

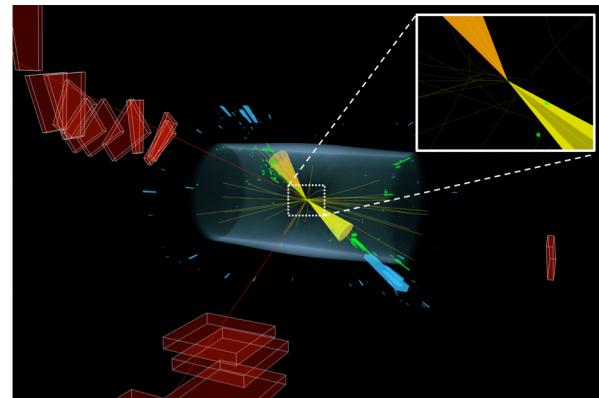
# Vector boson production with heavy flavor jets in CMS



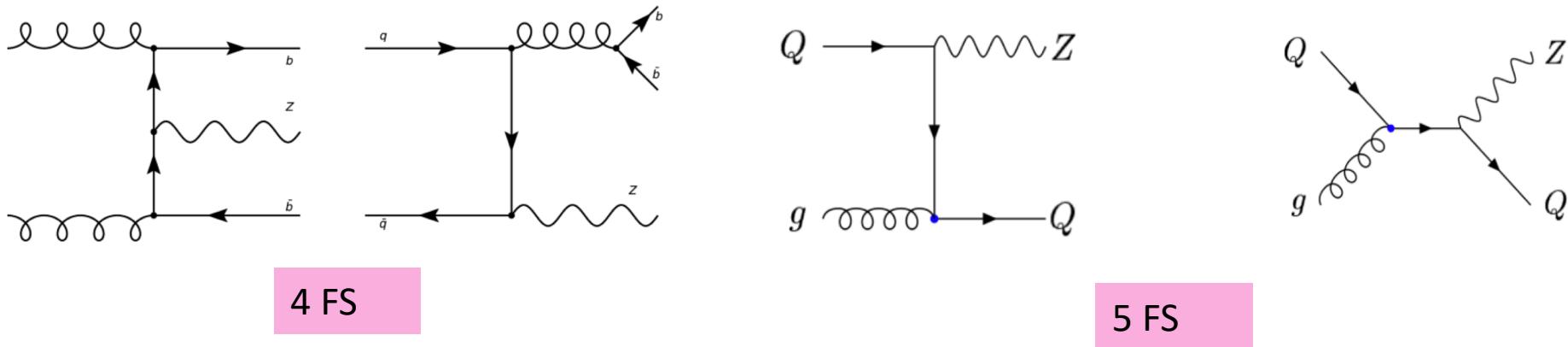
Elisabetta Gallo (DESY and UHH)  
On behalf of the CMS Collaboration



CLUSTER OF EXCELLENCE  
QUANTUM UNIVERSE



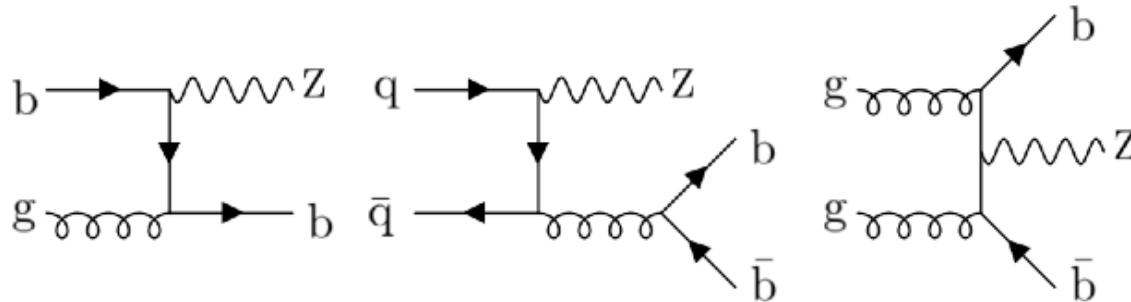
# V+HF at CMS



- Experimentally more challenging than V+light-jets due to heavy flavor ( $c,b$ ) tagging and larger background (i.e. top)
- Very important background to  $VH(bb,cc)$  production and other SM and BSM processes
- Provides crucial input for pQCD calculations (4 or 5 Flavour Scheme) and MC NLO+PS
- $W+c$  sensitive to strange PDFs,  $Z+c$  to intrinsic charm in the proton
- Here latest results from last year reported

# Z+b(b) at 13 TeV

CMS-SMP-20-015 accepted in Phys. Rev. D



- Cross sections for  $Z+ >=1, >=2$  b-jets, using all Run 2 data, combining the electron and muon Z decay channels
- Differential cross sections sensitive to  $bg$ ,  $q\bar{q}$  and  $gg$  in the initial state
- b-jet identified via a neural network algorithm (DeepCSV), with 0.1% (2-5%) misidentification for light- (c-) jets
- $Z+$ light-jets/ $Z+c$ -jets constrained by control region (CRs) + scale factors Data/MC applied to the signal region

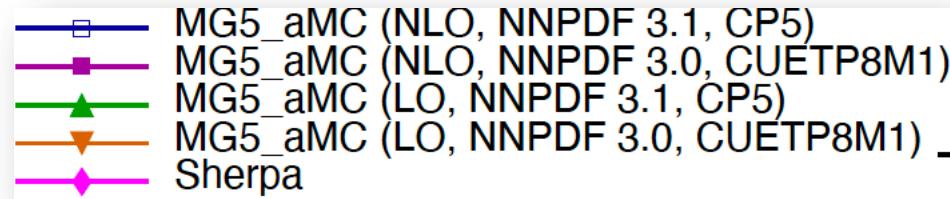
- Differential cross section unfolded in fiducial region:

Object	Selection
Dressed leptons	$p_T(\text{leading}) > 35 \text{ GeV}, p_T(\text{subleading}) > 25 \text{ GeV},  \eta  < 2.4$
Z boson	$71 < m_{\ell\ell} < 111 \text{ GeV}$
Generator-level b jet	b hadron jet, $p_T > 30 \text{ GeV},  \eta  < 2.4$

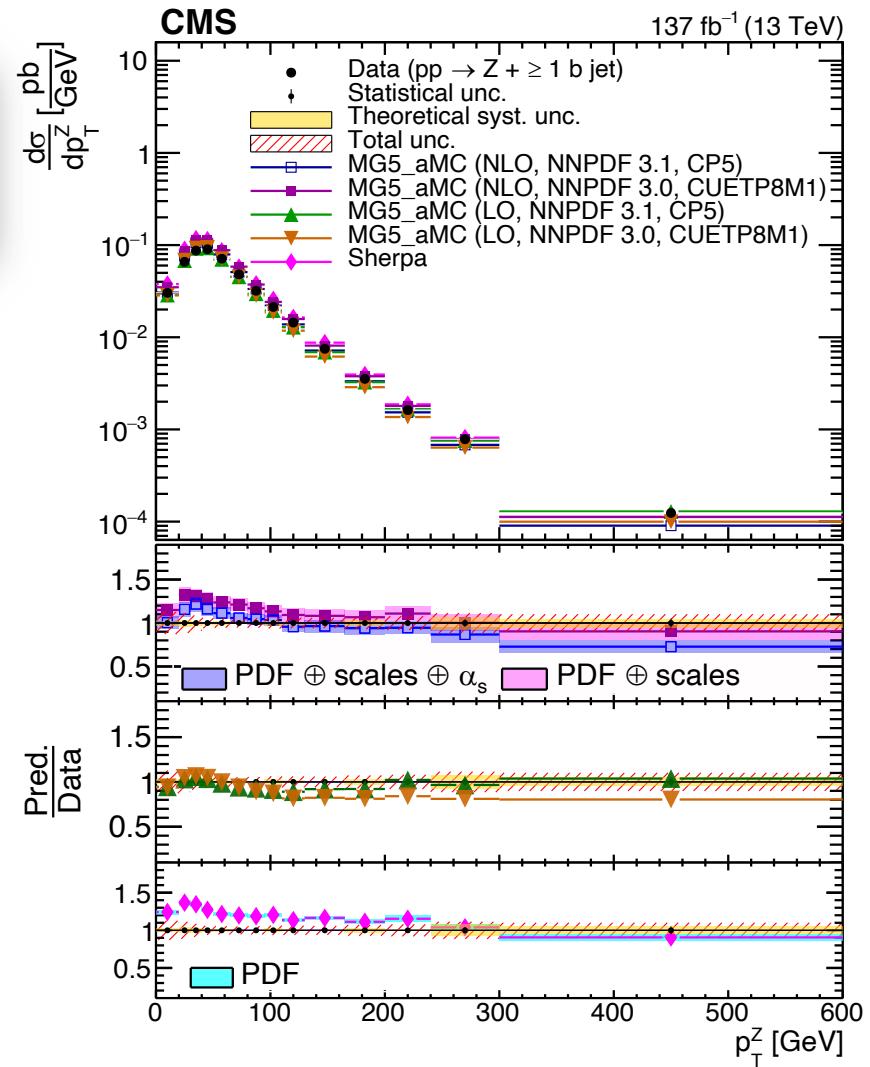
- Systematics uncertainties  $\sim 6\%$  for  $Z+>=1$  b-jet,  $\sim 10\%$  for  $Z+>= 2$  b-jets, main ones are b-tagging, JES, electron selection

