V+heavy flavor jets and constraints to PDFs

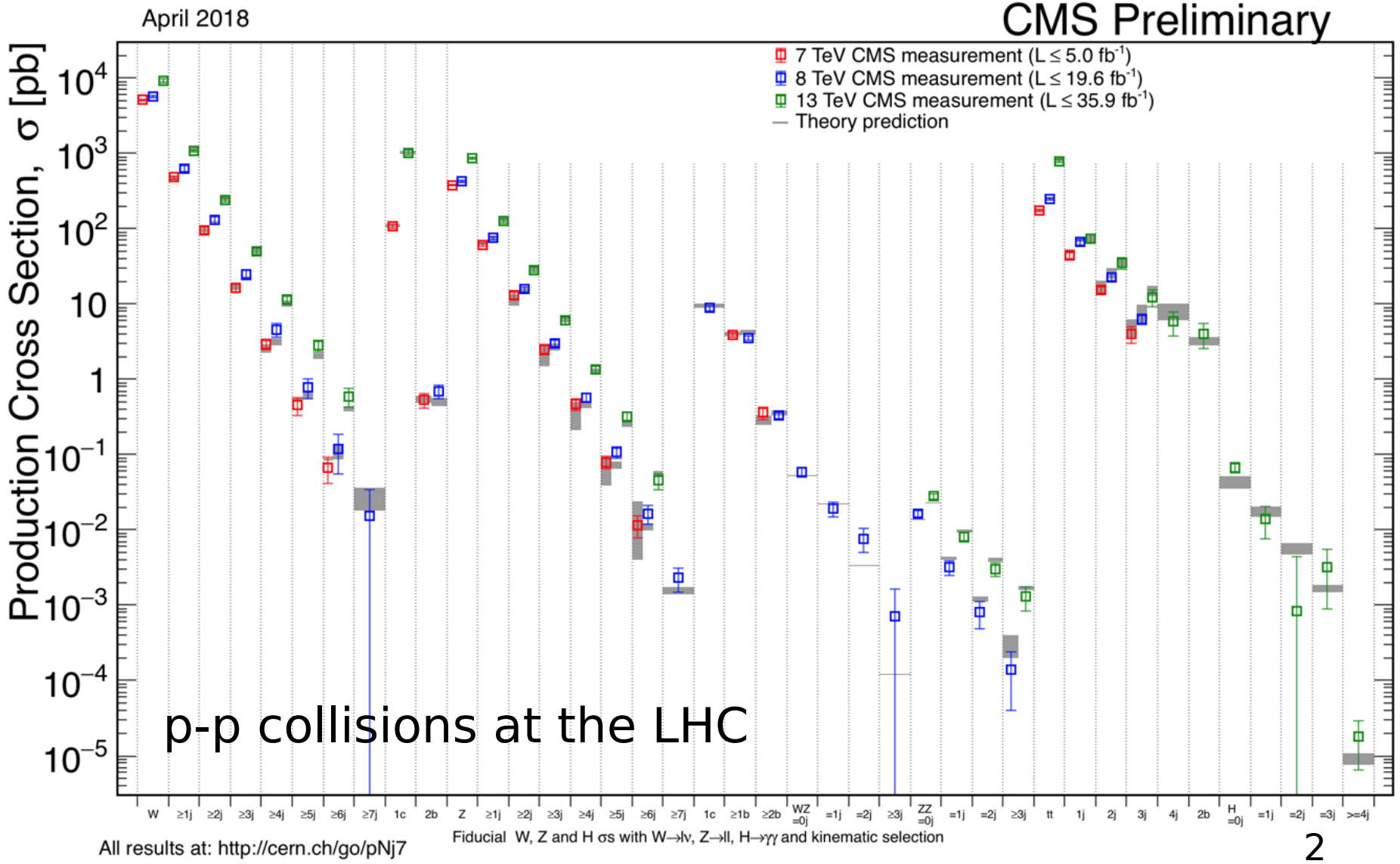
Juan Pablo Fernández Ramos (CIEMAT)
on behalf of the CMS Collaboration

