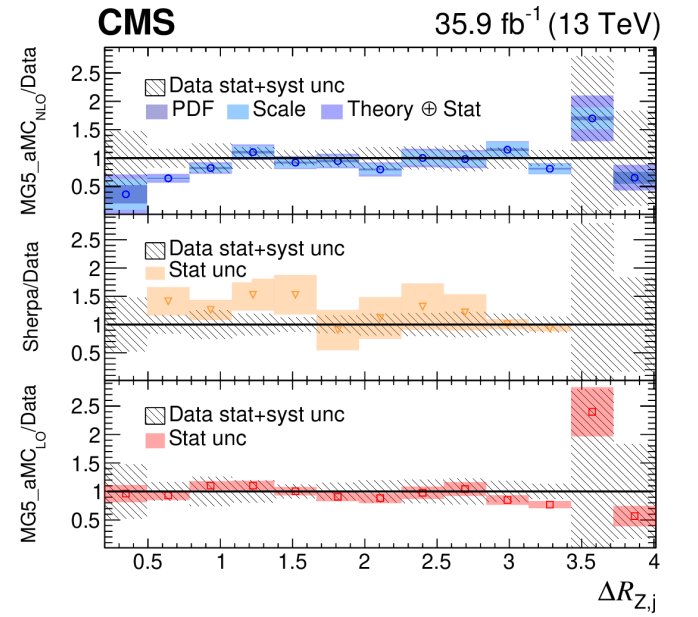
Collinear Z emission (cont.)

- Cross sections as a function of ΔR with two different requirements on leading jet p_T
- MG5 LO underestimates data at high ΔR regions
- MG5 NLO underestimates data at low ΔR regions,
- Sherpa+OpenLoops is higher than data below 1.8 but has large statistical uncertainty









CMS PAS SMP-20-009

Measurement of distributions sensitive to double parton scattering using Z bosons produced in association with jets at 13 TeV

Double parton scattering (DPS)

- DPS: two hard parton-parton interactions in a single pp collision
- Why measure DPS processes?
  - Irreducible backgrounds to SM measurement and new physics searches (hard DPS events)
  - Extracting parameters for DPS modelling in MC
- DPS cross section is related to effective cross section (a measure of transverse distribution of partons inside colliding hadrons and their overlap in a collision):

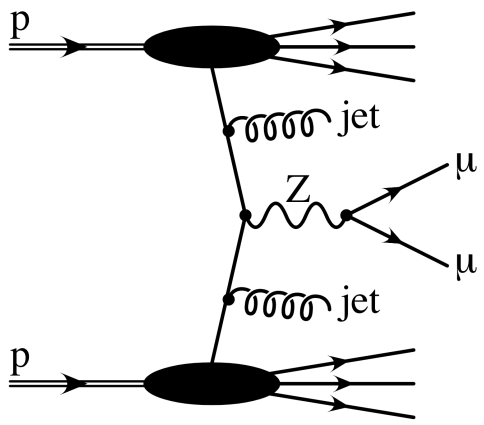
$$\sigma_{\text{eff}} = \frac{m}{2}\frac{\sigma_{\text{A}} \cdot \sigma_{\text{B}}}{\sigma_{\text{A+B}}^{\text{DPS}}}$$

m: symmetry factor

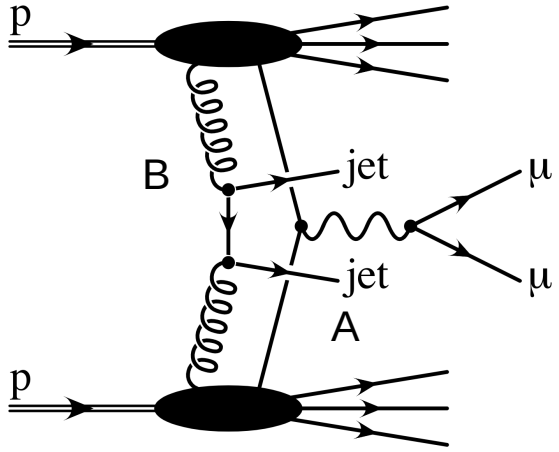
  - *m* = 1 (2) for indistinguishable (distinguishable) processes

- Effective cross section is expected to be independent of processes A and B

Single parton scattering (SPS)

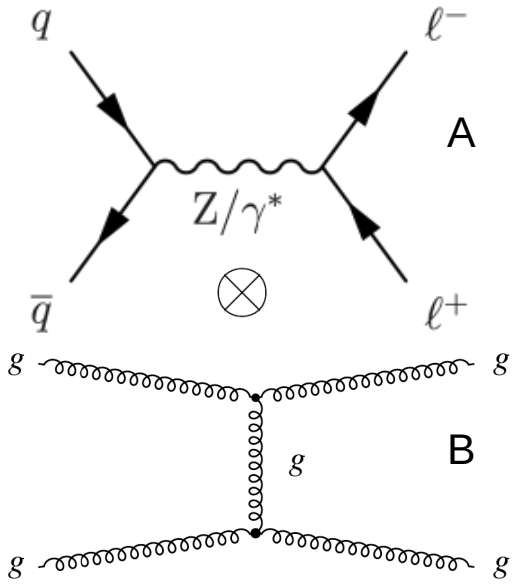


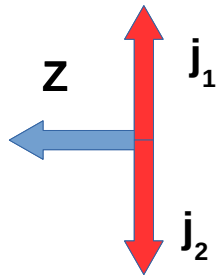
Double parton scattering (DPS)



DPS measurement in the Z+jets process

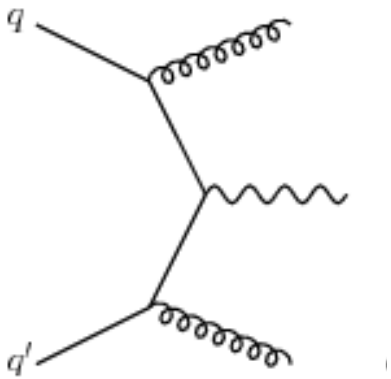
- Probe DPS effects in Z+jets production by measuring the Z+≥1 jet and Z+≥2 jets differential cross sections as functions of observables sensitive to DPS contributions:
  - A and B interactions in DPS are uncorrelated => back-to-back dijets and uncorrelated Z and jets
  - In SPS, Z and jets are correlated

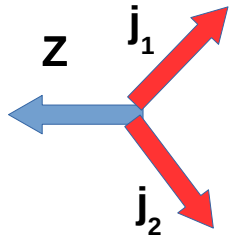




Z+≥1 jet

$$\Delta\phi(Z, j_1) \quad \Delta^{\mathrm{rel}}_{p_\mathrm{T}}(Z, j_1) = \frac{|\vec{p}_\mathrm{T}(Z) + \vec{p}_\mathrm{T}(j_1)|}{|\vec{p}_\mathrm{T}(Z)| + |\vec{p}_\mathrm{T}(j_1)|}$$





Z+≥2 jets

$$\Delta\phi\,(\mathrm{Z}, \mathrm{dijet})$$

$$\Delta^{\mathrm{rel}}_{p_\mathrm{T}}(\mathrm{Z}, \mathrm{dijet}) = \frac{|\vec{p}_\mathrm{T}(Z) + \vec{p}_\mathrm{T}(\mathrm{dijet})|}{|\vec{p}_\mathrm{T}(Z)| + |\vec{p}_\mathrm{T}(\mathrm{dijet})|}$$

$$\Delta^{\mathrm{rel}}_{p_\mathrm{T}}(j_1, j_2) = \frac{|\vec{p}_\mathrm{T}(j_1) + \vec{p}_\mathrm{T}(j_2)|}{|\vec{p}_\mathrm{T}(j_1)| + |\vec{p}_\mathrm{T}(j_2)|}$$