# Z+b(b) at 13 TeV

CMS-SMP-20-015 accepted in Phys. Rev. D

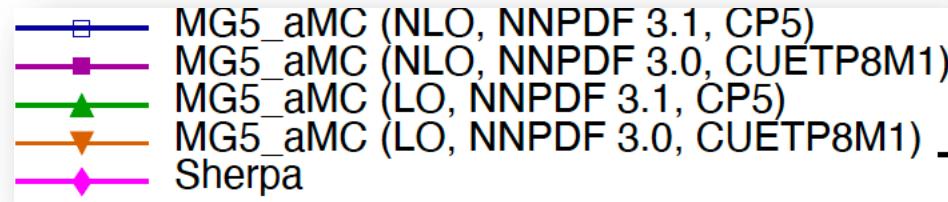


- $p_T^Z$  measured for  $Z+>=1$  b-jet
- Comparison to MG5\_aMC LO/NLO with 2 different tunes and PDF versions (CUETP8M1 and CP5)
- MG5\_aMC (MadGraph5\_amc@NLO) NLO overestimates data by ~20%, LO in good agreement (but different tunes and PDFs)
- SHERPA overestimates cross sections by up to 30%

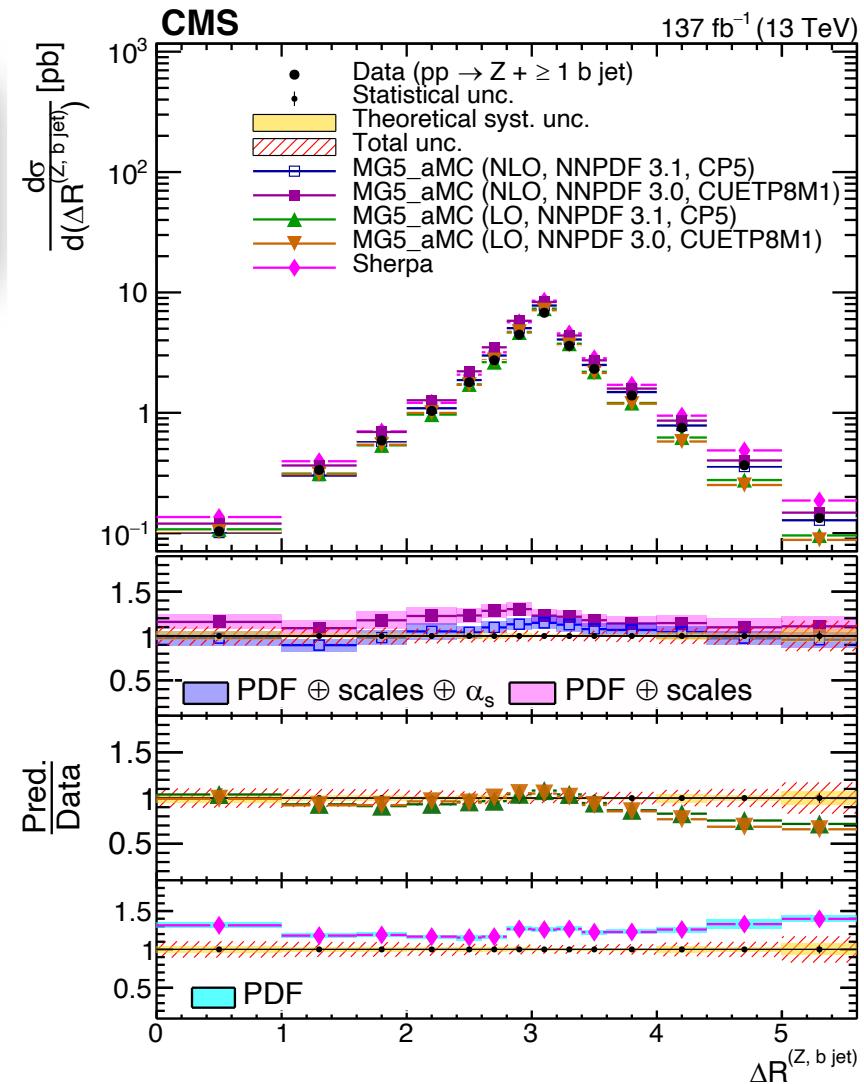


# Z+b(b) at 13 TeV

CMS-SMP-20-015 accepted in Phys. Rev. D

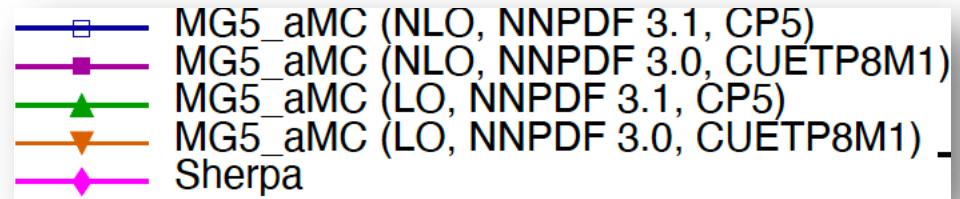


- Angular correlation between Z and the leading b-jet measured for  $Z+>=1$  b-jet
- Sensitive to pQCD effects
- MG5\_aMC LO gives the worst description, especially at high  $\Delta R$
- SHERPA gives the best description in shape