ICHEP @ Seoul
July 07, 2018

Vector boson+jet production overview

- Essential for tests of QCD predictions
- Focus on Vector boson + HF

- See V+jets for other CMS results



CMS V+HF measurements

channel	sqrt(s)	Int. Luminosity	documentation
Z + b	7 TeV	2.2 fb^{-1}	J. High Energy Phys. 06 (2012) 126 CMS-EWK-11-012
Z + b hadrons	7 TeV	5.2 fb^{-1}	J. High Energy Phys. 12 (2013) 039 CMS-EWK-11-015
Z + b(b)	7 TeV	5 fb^{-1}	J. High Energy Phys. 06 (2014) 120 CMS-SMP-13-004
W + bb	7 TeV	5 fb^{-1}	Phys. Lett. B 735 (2014) 204 CMS-SMP-12-026
W + c	7 TeV	5 fb^{-1}	J. High Energy Phys. 02 (2014) 013 CMS-SMP-12-002
Z + b(b)	8 TeV	19.8 fb^{-1}	EPJC 77 (2017) 751 CMS-SMP-14-010
W + bb	8 TeV	19.8 fb^{-1}	EPJC 77 (2017) 92 CMS-SMP-14-020
Z + c	8 TeV	19.8 fb^{-1}	EPJC 78 (2018) 287 CMS-SMP-15-009
W + c	13 TeV	35.7 fb^{-1}	CMS-PAS-SMP-17-014

- Started with V+b quark
- This talk : focus on latest results concerning the V+charm quark

Motivation

- Measurements of σ(pp→Z/W+c) provide tests QCD predictions. **Results sensitive to hard scattering process & associated soft QCD radiation**

- Allows better understanding of proton structure. Z/W+c jets tests PDF for c,s respectively.

- Intrinsic charm quark component inside the proton enhances σ(Z+c) @ high p_T(Z).

- Background to some SM processes and in searches

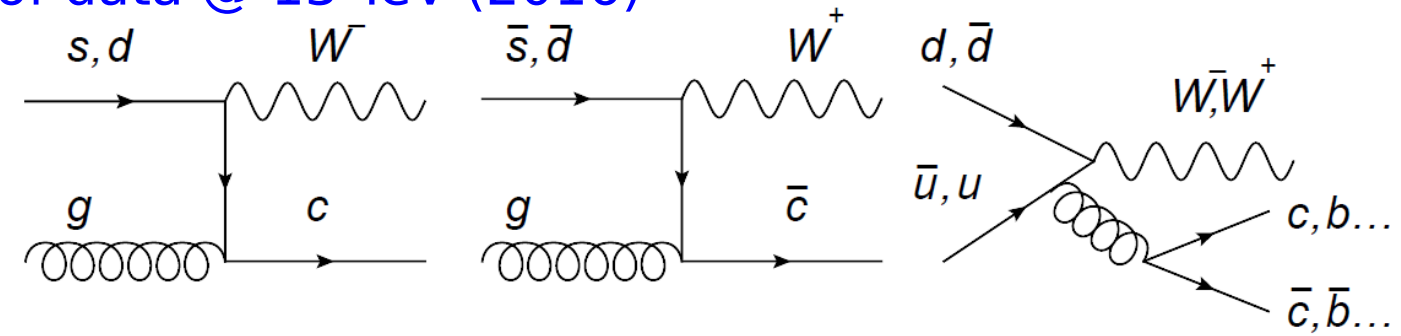
W+c analysis

Measurement of associated production of W bosons with charm quarks in proton-proton collisions at $\sqrt{s}$ = 13 TeV

CMS-PAS-SMP-17-014

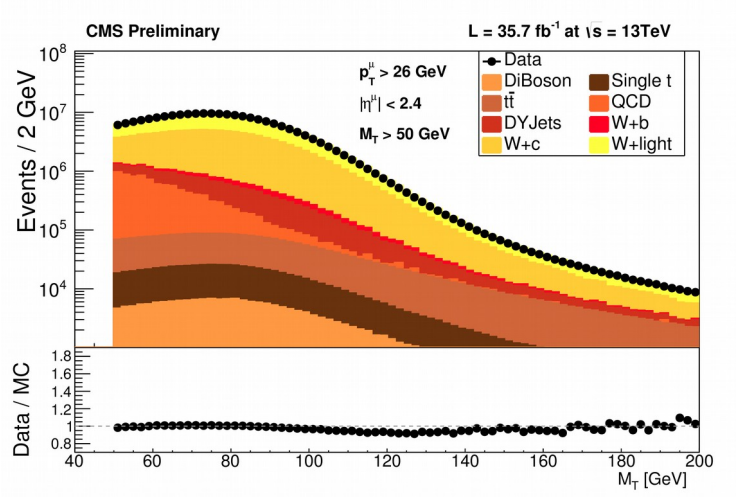
Event selection:

- 35.7 fb^{-1} of data @ 13 TeV (2016)

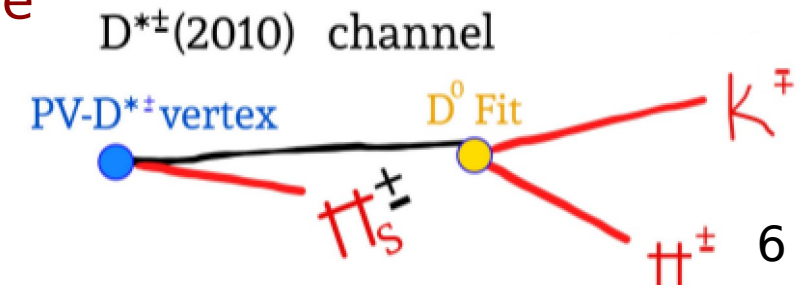


OS event | OS event | Bckg: 50% OS, 50% SS

- W→ μν selection
  - Isolated muon with $p_T(\mu)>26$ GeV and $|\eta(\mu)|<2.4$
  - $M_T(\mu,\nu) > 50$ GeV

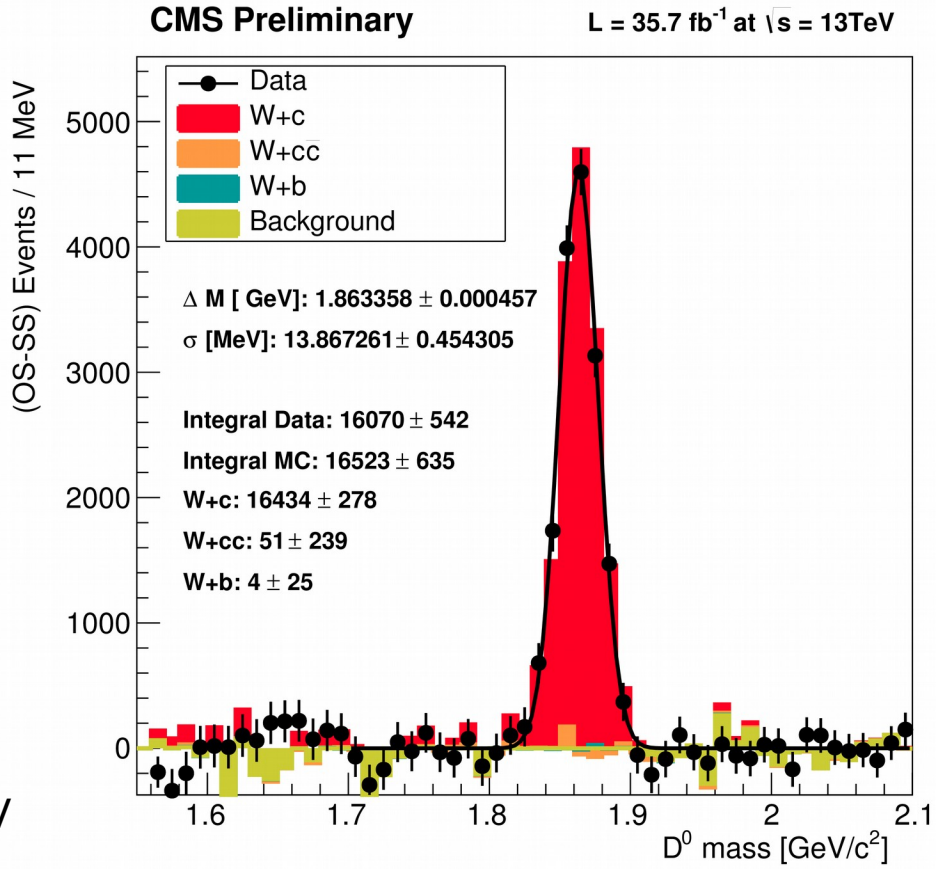


- W+c selection:
  - Identify jets from heavy flavor quarks through the $D^{*\pm}$ exclusive decay