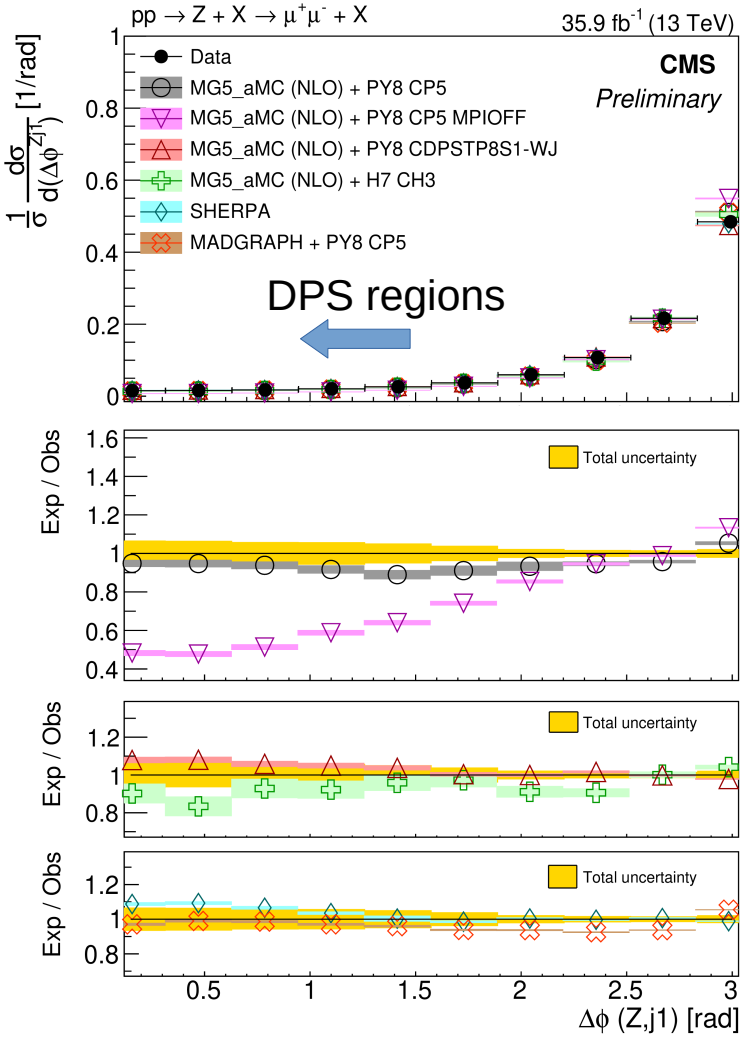
Integrated cross sections

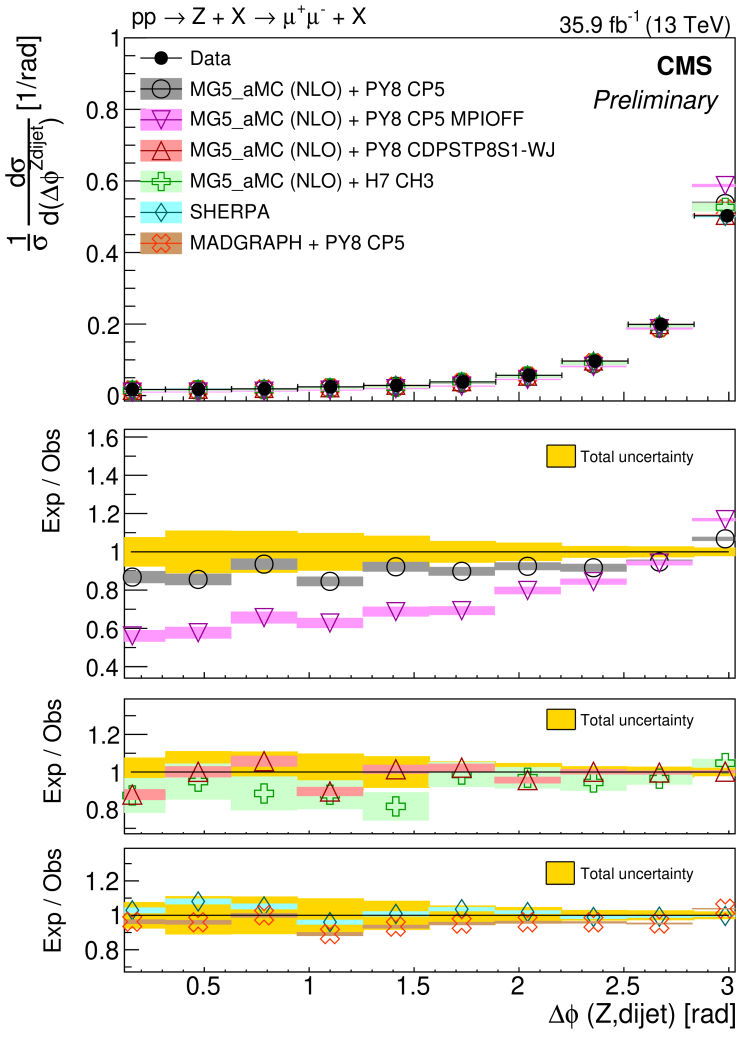
- Fiducial selections at particle level
  - Dressed muons: $p_T > 27$ GeV, $|\eta| < 2.4$
  - Z boson: $70 < M_{\mu\mu} < 110$ GeV
  - At least 1 (Z+≥1 jet) or 2 jets (Z+≥2 jets): $p_T > 20$ GeV, $|\eta| < 2.4$

Cross-section (pb)		$Z + \geq 1$ Jets	$Z + \geq 2$ Jets
Measurement		158.5 ± 0.3 (stat) ± 7.0 (syst) ± 1.2 (theo) ± 4.0 (lumi) pb	44.8 ± 0.4 (stat) ± 3.7 (syst) ± 0.5 (theo) ± 1.1 (lumi) pb
MG5_aMC (NLO)	PYTHIA 8, CP5 tune	167.4 ± 9.7	47.0 ± 3.9
	PYTHIA 8, CDPSTP8S1-WJ tune	178.4 ± 0.3	50.5 ± 0.2
	HERWIG 7, CH3 tune	158.3 ± 1.1	44.4 ± 0.6
MADGRAPH + PYTHIA 8, CP5 tune (LO)		161.2 ± 0.1	45.3 ± 0.1
SHERPA (NLO+LO)		149.8 ± 0.2	41.6 ± 0.1

- Measured cross sections are compared to MG5+Pythia8 with different tunes and with other hadronization models (HERWIG7 and SHERPA)
- Predictions with DPS-specific tune (CDPSTP8S1-WJ) is ~10% higher than data

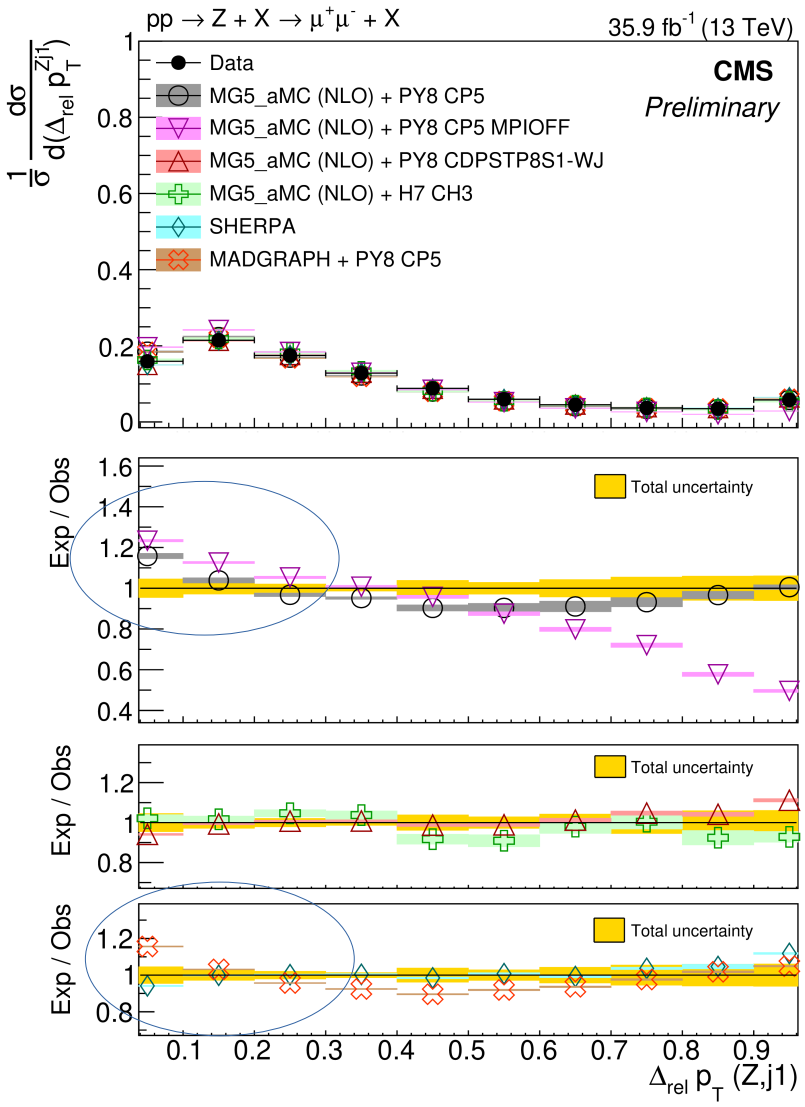
Normalized differential cross sections: ΔΦ

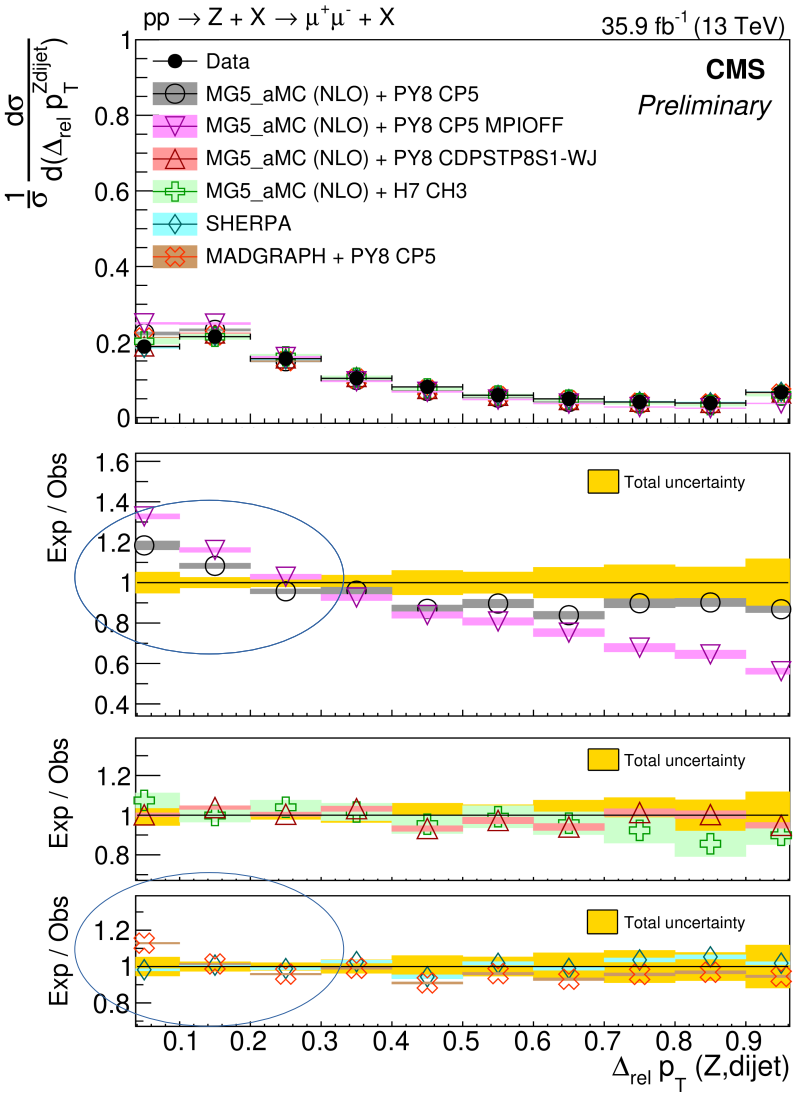




- Normalized differential cross sections are shown
- Predicted cross sections with multi-parton interaction (MPI) turned off are significantly lower than data
- Generally, predictions with MPI describe data shape well (consistent within uncertainty)