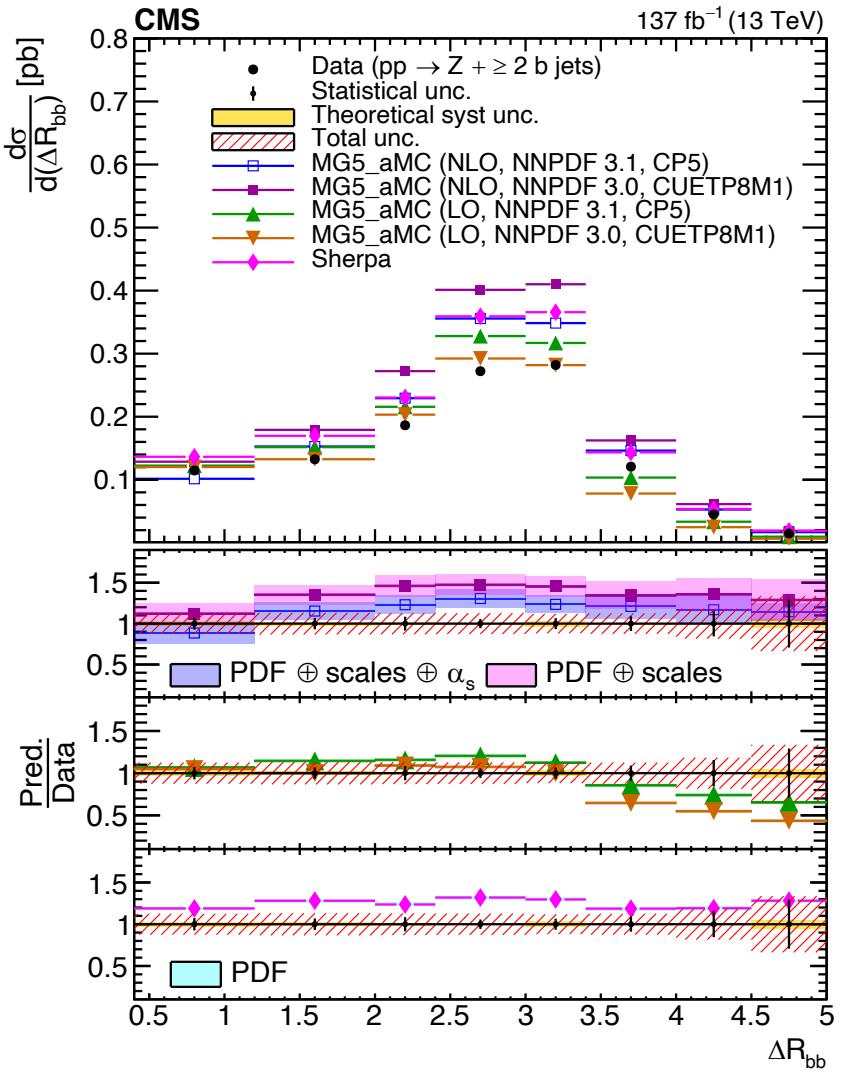


# Z+b(b) at 13 TeV

CMS-SMP-20-015 accepted in Phys. Rev. D

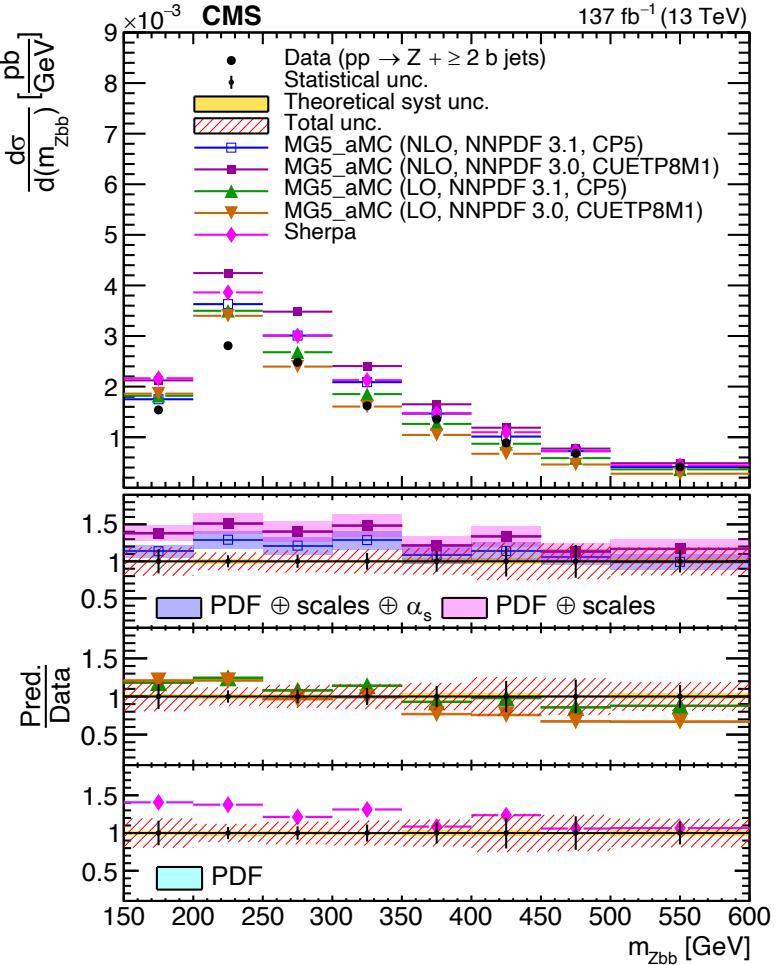
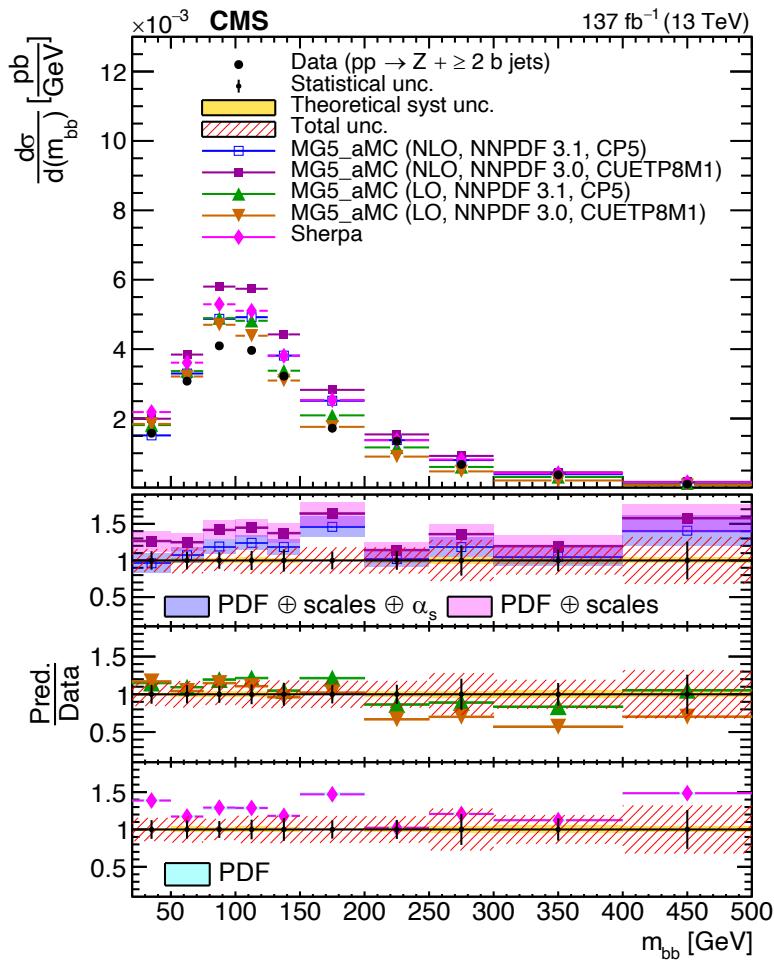


- Angular correlation between the 2 leading b-jets measured for  $Z+>=2$  b-jet events
- At small values, sensitive to gluon splitting
- MG5\_aMC NLO higher than the data, also here MG5\_aMC LO gives the worst description, especially at high  $\Delta R$
- SHERPA gives the best description in shape



# Z+b(b) at 13 TeV

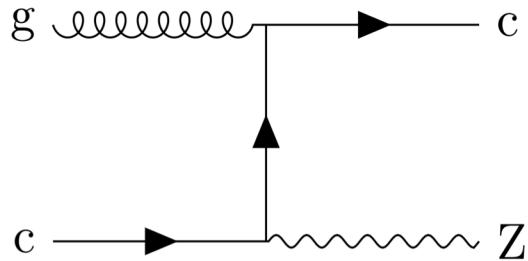
CMS-SMP-20-015 accepted in Phys. Rev. D



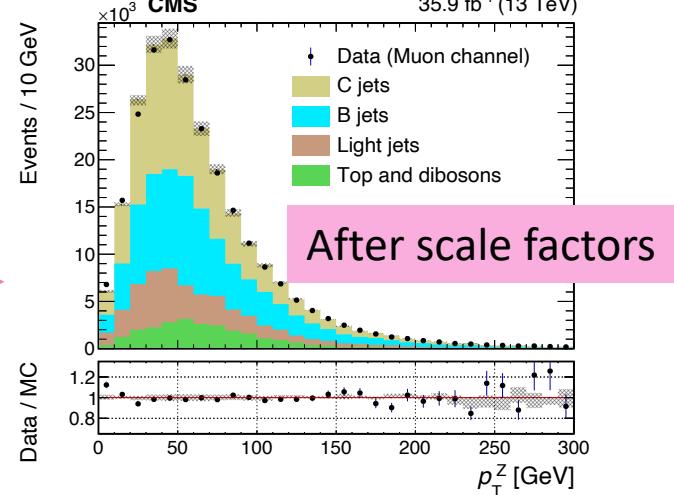
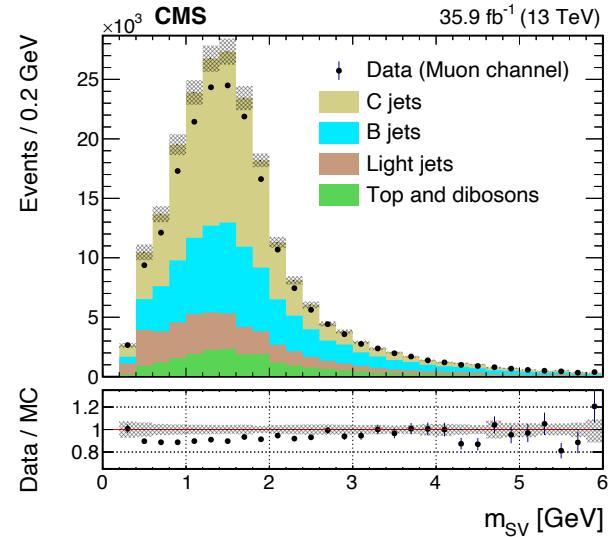
Invariant masses of bb or of Zbb system important for searches