- OS – SS to remove symmetric background

$D(2010)^{*\pm}$ reconstruction



- $D^{*\pm} \rightarrow D^0 \pi_s^{\pm}$ [$D^0 \rightarrow K^-\pi^+(+c.c.)$] decay.
- All track combinations
- Kaon: track with sign opposite to π_s
- $p_T(D^*)>4$ && $p_T(D^*)/\Sigma p_{T(cone\ 0.4)} > 0.2$
- $p_T(K)>1$ && $p_T(\pi)>1$ GeV, $p_T(\pi_s)>0.35$ GeV
- $|\Delta R(D^*, jet)| < 0.5$, $|\Delta R(D^0, \pi_s)| < 0.15$
- $|m(D^0)-1.865|<35$ MeV, $|\Delta m-145|<1$ MeV

Non resonant background in the signal region subtracted using wrong charge $D^0 \rightarrow K^-\pi^-(+c.c.)$ combinations

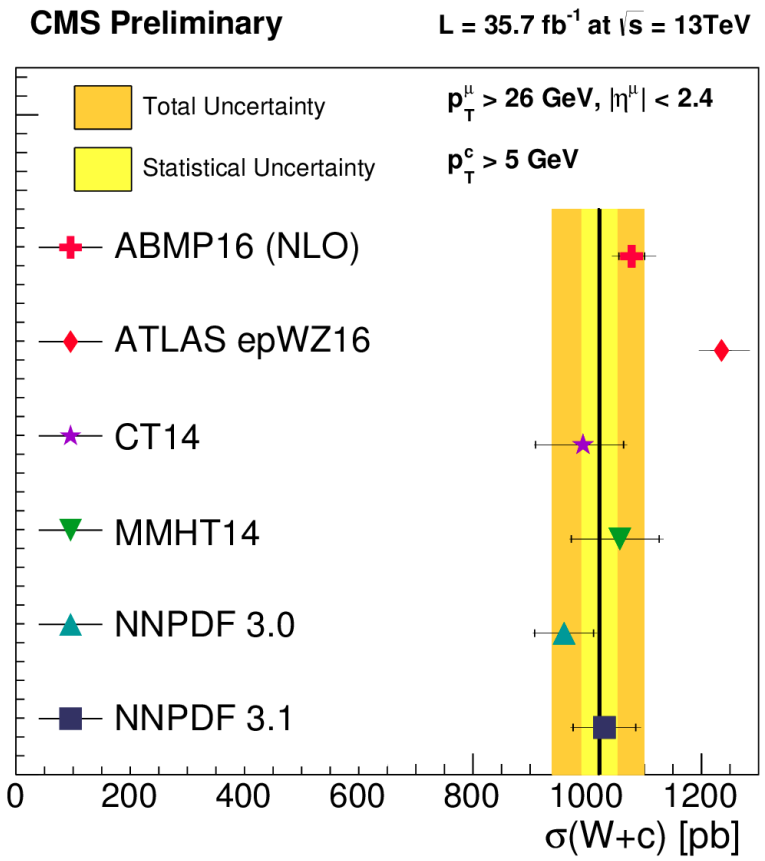
$19.2 \pm 0.6 \times 10^3$ $D^{*\pm}$ ($W \rightarrow \mu\nu$) candidates