Normalized differential cross section: $\Delta_{rel}p_T$

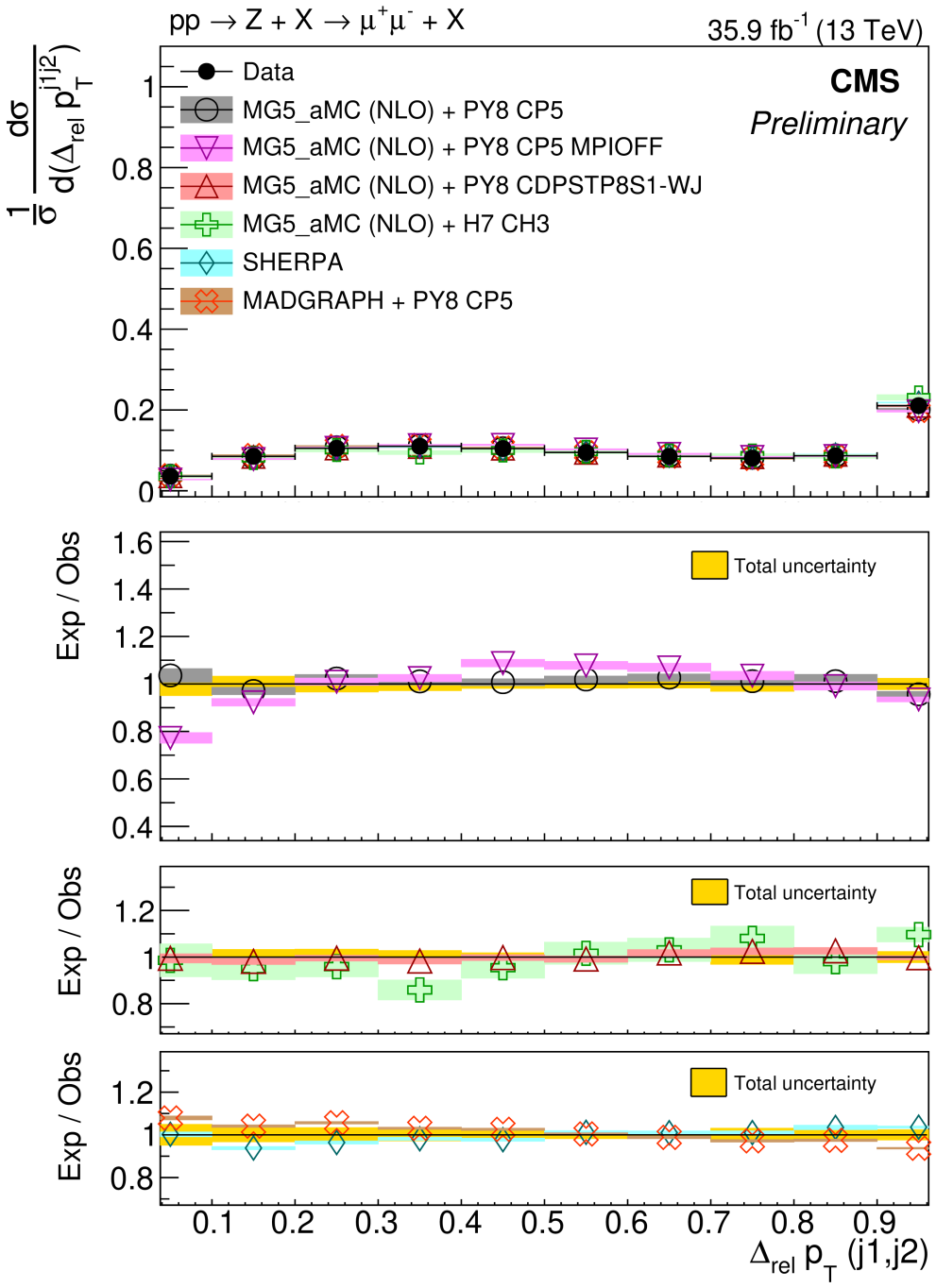




- MG5_aMC using CP5 tunes disagree with data at low $\Delta_{rel}p_T$
- SHERPA and MG5_aMC (NLO)+H7 using CH3 tune are consistent with data

Normalized differential cross section: $\Delta_{rel}p_T(j1, j2)$

- Data are compared with MG5_aMC with different tunes and SHERPA
- Various predictions agree with data as long as MPI is included



arXiv:2012.09254

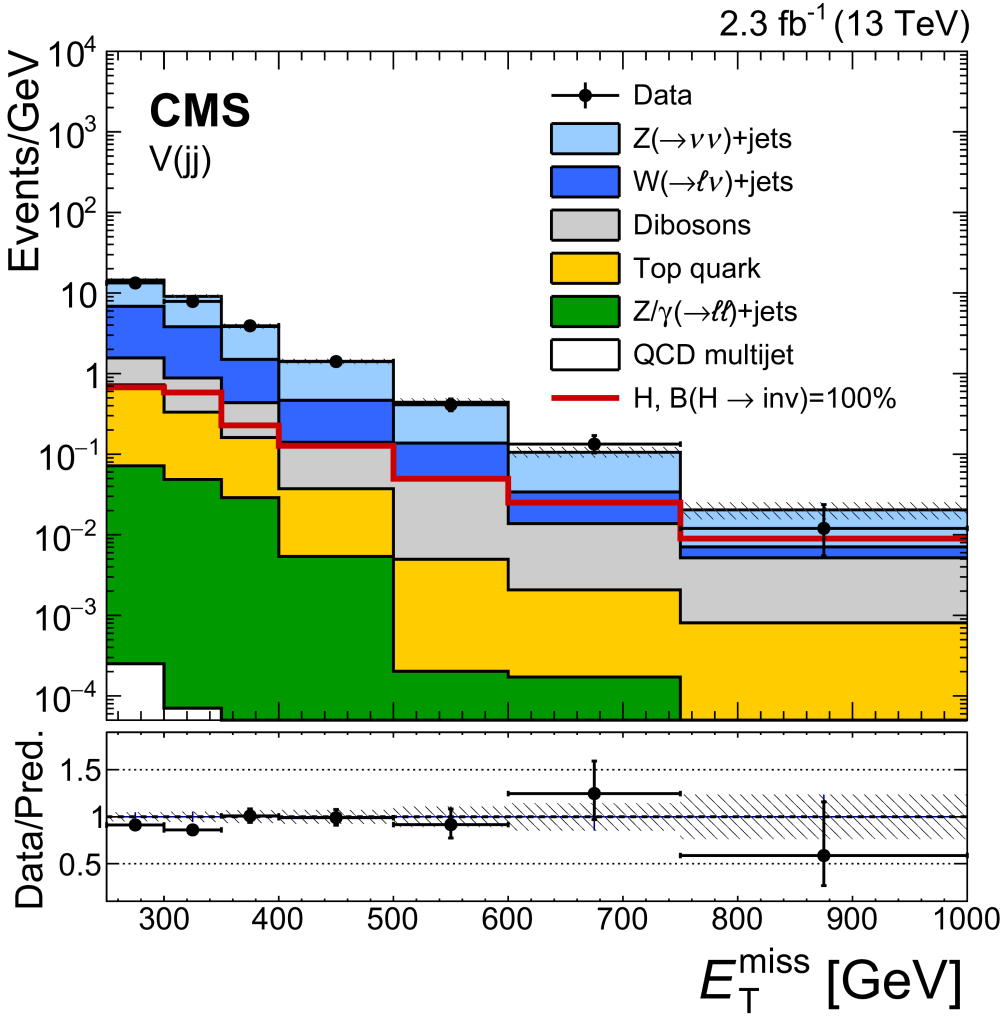
Measurement of the Z boson differential production cross section using its invisible decay mode ($Z \to \nu\bar{\nu}$) in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

Z(→νν) +jets measurement

- Complementary cross section measurements at high Z p_T for those based on leptonic decay channels, JHEP 12 (2019) 061
  - $Z \to \nu\nu$ branching ratio is higher than that of leptonic decays => improve statistical uncertainty
  - Missing transverse momentum (p_T^{miss}) is measured with sufficient accuracy at large Z p_T => reduced systematic uncertainty
- Z(→νν)+jets is significant background contributions to BSM searches and Higgs measurements
- Cross sections from $Z \to \nu\nu$ and $Z \to \ell\ell$ measurements are combined

Search for SM invisible Higgs decays

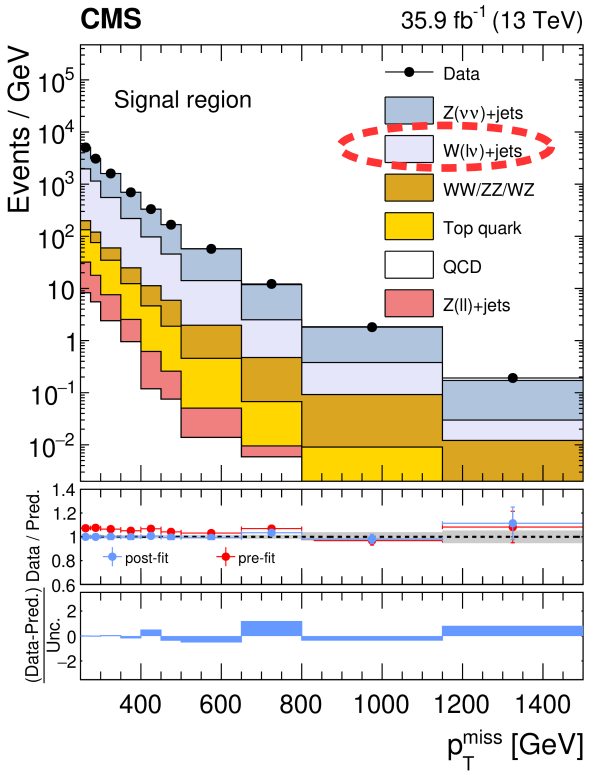
JHEP 02 (2017) 135



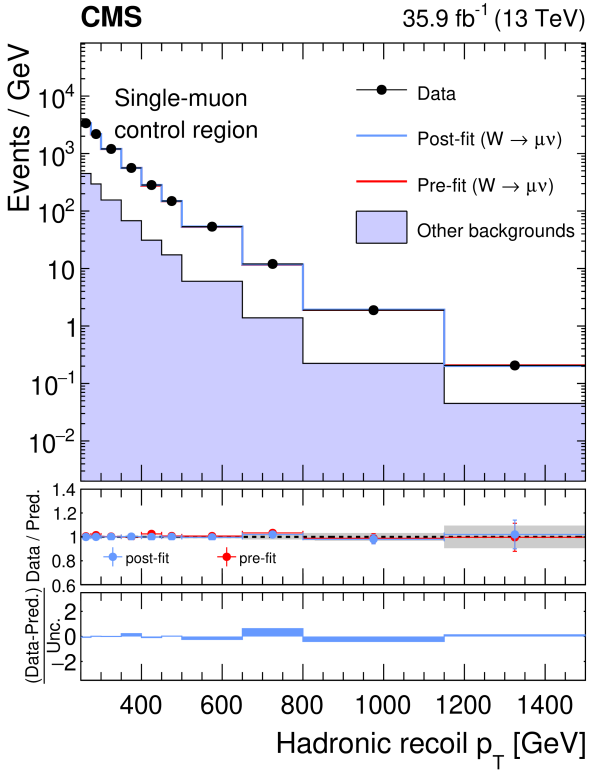
Z(→νν)+jets: Backgrounds

- Challenging measurement with a large background from W→ℓν (40% and 10% at low and high p_T^{miss} , respectively)
- Use W→ℓν control regions to handle this background
- Perform a simultaneous fit in signal and W→ℓν control regions to obtain Z(→νν) signal cross section and W background rates
- QCD background is small and estimated by a data-driven method