# Z+c at 13 TeV

[JHEP 04 \(2021\) 109](#)

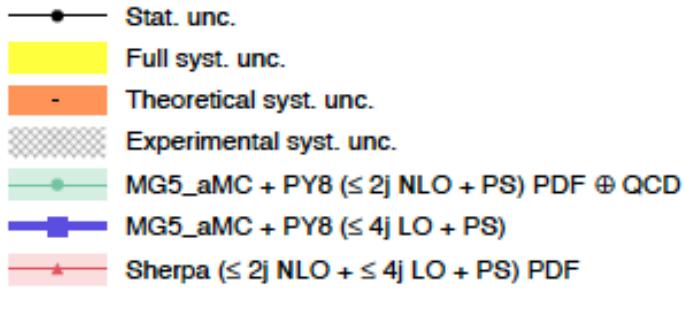


- Measurement of total and differential distributions for Z+c in fiducial region with 2016 data, in both electron and muon Z decay channels
- c-contribution obtained fitting templates of Z+c, Z+b and Z+light-jet contributions to the secondary vertex mass  $m_{SV}$  in each Z or leading c-jet  $p_T$  bin
- Scale factors applied to the contributions achieve good agreement between data and sum of contributions
- c-tagging and JES are the largest experimental uncertainties

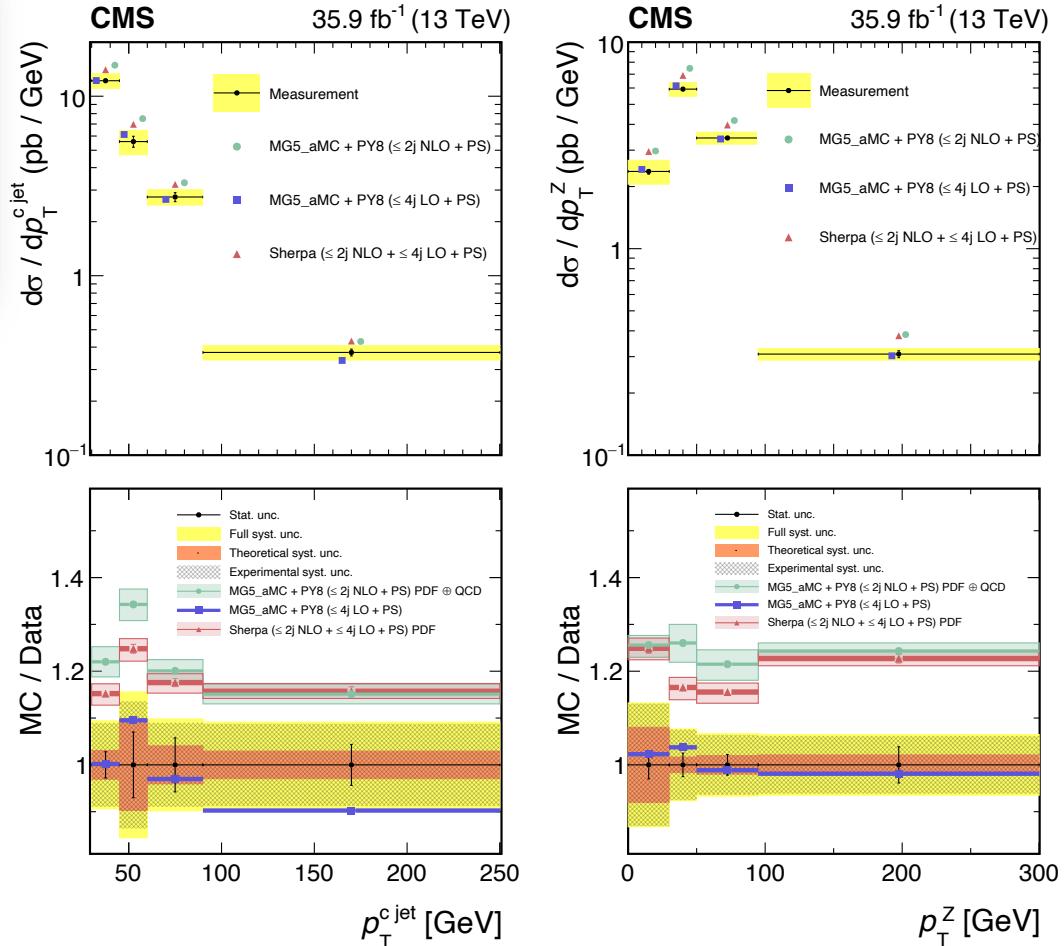


# Z+c at 13 TeV

[JHEP 04 \(2021\) 109](#)

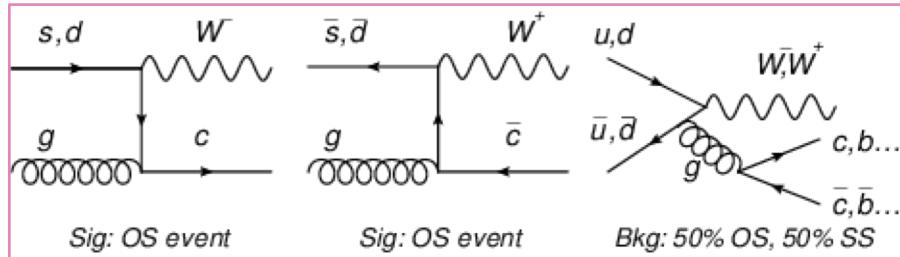


- Good agreement with MG5\_aMC LO
- SHERPA and MG5\_aMC NLO above the data
- As all predictions are normalized to FEWZ, there could be an overestimation of charm PDF

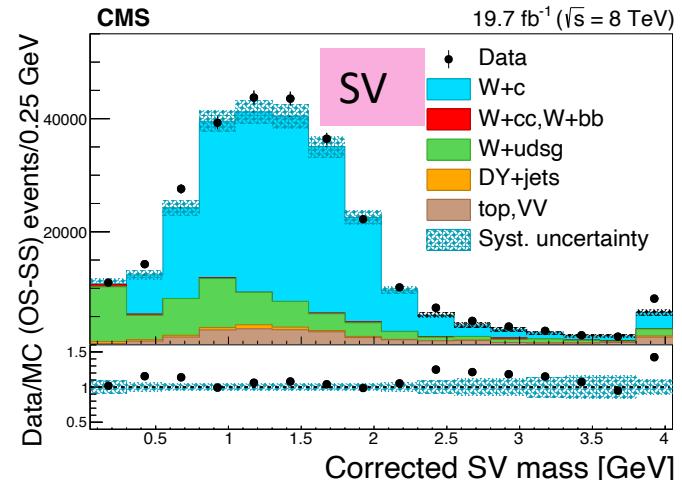
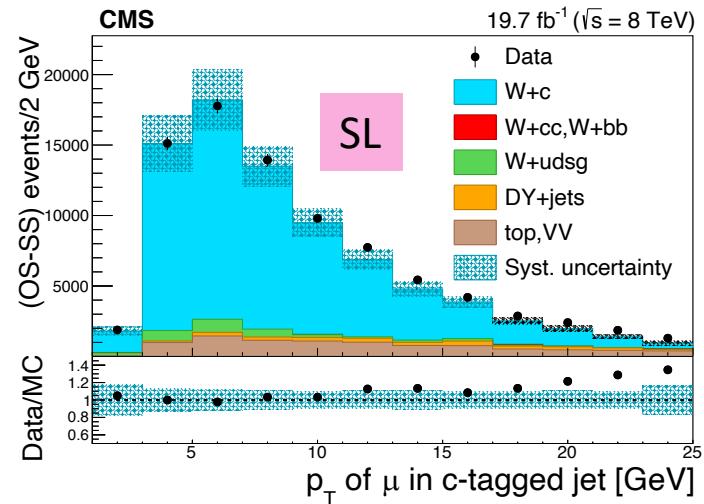