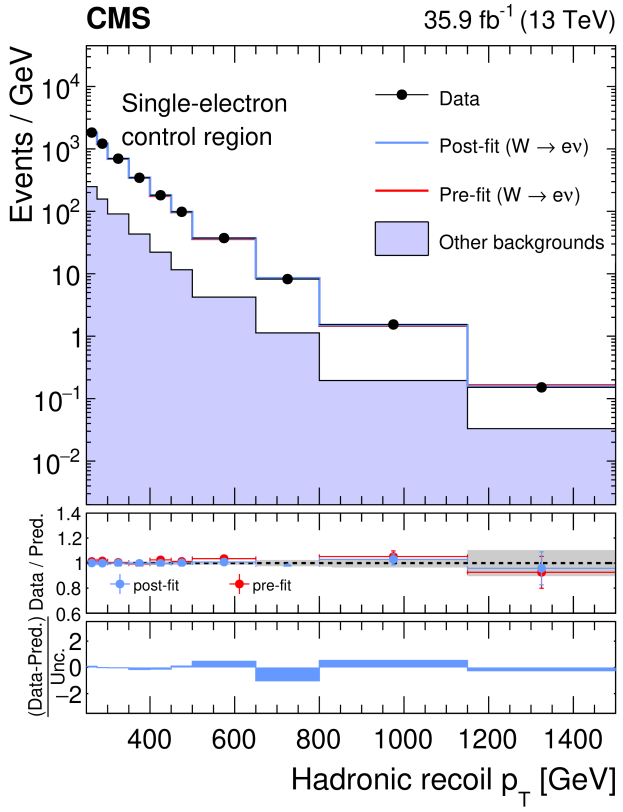
Signal region



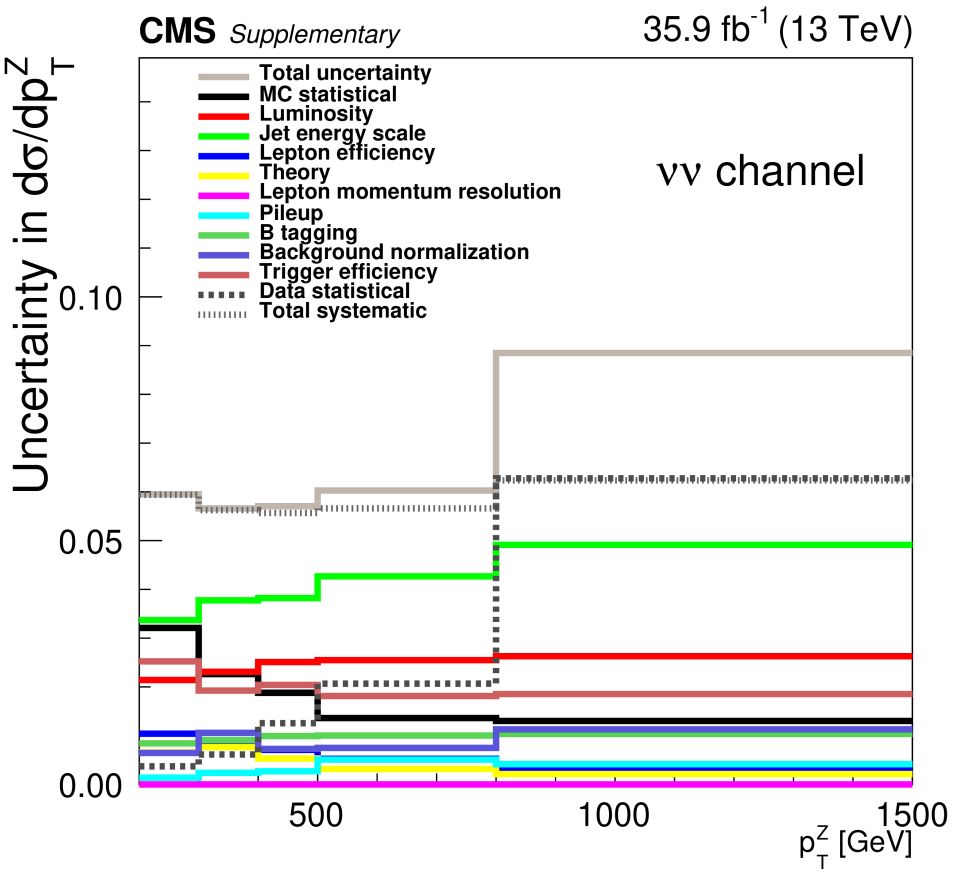
W→μν control region

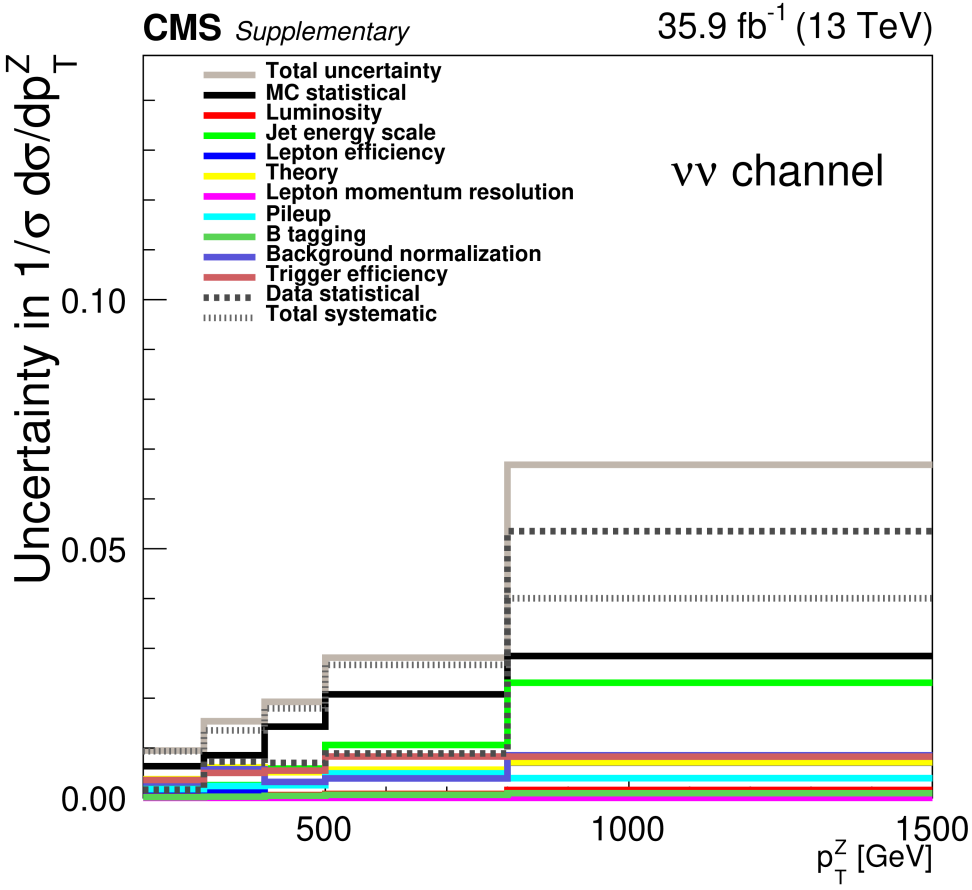


W→eν control region



Uncertainties

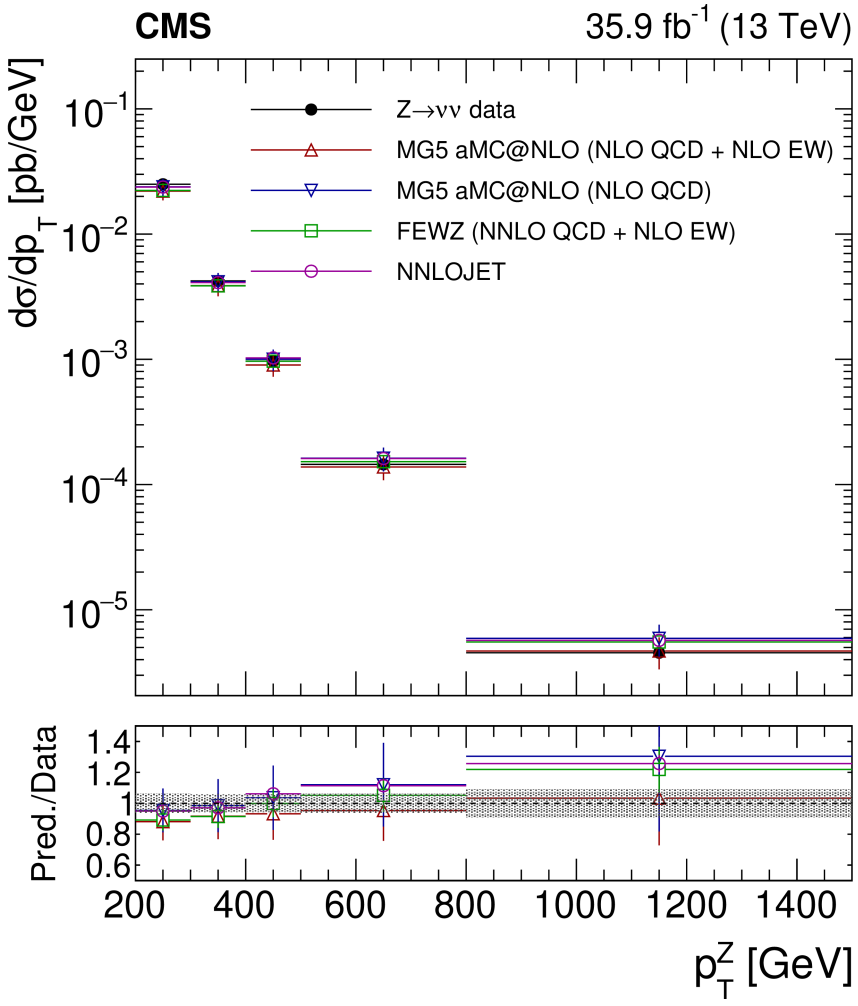




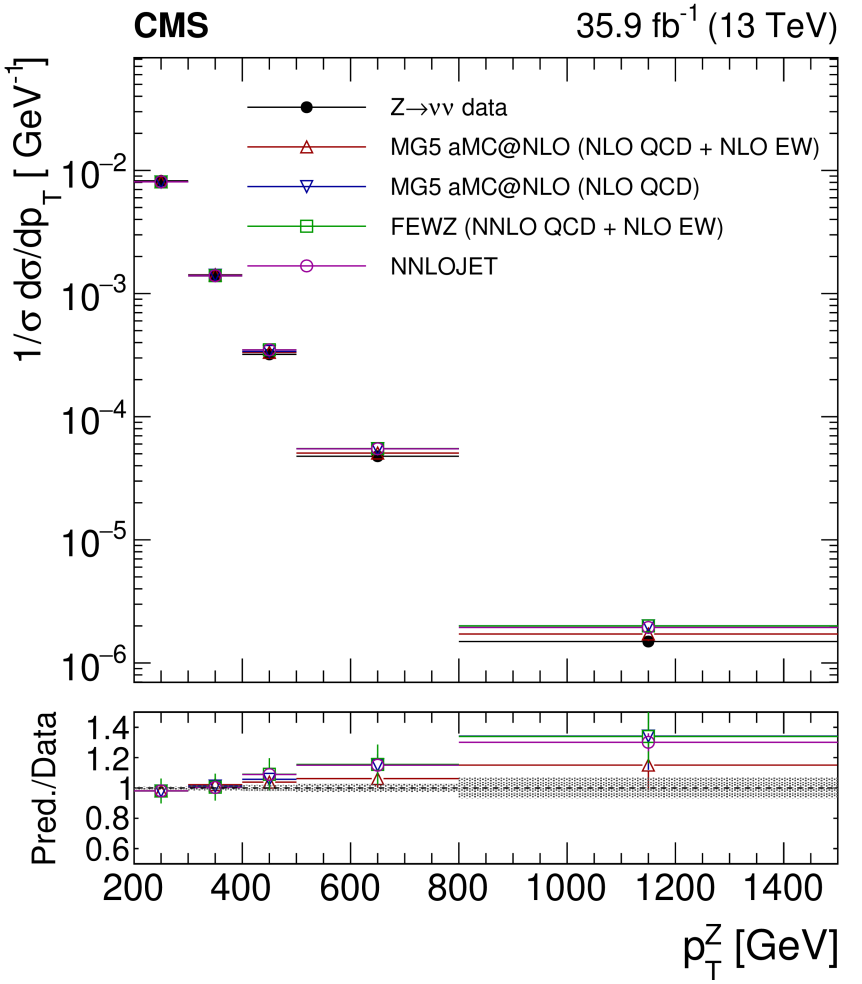
- At high Z p_T bins, uncertainties are dominated by the jet energy scale and data statistical uncertainties

Measured cross sections

Differential cross sections



Normalized differential cross sections



- MG5_aMC@NLO (NLO QCD + NLO EW) is in good agreement with data
- Other predictions are higher than data in p_T > 500 GeV

Combination of $Z \to \ell\ell$ and $Z \to \nu\nu$ channels

- Improve statistical uncertainty at high Z p_T by combining results from $Z \to \ell\ell$ and $Z \to \nu\nu$ analyses
- The fiducial regions of leptonic channels are modified
  - No requirements on particle-level jets
  - No requirements on lepton p_T and η
- Except for the luminosity, most of the systematic uncertainties and background processes are different