# W+c at 8 TeV

CMS-SMP-18-013



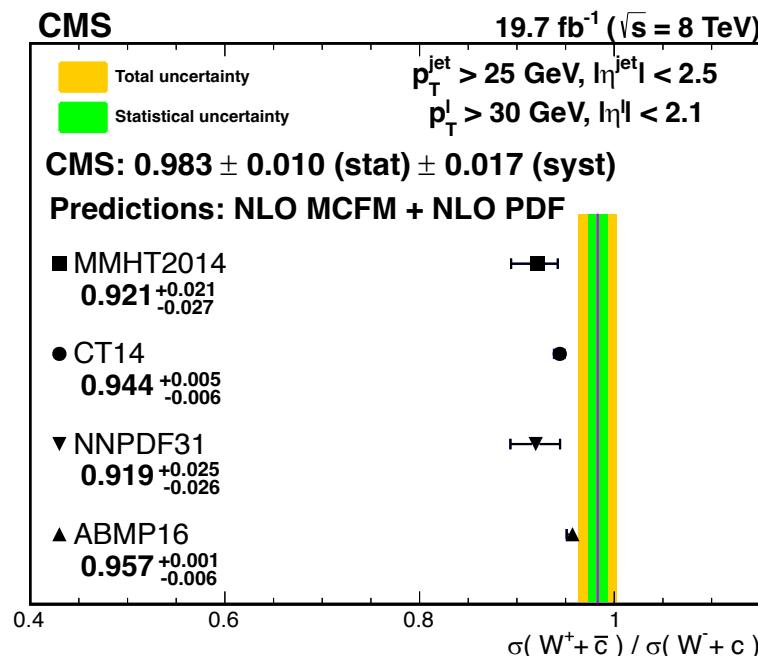
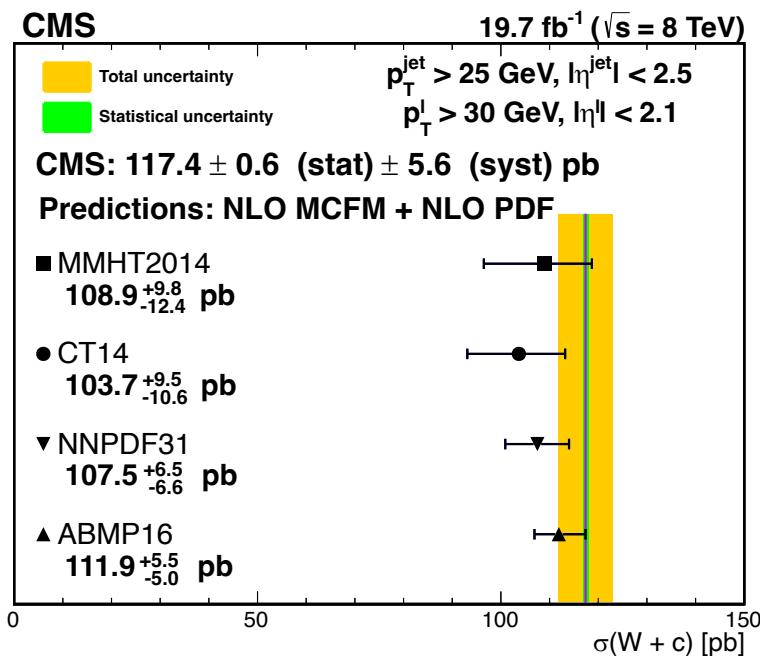
- Total and differential distributions measured, as a function of the  $p_T$  and  $\eta$  of the lepton ( $e, \mu$ ) from the W decay
- W+c events have opposite sign (OS), contrary to gluon splitting where both same sign (SS) and OS are present -> used for background subtraction
- Charm jets identified by:
  - a muon inside the jet (SL)
  - a displaced secondary vertex inside the jet (SV)
- 4 exclusive categories (SV/SL e,  $\mu$ ) combined



# W+c at 8 TeV

CMS-SMP-18-013

- Cross sections and the ratio  $W^+/W^-$  measured in fiducial region
- Main systematics in cross section : SV efficiency (1.8%), charm fragmentation (2.6%), luminosity (2.6%). Systematics much reduced in the ratio
- Measurements slightly higher than MCFM NLO calculation with different PDF sets



$m_c = 1.5 \text{ GeV}, \alpha_S = 0.118(0.119)$  for MMHT2014, CT14, NNPDF3.1 (ABMP16)  
factorization and renormalization scale set to  $m_W$

# W+c at 8 TeV

CMS-SMP-18-013

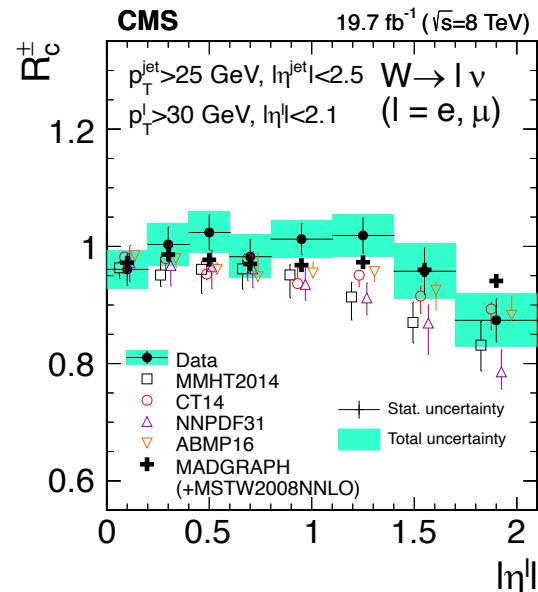
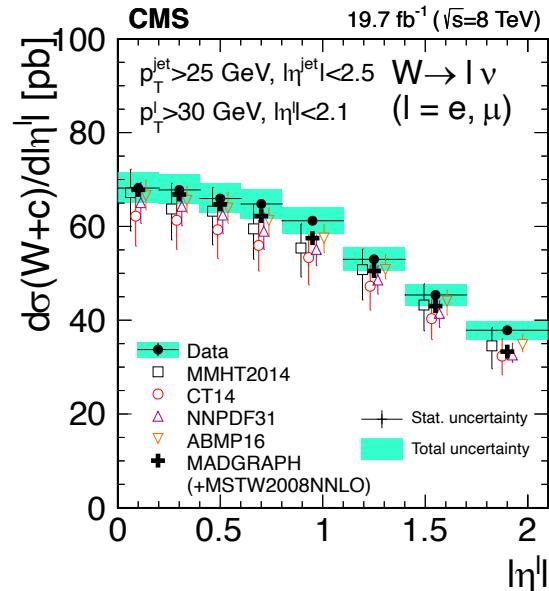
- Differential distribution in  $\eta$  used to extract strange parton density in the proton
- $Q^2=m_W^2$ ,  $0.001 < x < 0.080$
- Combined fit of
  - HERA inclusive DIS data
  - CMS W charge asymmetry at 7, 8 TeV
  - W+c at 7, 13 TeV
  - these data
- xFitter framework used
  - Roberts-Thorne GMVFN scheme at NLO

$$xf(x) = Ax^B(1-x)^C(1+Dx+Ex^2)$$

$$x\bar{s}(x) = A_{\bar{s}}x^{B_{\bar{s}}}(1-x)^{C_{\bar{s}}}$$

$$x\bar{s} = xs$$

$$B_{\bar{s}} = B_{\bar{d}} \quad (\text{released for systematics})$$



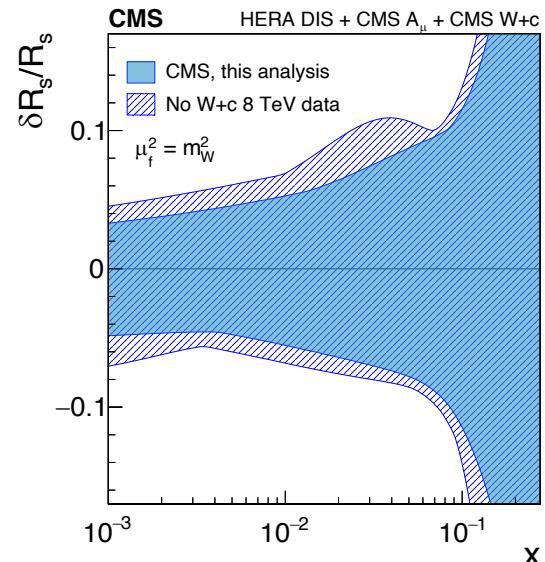
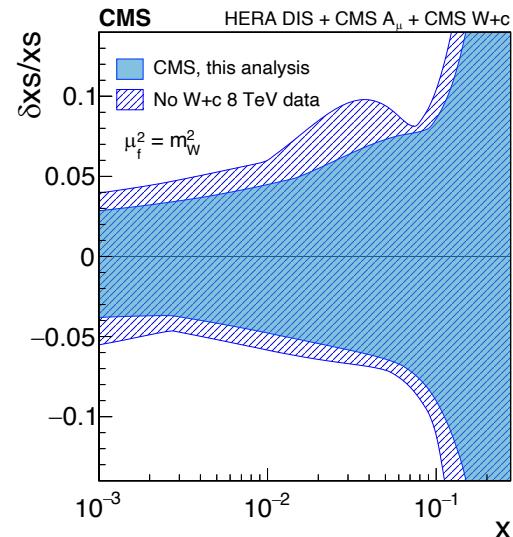
# W+c at 8 TeV

CMS-SMP-18-013

- Fit with and without these data show impact on strange distribution and on strangeness suppression factor:

$$R_s(x, \mu_f^2) = (s + \bar{s}) / (\bar{u} + \bar{d})$$

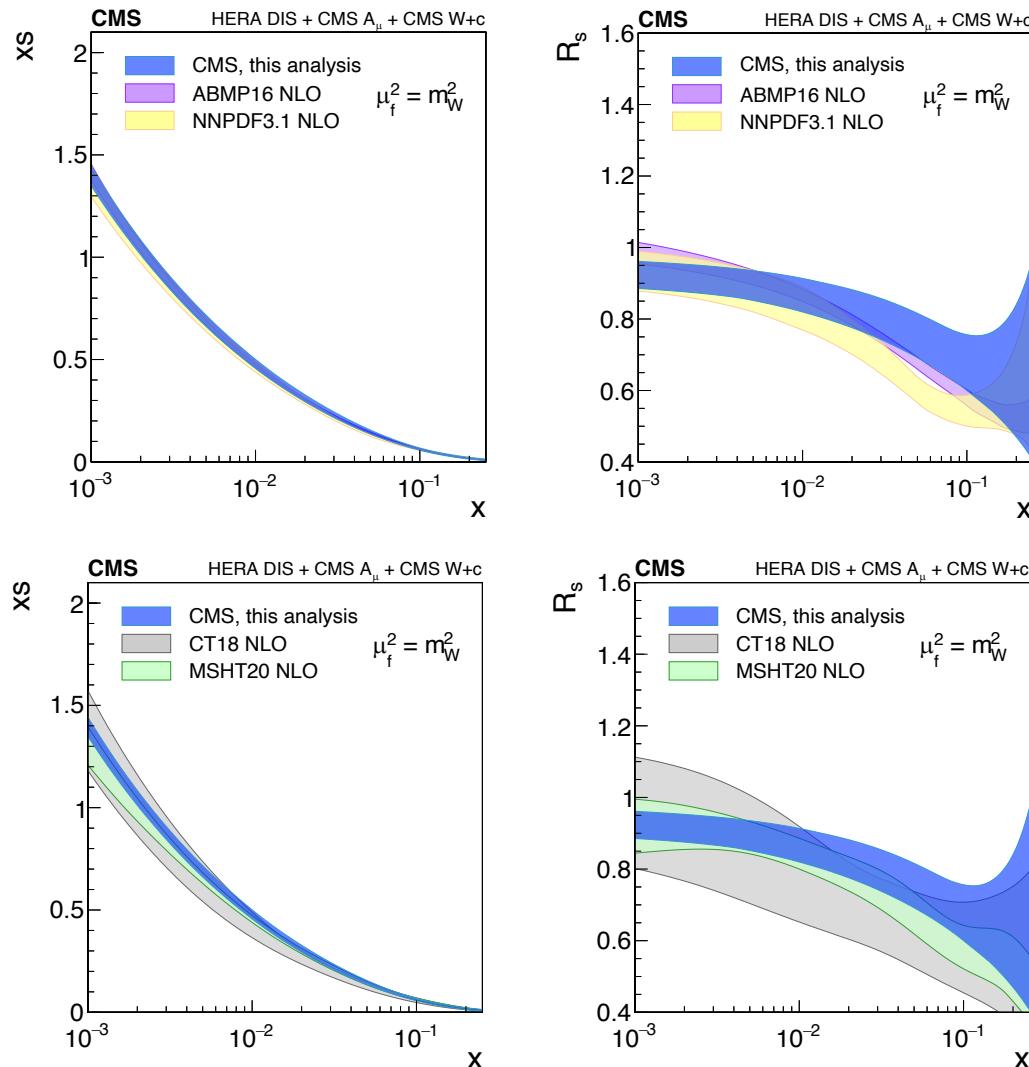
Data set	$\chi^2 / n_{\text{dp}}$
HERA I+II charged current	$e^+ p, E_p = 920 \text{ GeV}$ 41 / 39
HERA I+II charged current	$e^- p, E_p = 920 \text{ GeV}$ 59 / 42
HERA I+II neutral current	$e^- p, E_p = 920 \text{ GeV}$ 220 / 159
HERA I+II neutral current	$e^+ p, E_p = 820 \text{ GeV}$ 69 / 70
HERA I+II neutral current	$e^+ p, E_p = 920 \text{ GeV}$ 445 / 377
HERA I+II neutral current	$e^+ p, E_p = 460 \text{ GeV}$ 217 / 204
HERA I+II neutral current	$e^+ p, E_p = 575 \text{ GeV}$ 220 / 254
CMS W muon charge asymmetry 7 TeV	13.5 / 11
CMS W muon charge asymmetry 8 TeV	3.8 / 11
W+c 7 TeV	2.9 / 5
W+c 13 TeV	2.8 / 5
W+c 8 TeV	3.0 / 8
Correlated $\chi^2$	86
Log penalty $\chi^2$	5
Total $\chi^2 / n_{\text{dof}}$	1387 / 1171