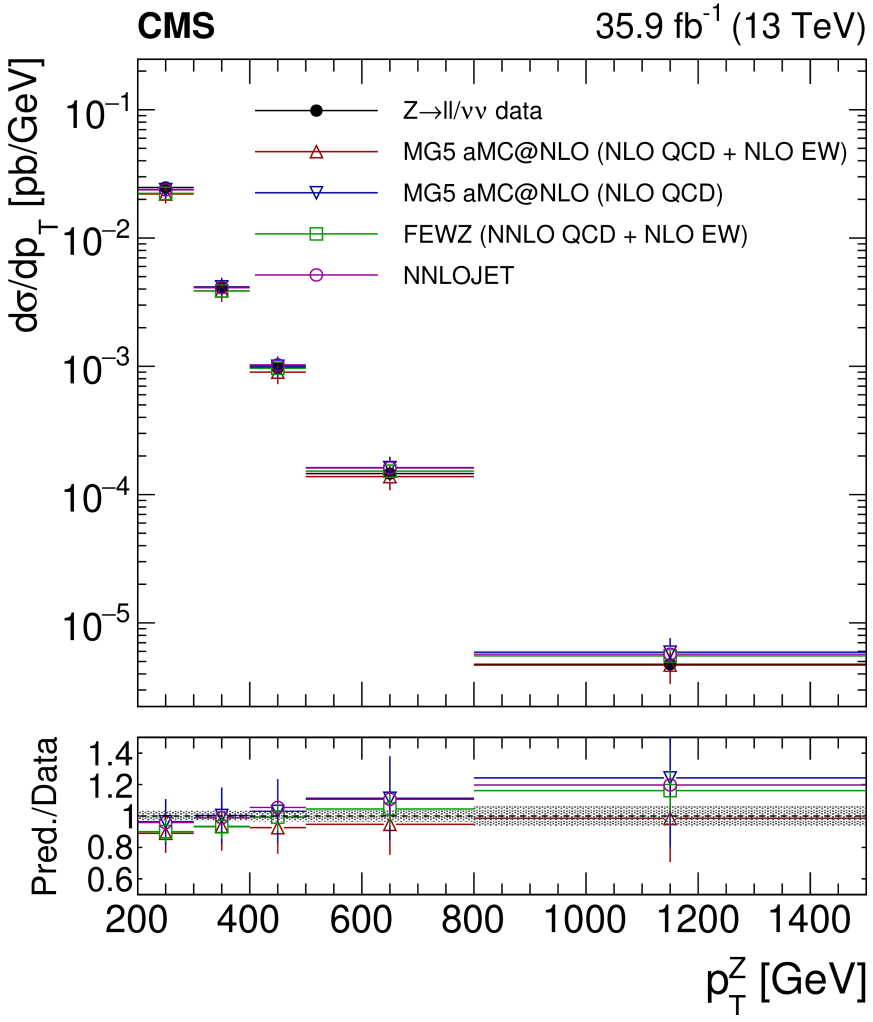
Signal strengths in Z p_T bins

p_T^Z (GeV)	$Z \to \ell\ell$	$Z \to \nu\nu$	$Z \to \ell\ell + \nu\nu$
200–300	$1.012^{+0.006}_{-0.007}$	$1.019^{+0.009}_{-0.009}$	$1.011^{+0.004}_{-0.004}$
300–400	$0.943^{+0.025}_{-0.025}$	$0.979^{+0.015}_{-0.015}$	$0.963^{+0.015}_{-0.014}$
400–500	$1.006^{+0.031}_{-0.030}$	$0.963^{+0.019}_{-0.019}$	$0.971^{+0.017}_{-0.016}$
500–800	$0.993^{+0.036}_{-0.036}$	$0.942^{+0.024}_{-0.024}$	$0.949^{+0.021}_{-0.020}$
800–1500	$1.036^{+0.099}_{-0.095}$	$0.869^{+0.059}_{-0.057}$	$0.914^{+0.052}_{-0.051}$

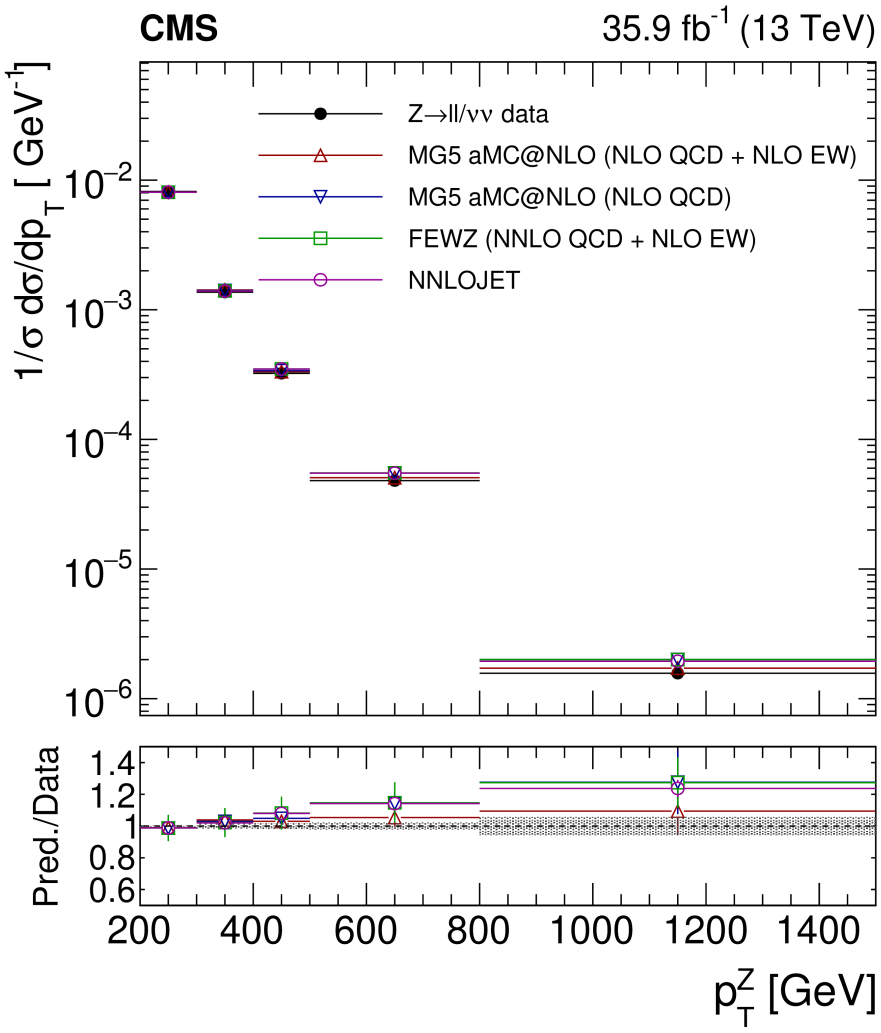
- Significant reduction in uncertainties of the combined cross sections

Combined cross sections

Differential cross sections



Normalized differential cross sections



- Predictions overestimate data at high Z p_T except MG5_aMC@NLO (NLO QCD + EW)

Phys. Rev. D 102 (2020) 032007

Measurement of the associated production of a Z boson with charm or bottom quark jets in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

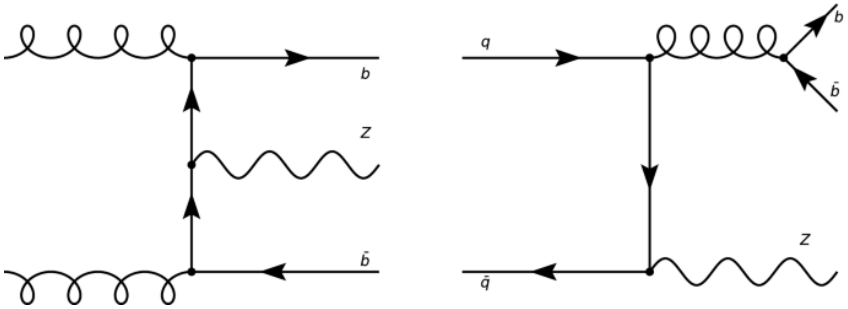
arXiv:2012.04119

Measurement of differential cross sections for Z bosons produced in association with charm jets in pp collisions at $\sqrt{s} = 13\,\text{TeV}$

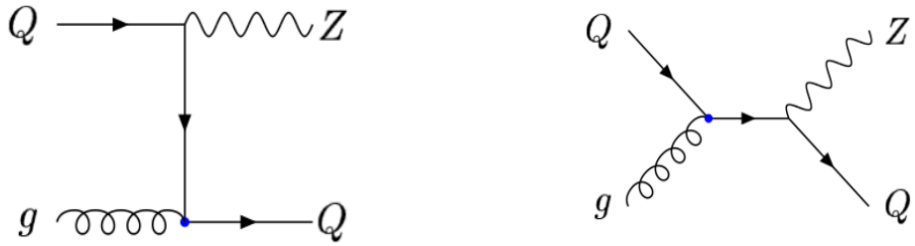
Z+HF (b, c) jets measurement

- Provide important tests of pQCD calculations
  - NLO pQCD in 4FS and 5FS schemes, NLO ME+PS
- Sensitive to the c and b contents of protons.
  - c and b partons in hadrons perturbative versus intrinsic components?
- Major background contributions for Higgs measurements and new physics searches using HF jets
- These measurements are challenging
  - Small production cross sections and high backgrounds from top quark pair production
  - Difficulties in HF tagging (machine learning techniques) and distinguishing b and c jets

4FS

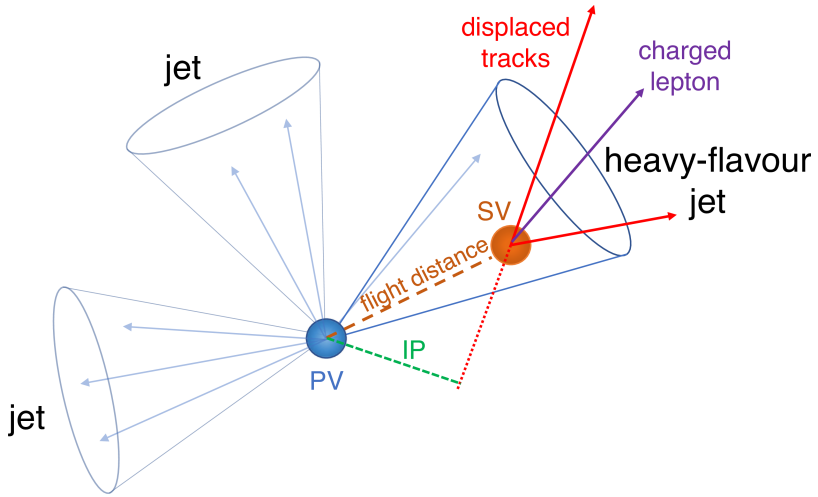


5FS

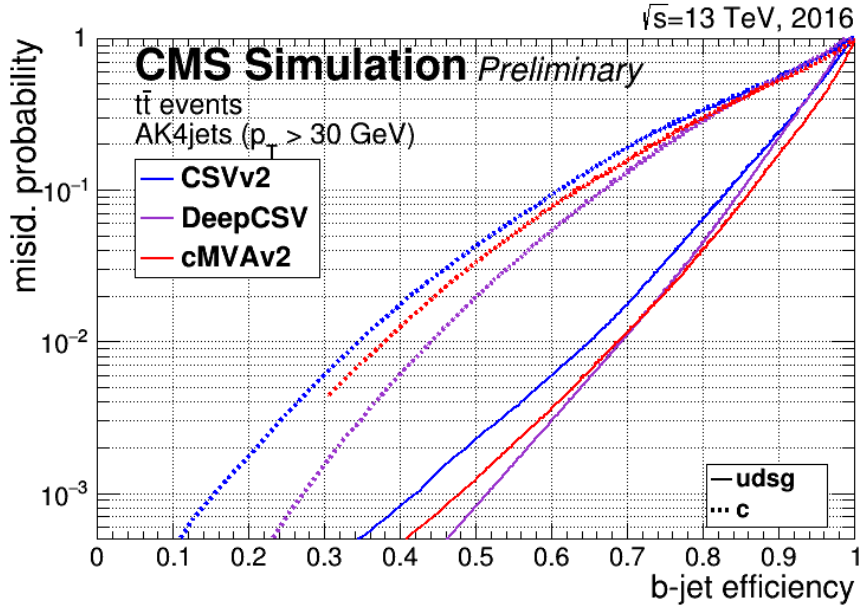


Heavy flavor tagging

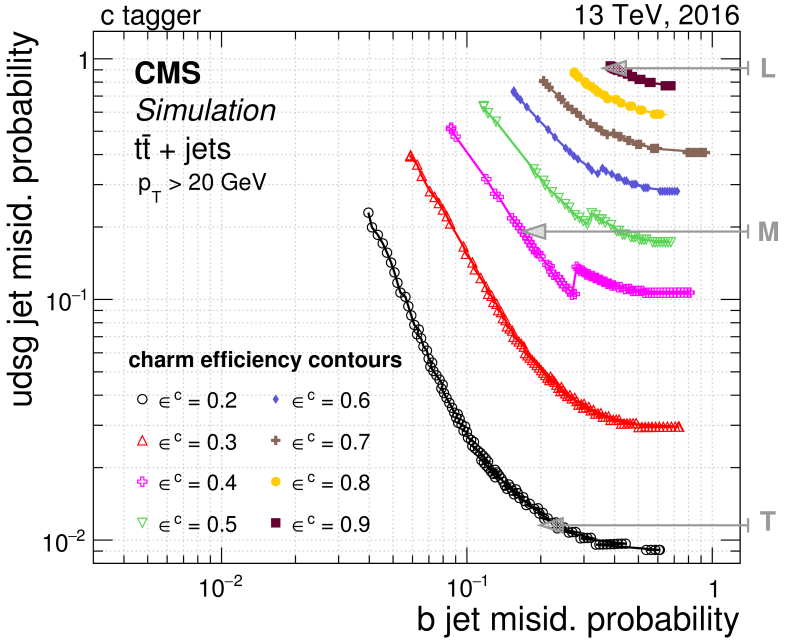
- b and c jets are tagged based on track, secondary vertex and soft lepton information
- A tagger is constructed using machine learning techniques



CMS DP-2017/012

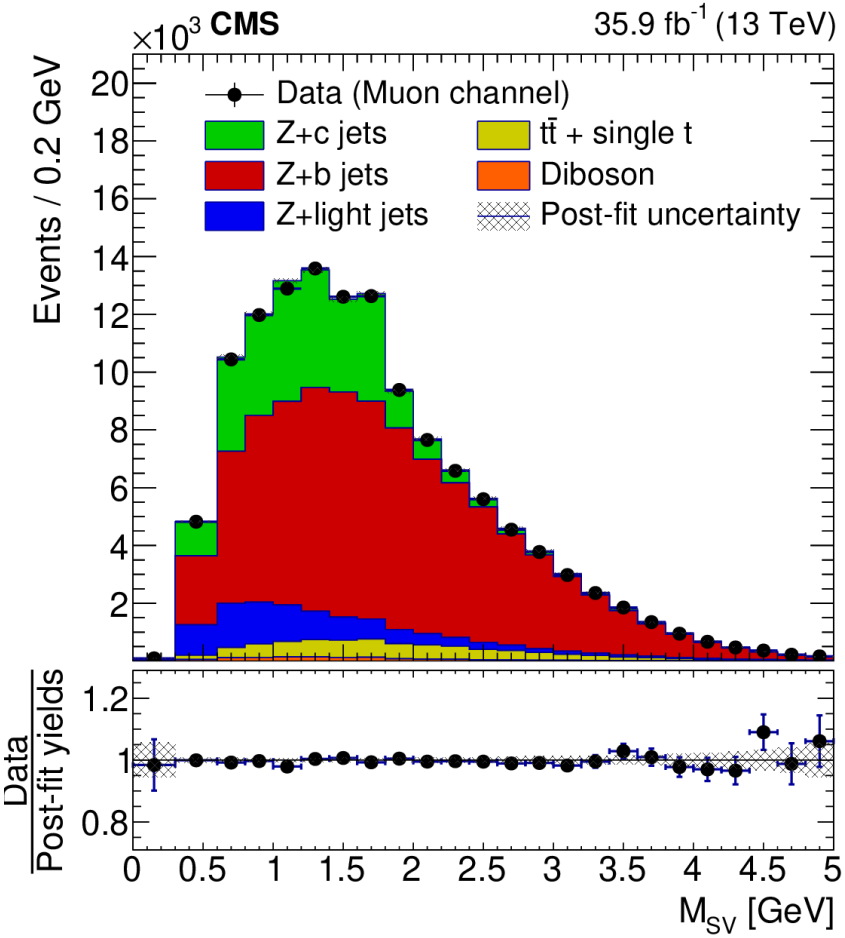


JINST 13 (2018) P05011



Measurement

- Select a sample enriched in Z+HF events by applying HF tagging
- A template fit of the secondary vertex mass (M_{SV}) is used to extract Z+b and Z+c signal yields simultaneously
- The measured cross sections are unfolded to those at the particle level:
  - Anti-k_T gen jets (R = 0.4) containing b or c hadrons
  - Dressed leptons

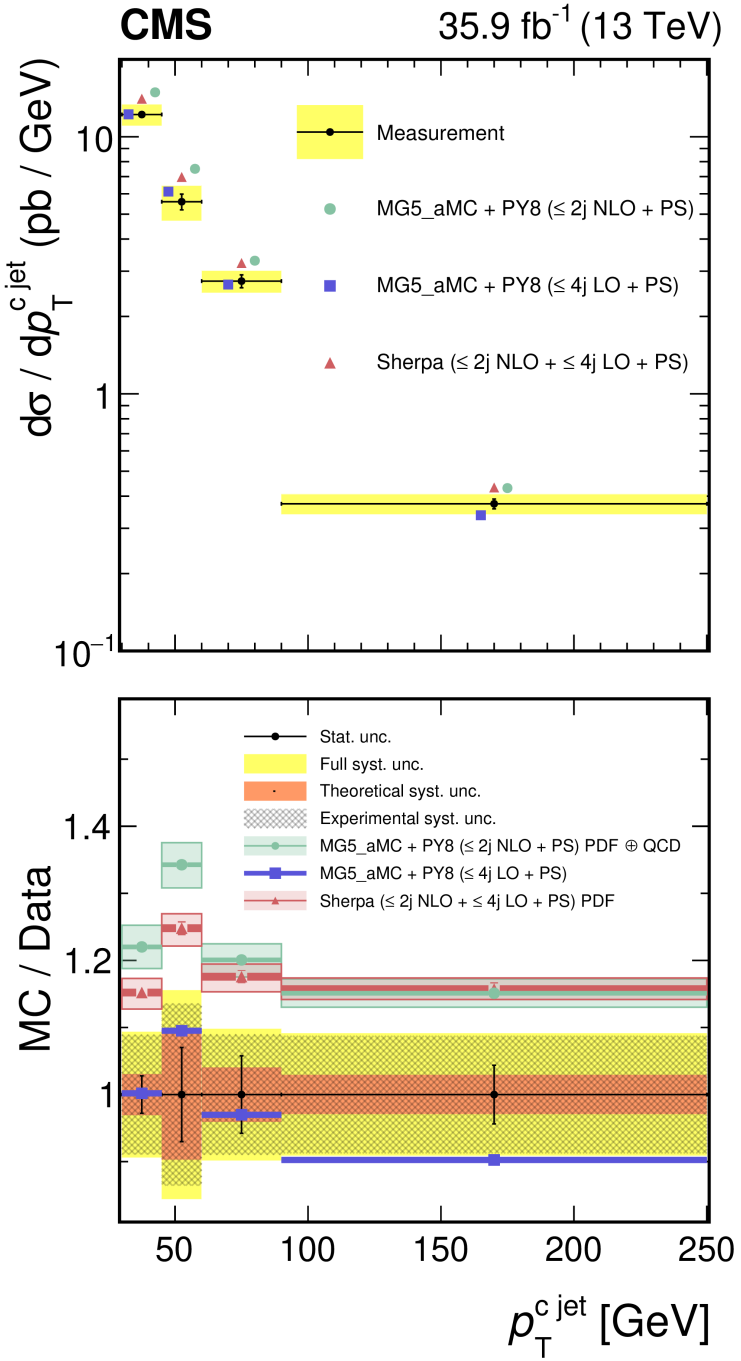


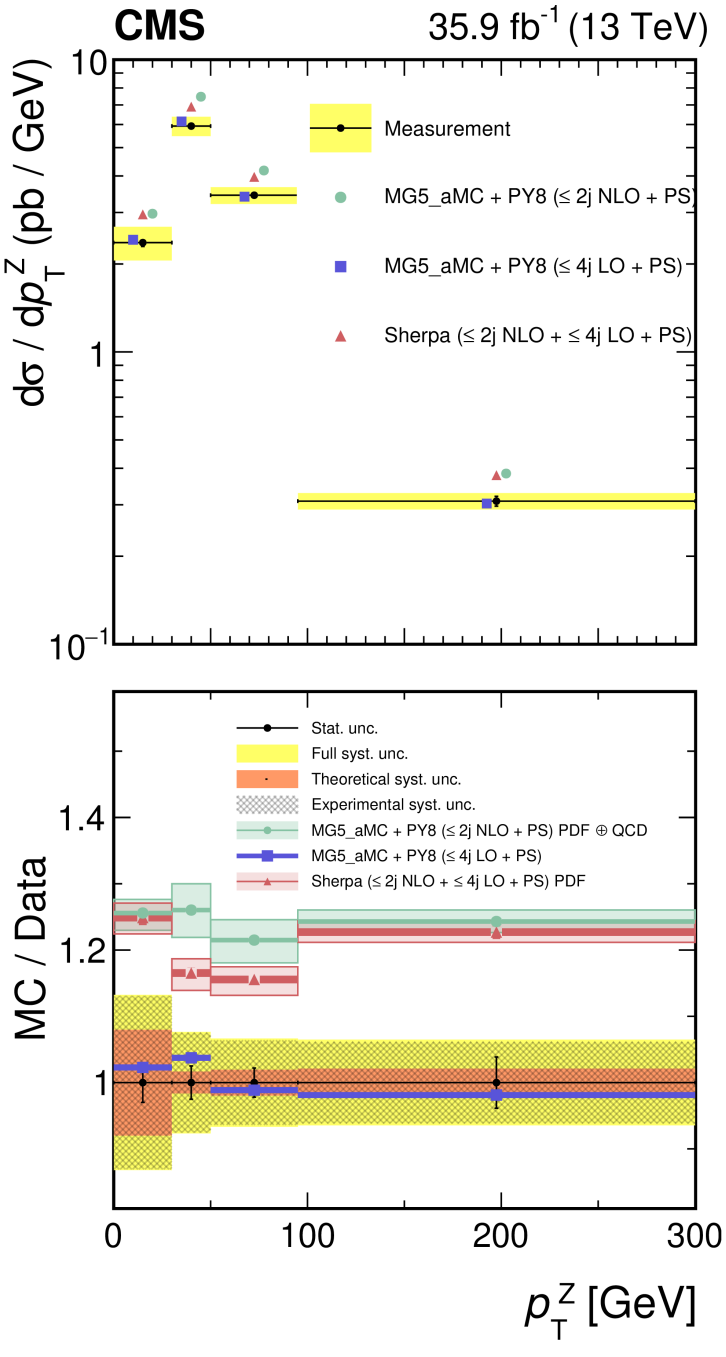
Systematic uncertainties in Z+c cross sections