# W+c at 8 TeV

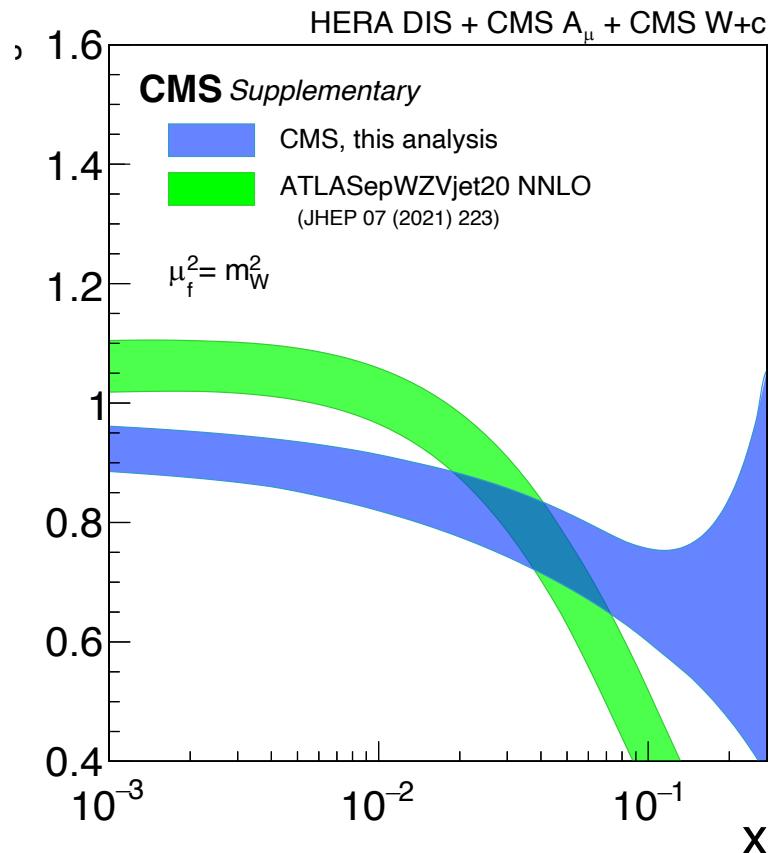
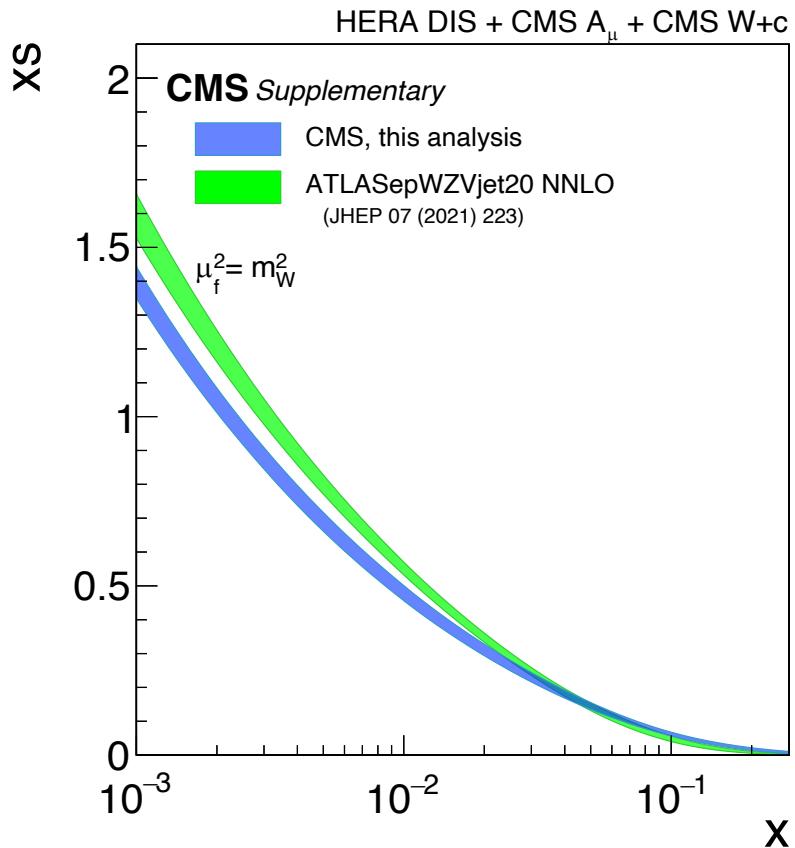
CMS-SMP-18-013

- Extracted strange distribution and strangeness suppression factor in agreement with other PDF sets
- Uncertainty bands in the fit include the experimental, model and parametrization unc.:
  - change of  $m_c$  ( $1.45 \rightarrow 1.55$  GeV),  $m_b$  and  $Q^2_{\min}$  for HERA data
  - change of parametrizations, also for the s-PDF
  - change of  $Q^2_0$  ( $1.6 \rightarrow 2.2$   $\text{GeV}^2$ )



# W+c at 8 TeV

CMS-SMP-18-013



Comparison with ATLAS

# Summary

- Extensive measurement of total and differential cross sections for  $Z+>= 1$  b-jet ,  $>=2$  b-jets, using all Run 2 data. SHERPA describes shape in general, but higher cross section, compared to MG5\_LO/NLO.
- Measurement of total and differential cross sections for  $p_T$  spectra of Z and charm jet  $Z+>= 1$  c-jet , NLO MC models SHERPA and MG5\_AMC above the data.
- Analysis of W+c data at 8 TeV allows to extract from a QCD fit the strange distribution: s-PDF and strangeness suppression factor compatible with those of previous CMS results and of various PDF sets.

# Backup

- Available MC models and calculations in differential Z+jets distributions:
  - **MADGRAPH5\_aMC at LO +PYTHIA8**, Z+N jets (N=0,..4,) at LO, MLM matching, NNPDF 3.0
  - **MADGRAPH5\_AMC@NLO 2.2.2 and 2.3.2** with FxFx merging scheme using PYTHIA8 for PS and CUETP8M1 tune for 2016, CP5 tune for 2017,2018. Z+ up to 2 jets at NLO, LO for Z+3 jets
  - **SHERPA 2.2 NLO** up to 2 additional partons, or LO up to 4 additional partons

# Backup

CMS-SMP-20-015 accepted in Phys. Rev. D

Table 2: Summary of the uncertainties (in percent) in the integrated cross sections for the dielectron, dimuon, and combined channels in the  $Z + \geq 1$  b jet and  $Z + \geq 2$  b jets events.

Uncertainty (%)	$Z + \geq 1$ b jet			$Z + \geq 2$ b jets		
	ee	$\mu\mu$	$\ell\ell$	ee	$\mu\mu$	$\ell\ell$
Statistical	1.0	0.7	0.6	7.7	5.9	4.6
JES, JER	2.7	3.0	2.9	6.9	5.4	5.8
b tagging/mistagging	3.0	2.9	2.9	5.4	6.0	5.8
Unclustered energy of $p_T^{\text{miss}}$	2.8	2.8	2.8	3.5	3.7	3.6
Background estimation	2.2	2.0	2.1	2.3	2.4	2.4
Pileup reweighting	2.4	1.7	1.9	2.9	2.1	2.4
Electron selection	4.6	—	1.5	4.3	—	1.4
Luminosity	1.6	1.6	1.6	1.6	1.6	1.6
Muon selection	—	0.6	0.4	—	1.0	0.7
Pileup jet identification	0.3	0.3	0.3	0.6	0.7	0.7
L1 prefiring	0.3	0.2	0.2	0.3	0.2	0.3
$\mu_R$ and $\mu_F$ scales	2.6	2.9	2.1	2.5	2.3	2.5
PDF	0.4	0.3	0.3	0.3	0.3	0.3
$\alpha_S$	0.3	0.2	0.2	0.1	0.1	0.1
Total experimental	7.6	5.9	6.1	11.4	9.7	9.9
Total theoretical	2.6	2.9	2.1	2.5	2.3	2.5

Table 4: Summary of the uncertainties (in percent) in the normalized differential distributions for the combined channel in the  $Z + \geq 1$  b jet events.