Channel	QCD (%)	PDF (%)	c tag/mistag (%)	JER (%)	JES (%)	Pileup (%)	Top Pair (%)	ID\Iso (%)	$\mathcal{L}$ (%)	MC stat. (%)
$\mu\mu, p_T^{c\ jet}$	5.5	0.5	4.2	3.9	4.8	1.5	0.6	1.0	2.5	4.2
$\mu\mu, p_T^{Z}$	1.9	0.5	4.2	1.1	3.9	1.6	0.8	1.0	2.5	3.1
ee, $p_T^{c\ jet}$	6.4	0.6	4.2	3.1	6.4	3.0	0.7	2.6	2.5	6.3
ee, p_T^{Z}	2.6	0.5	4.1	1.1	4.8	1.8	0.6	2.6	2.5	3.8

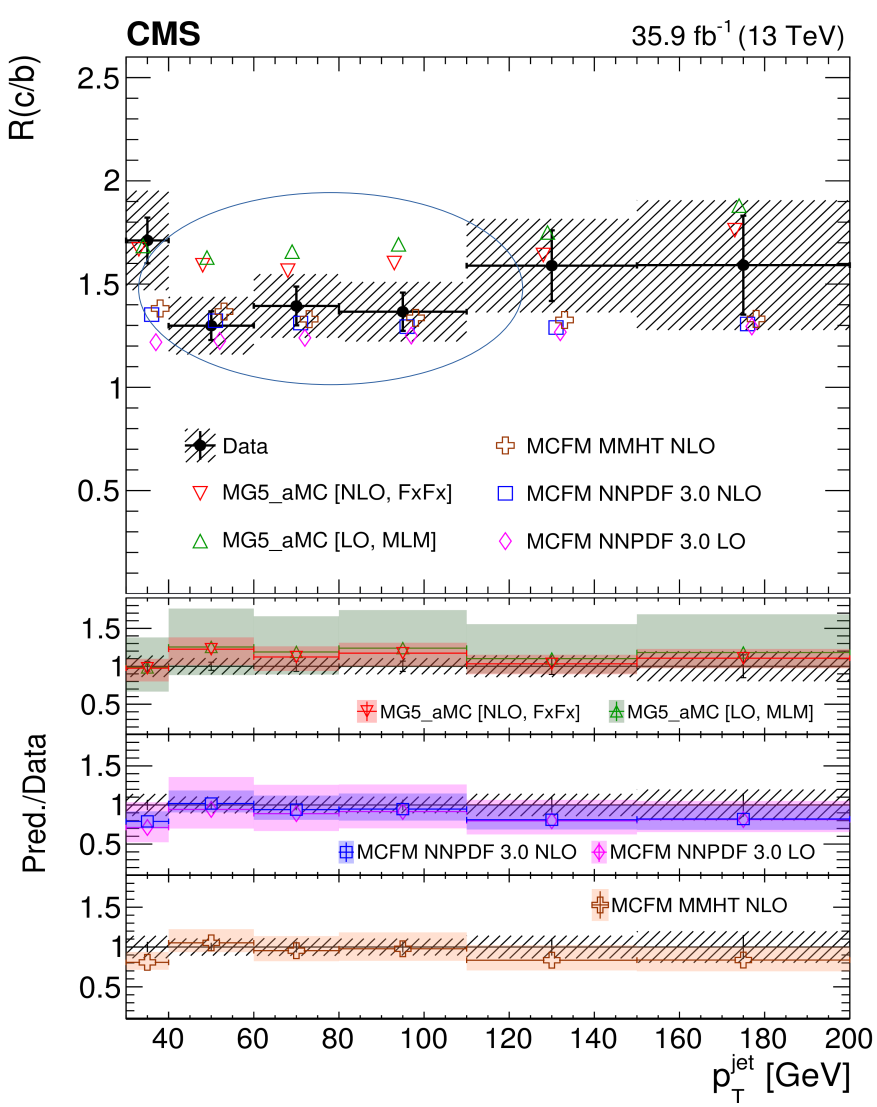
Z+c differential cross sections

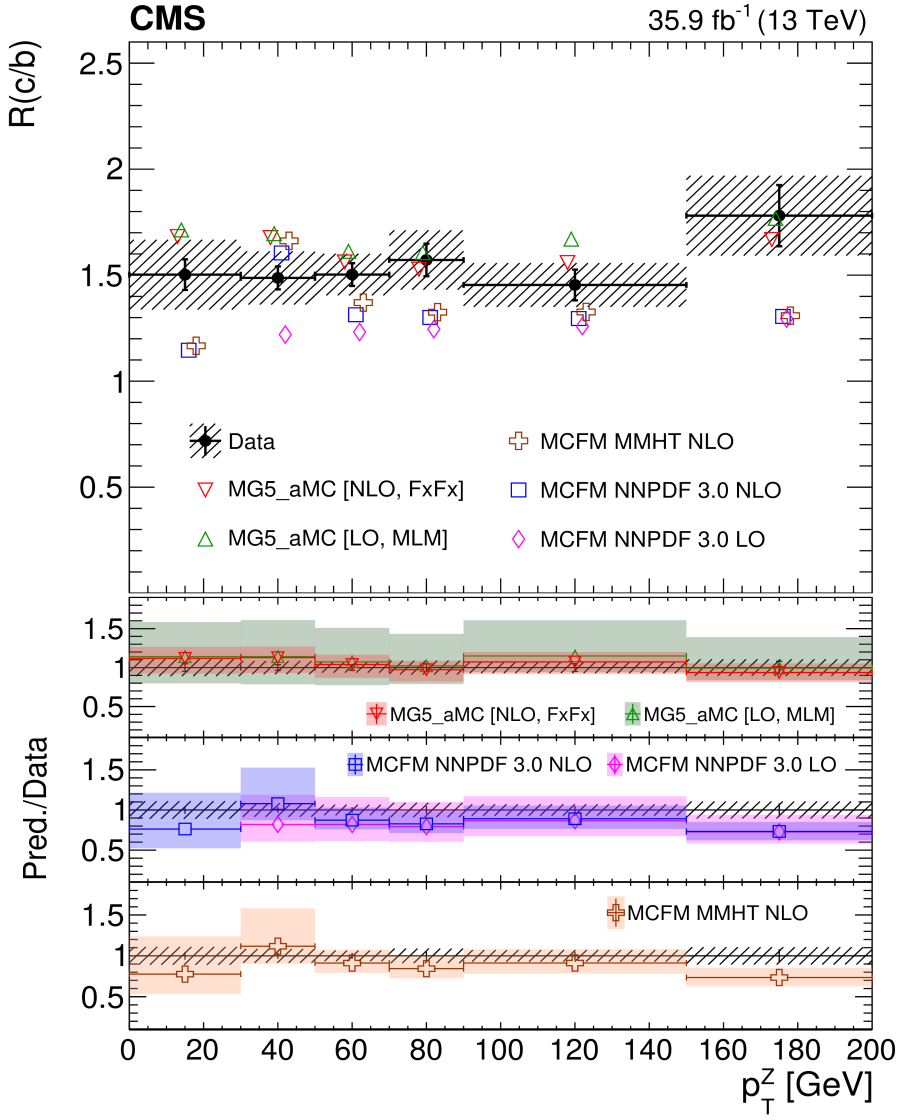
- Predictions are normalized to an inclusive NNLO Z+jets cross section (FEWZ)
- MG5 LO generally agrees with data
- Both MG5 NLO and SHERPA overestimate data