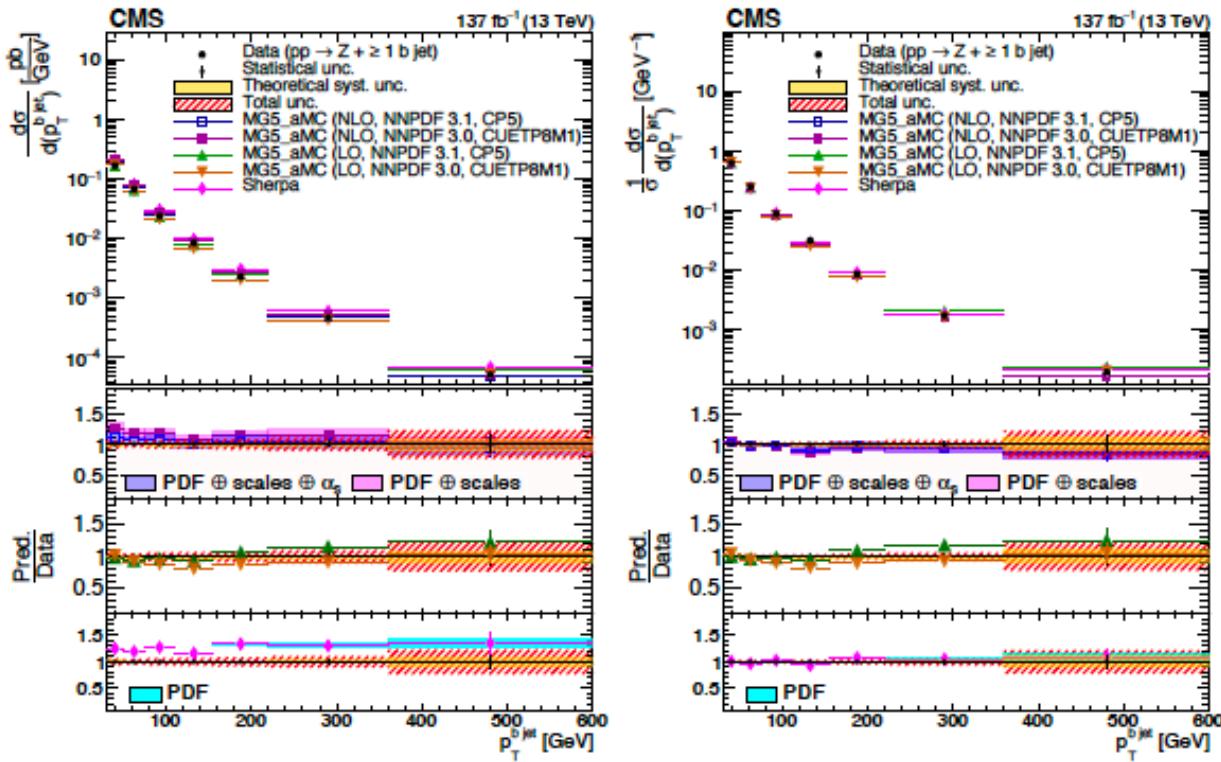
Observable/Uncertainty (%)	$p_T^Z$	$p_T^{\text{b jet}}$	$ \eta^{\text{b jet}} $	$\Delta\phi^{(Z,\text{b jet})}$	$\Delta y^{(Z,\text{b jet})}$	$\Delta R^{(Z,\text{b jet})}$
Statistical	1.0–5.2	1.0–12	0.7–1.4	1.0–2.1	0.7–4.7	0.8–3.3
JES, JER	1.0–5.8	1.0–6.7	0.5–4.5	0.5–2.1	0.5–4.2	0.5–4.2
b tagging/mistagging	0.5–3.9	0.5–5.6	0.5–1.7	0.5–2.9	0.5–3.7	0.5–1.7
Unclustered energy of $p_T^{\text{miss}}$	0.5–0.7	0.5–0.7	0.5–1.1	0.5–1.0	0.5	0.5–1.0
Background estimation	0.5–1.9	0.6–1.7	0.5–1.4	0.5–2.0	0.5–2.4	0.5–3.3
Pileup reweighting	0.5–0.9	0.5–1.3	0.5	0.5–0.7	0.5	0.5–0.8
Model dependency	0.5–1.4	0.5–0.8	0.5	0.5–1.2	0.5	0.5–0.9
Electron selection	0.5–1.3	0.5	0.5	0.5	0.5–0.9	0.5–0.6
Muon selection	0.5–1.0	0.5	0.5	0.5	0.5	0.5
Pileup jet identification	0.5	0.5	0.5	0.5	0.5	0.5
L1 prefiring	0.5	0.5	0.5	0.5	0.5	0.5
$\mu_R$ and $\mu_F$ scales	1.0–7.0	0.5–9.4	0.5–2.0	1.0–2.1	0.5–2.1	0.5–3.7
PDF	0.5–2.0	0.5–1.6	0.5–0.7	0.5–1.8	0.5–1.6	0.5–3.3
$\alpha_S$	0.5–1.0	0.5–0.7	0.5	0.5–0.9	0.5	0.5–1.1

Table 7: Measured and predicted cross sections (in pb) for the  $Z + \geq 1$  b jet and  $Z + \geq 2$  b jets final states. The cross section ratios between the  $Z + \geq 2$  b jets and  $Z + \geq 1$  b jet are shown in the last three rows for the dielectron, dimuon, and combined channels. In the measured results the first, second, and third uncertainties correspond to the statistical, systematic, and theoretical sources, respectively. The MG5\_aMC (NLO) predictions include theoretical uncertainties (PDF, and renormalization and factorization scales).

Channel	Measured	MG5_aMC	MG5_aMC	MG5_aMC	MG5_aMC	SHERPA
		LO	LO	NLO	NLO	
		NNPDF 3.0	NNPDF 3.1	NNPDF 3.0	NNPDF 3.1	
$Z + \geq 1$ b jet	ee	$6.45 \pm 0.06 \pm 0.49 \pm 0.17$	6.25	6.33	$7.86 \pm 0.52$	$7.05 \pm 0.48$
	$\mu\mu$	$6.55 \pm 0.05 \pm 0.39 \pm 0.19$	6.26	6.34	$7.86 \pm 0.51$	$7.02 \pm 0.47$
	$\ell\ell$	$6.52 \pm 0.04 \pm 0.40 \pm 0.14$	6.25	6.34	$7.86 \pm 0.51$	$7.03 \pm 0.47$
$Z + \geq 2$ b jets	ee	$0.66 \pm 0.05 \pm 0.07 \pm 0.02$	0.62	0.72	$0.89 \pm 0.08$	$0.77 \pm 0.07$
	$\mu\mu$	$0.65 \pm 0.04 \pm 0.06 \pm 0.02$	0.64	0.71	$0.91 \pm 0.09$	$0.77 \pm 0.07$
	$\ell\ell$	$0.65 \pm 0.03 \pm 0.07 \pm 0.02$	0.63	0.71	$0.90 \pm 0.09$	$0.77 \pm 0.07$
Ratio	ee	$0.102 \pm 0.008 \pm 0.008 \pm 0.004$	0.100	0.113	$0.113 \pm 0.016$	$0.110 \pm 0.013$
	$\mu\mu$	$0.100 \pm 0.006 \pm 0.006 \pm 0.004$	0.103	0.112	$0.116 \pm 0.016$	$0.110 \pm 0.013$
	$\ell\ell$	$0.100 \pm 0.005 \pm 0.007 \pm 0.003$	0.102	0.112	$0.114 \pm 0.016$	$0.110 \pm 0.013$

# Backup

CMS-SMP-20-015 accepted in Phys. Rev. D



# Backup

CMS-SMP-18-013

PDF set	$\sigma(W+c)$ [pb]	$\delta_{\text{PDF}} [\%]$	$\delta_{\text{scales}} [\%]$	$\delta_{\alpha_s} [\%]$	Total uncert. [pb]
MMHT2014	108.9	+6.0 -9.1	+4.4 -4.6	$\pm 5$	+9.8 -12.4
CT14	103.7	+7.6 -8.7	+4.5 -4.6	$\pm 2.2$	+9.5 -10.6
NNPDF3.1	107.5	$\pm 3.5$	+4.4 -4.5	$\pm 2.2$	+6.5 -6.6
ABMP16	111.9	$\pm 0.9$	+4.8 -4.4	—	+5.5 -5.0
CMS		$117.4 \pm 0.6 \text{ (stat)} \pm 5.4 \text{ (syst)} \text{ pb}$			

Source	Uncertainty [%]
Lepton efficiency	0.7
Jet energy scale and resolution	0.8
$p_T^{\text{miss}}$ resolution	0.3
Pileup modelling	0.4
$\mu$ in jet reconstruction efficiency	0.9
Secondary vertex reconstruction efficiency	1.8
Secondary vertex charge determination	1.0
Charm fragmentation and decay fractions	2.6
Charm fragmentation functions	0.3
Background subtraction	0.8
PDF	1.0
Limited size of MC samples	0.6
Integrated luminosity	2.6

$$\begin{aligned}
 xg(x) &= A_g x^{B_g} (1-x)^{C_g}, \\
 xu_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} \left(1 + E_{u_v} x^2\right), \\
 xd_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \\
 x\bar{u}(x) &= A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} \left(1 + D_{\bar{u}} x\right), \\
 x\bar{d}(x) &= A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}}, \\
 x\bar{s}(x) &= A_{\bar{s}} x^{B_{\bar{s}}} (1-x)^{C_{\bar{s}}}.
 \end{aligned}$$