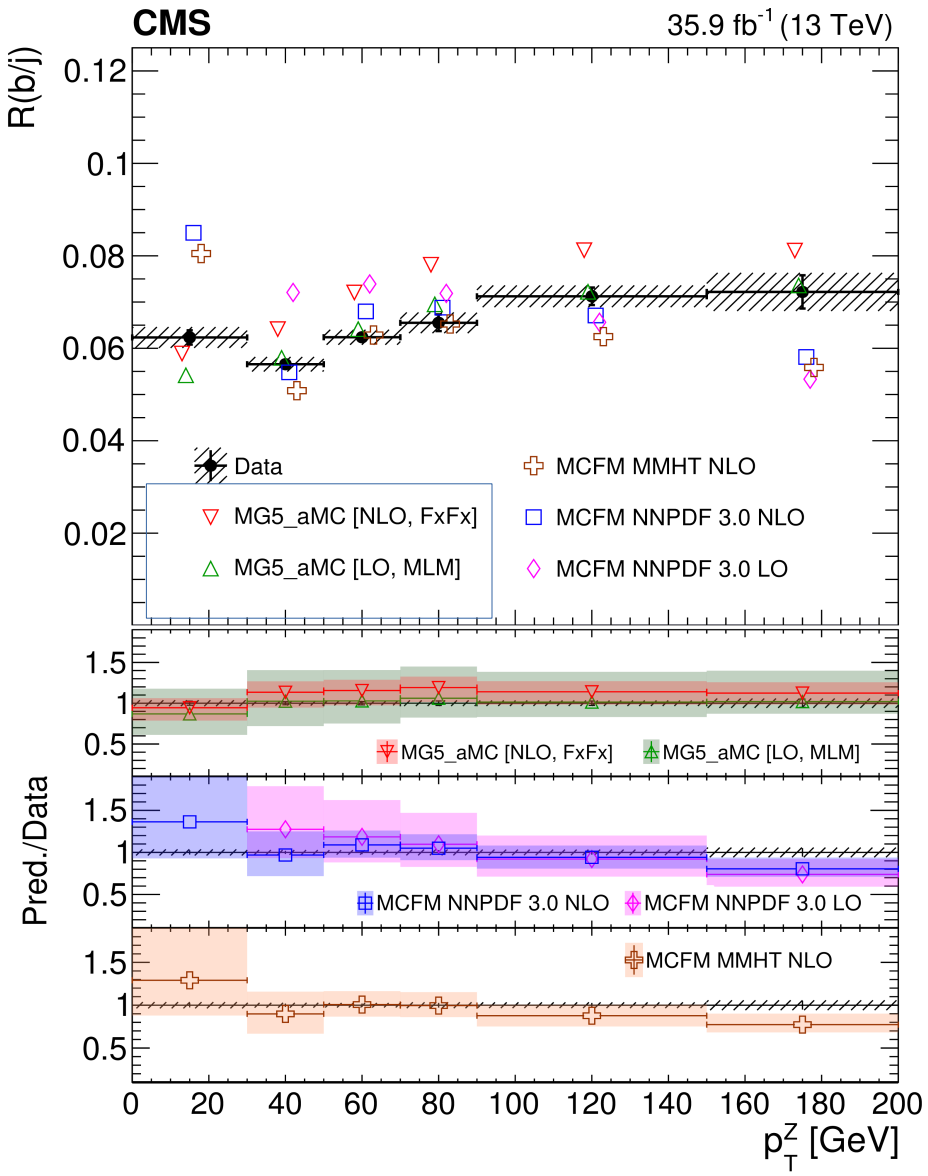
Z+c/Z+b cross section ratios

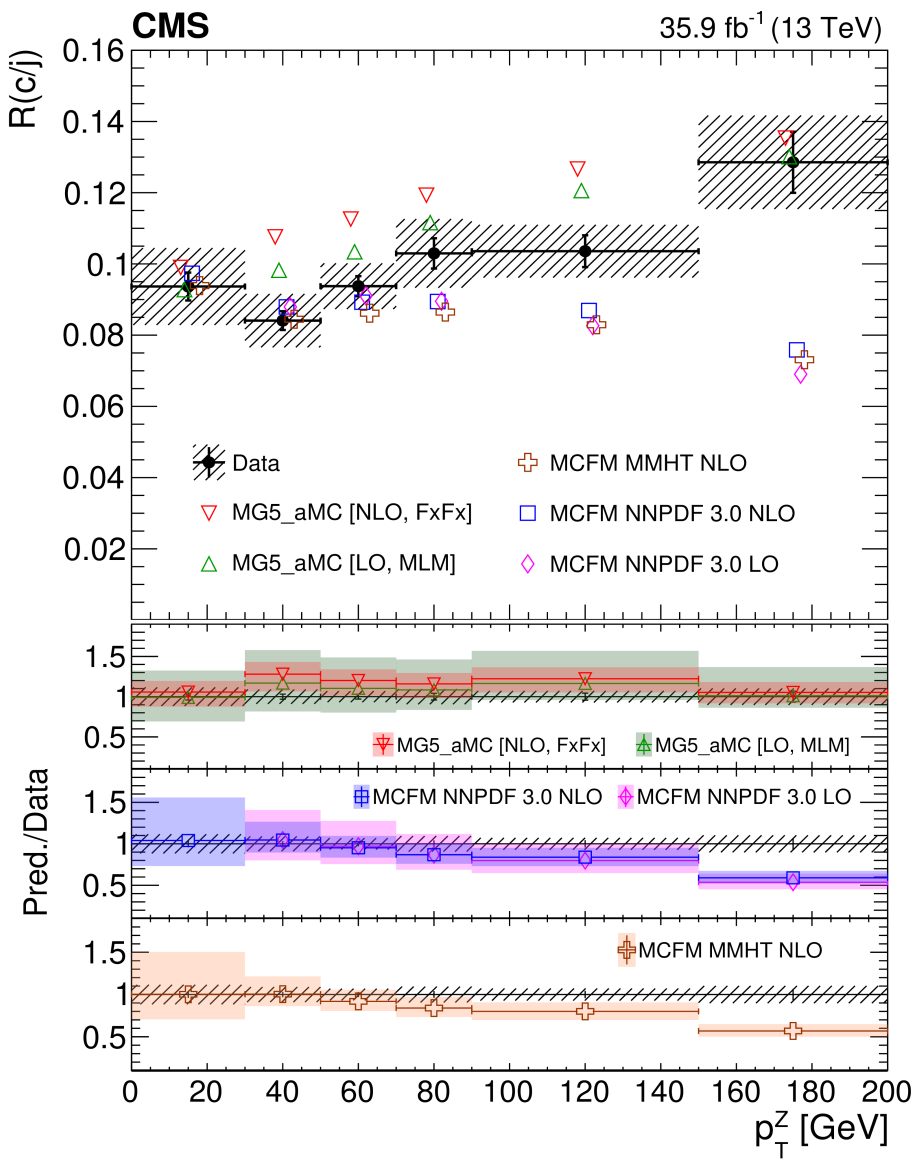




- MCFM cross section ratios are at parton level
- MG5_aMC predictions are higher than data at low jet and Z p_T
- No significant differences in MCFM ratios with choices of PDFs

Z+c/Z+j and Z+b/Z+j cross section ratios





- Predictions from MG5_aMC are consistent with data within uncertainties
- MG5_aMC LO describes data better than MG5_aMC NLO

Summary

- CMS has performed new measurements in V+jets
  - Extended to extreme phase spaces in Z+jets:

    High Z p_T and collinear Z emission regions
  - Included Z/γ cross section ratios: help to constrain backgrounds for searches and measurements in invisible decay channels
  - Studied double parton scattering using Z+≥1 jet, Z+≥2 jets
  - Obtained Z+HF (b, c) jets cross sections at 13 TeV
- Observed tensions between data and predictions in some phase space regions
- NLO predictions in Z+HF jets show disagreement with data

Stay tuned for more results with Run II data!!!

Backups

Uncertainties

- Uncertainties in differential cross sections

Observable/Uncertainty	$\Delta\phi(\mathrm{Z}, j1)$	$\Delta_{rel}p_{\mathrm{T}}(\mathrm{Z}, j1)$	$\Delta\phi(\mathrm{Z}, \mathrm{dijet})$	$\Delta_{rel}p_{\mathrm{T}}(\mathrm{Z}, \mathrm{dijet})$	$\Delta_{rel}p_{\mathrm{T}}(j1, j2)$
JES	2–7%	2–7%	5–7%	5–7%	5–6%
JER	2–5%	2–5%	2–6%	2–6%	1–3%
Pileup jet identification	1–2%	1–2%	1–2%	1–2%	1–2%
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%
Pileup modelling	0.5–1%	0.5–1%	1.5–1.7%	1.5–1.7%	1–1.2%
Model dependence	2–3%	2–3%	5–6%	5–6%	5–6%
QCD renormalization scale	2–3%	2–3%	2–3%	2–3%	2–3%
PDF	0.5–1%	0.5–1%	0.5–1%	0.5–1%	0.5–1%
Muon selection	<0.2%	<0.2%	<0.2%	<0.2%	<0.2%
Background modelling	<0.2%	<0.2%	<0.6%	<0.6%	<0.4%
Total	5–10%	5–10%	10–13%	10–13%	8–10%

- Uncertainties in normalized differential cross sections

Observable/Uncertainty	$\Delta\phi(\mathrm{Z}, j1)$	$\Delta_{rel}p_{\mathrm{T}}(\mathrm{Z}, j1)$	$\Delta\phi(\mathrm{Z}, \mathrm{dijet})$	$\Delta_{rel}p_{\mathrm{T}}(\mathrm{Z}, \mathrm{dijet})$	$\Delta_{rel}p_{\mathrm{T}}(j1, j2)$
JES	1–3%	1–3%	2–5%	1–4%	1–2%
JER	1–3%	1–3%	2–4%	2–4%	1–2%
Pileup jet identification	< 1%	< 1%	< 1%	< 1%	< 1%
Pileup modelling	0.5%	0.5%	0.5–1%	1.5–1.7%	1–1.2%
Model dependence	1–2%	1–2%	3–4%	3–4%	3–4%
QCD renormalization scale	0.5–1%	0.5–1%	0.5–1%	0.5–1%	0.5–1%
PDF	< 0.2%	< 0.2%	< 0.2%	< 0.2%	< 0.2%
Muon selection	<0.2%	<0.2%	<0.2%	<0.2%	<0.2%
Backgrounds modelling	<0.1%	<0.1%	<0.2%	<0.2%	<0.2%
Total	2–5%	2–5%	5–7%	5–7%	2–4%

DRAFT

Signal extraction

- Perform a simultaneous fit in signal and W→lν control regions to obtain signal cross section and W background rates
  - Binned maximum likelihood fit based on p_T^{miss} and hadronic recoil p_T distributions

$$\mathcal{L}(\boldsymbol{\mu}^{W\to\ell\nu}, \boldsymbol{\mu}, \boldsymbol{\theta}) = \prod_i \text{Poisson}\left(d_i^W | B_i^W(\boldsymbol{\theta}) + \frac{\mu_i^{W\to\ell\nu}}{R_i^W(\boldsymbol{\theta})}\right) \times \prod_i \text{Poisson}\left(d_i | B_i(\boldsymbol{\theta}) + \mu_i^{W\to\ell\nu} + \boldsymbol{\mu} S_i(\boldsymbol{\theta})\right)$$

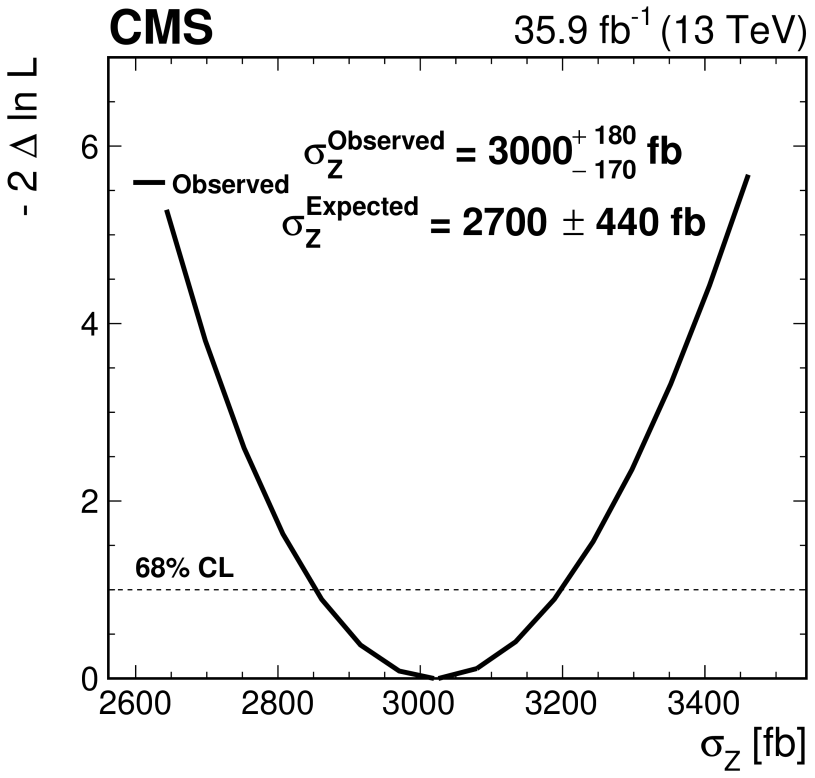
d_i d_i^W : observed data in SR and CRs

B_i B_i^W : backgrounds in SR and CRs

$\mu_i^{W\to\ell\nu}$: W(→ lν)+jets yield

R_i^W : transfer factors

$\boldsymbol{\mu}$: signal strength

S_i : signal yield

$\mu_i^{W\to\ell\nu}$ $\boldsymbol{\mu}$ are allow to freely floating in the fit