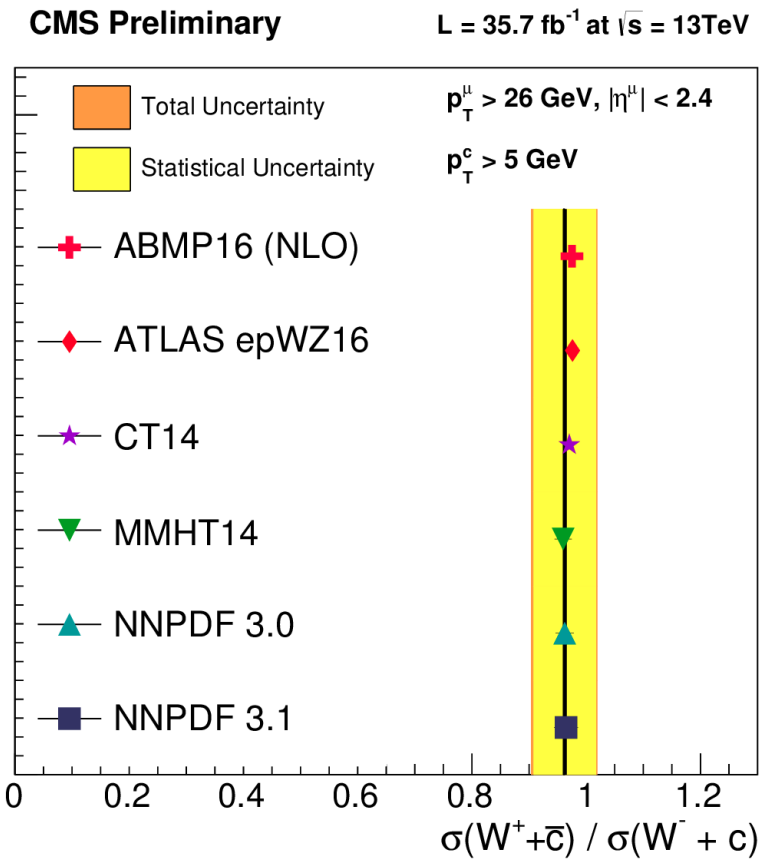
- Main sources of systematic uncertainties : fragmentation, MC-Statistics

Others (lumi, tracking efficiency, BR, PDF). Total : ~7%

- Signal MC: W+jets generated w. MADGRAPH5@NLO+PYTHIA8 (NNPDF3.0)

$\sigma(W+c)$ and $\sigma(W^{+}+\bar{c})/\sigma(W^{-}+c)$ ratio





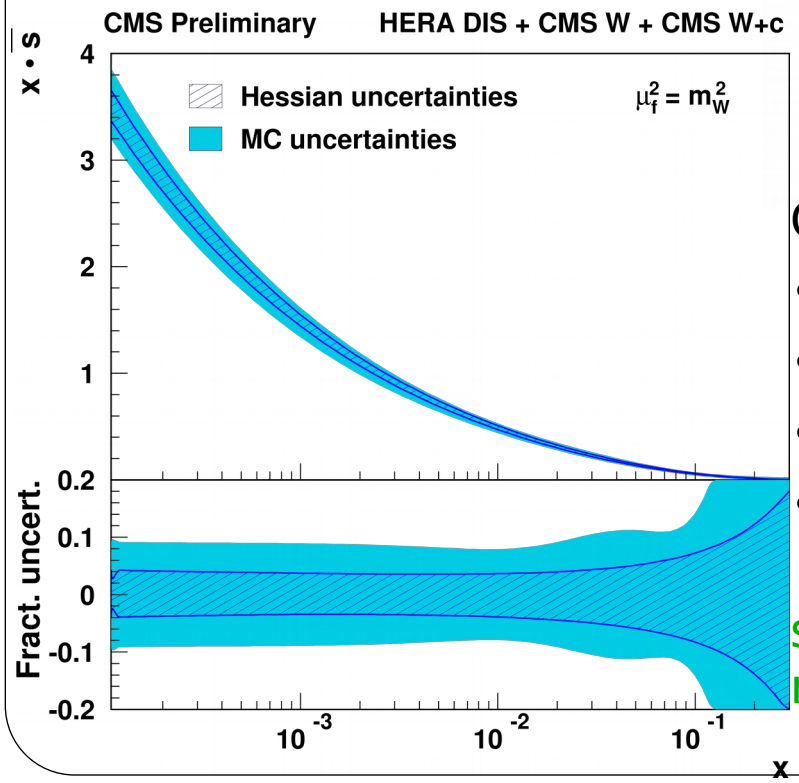
$$\sigma(\mathrm{W}+\mathrm{c}) = 1026 \pm 31\ (\mathrm{stat})\ ^{+76}_{-72}\ (\mathrm{syst})\ \mathrm{pb}$$

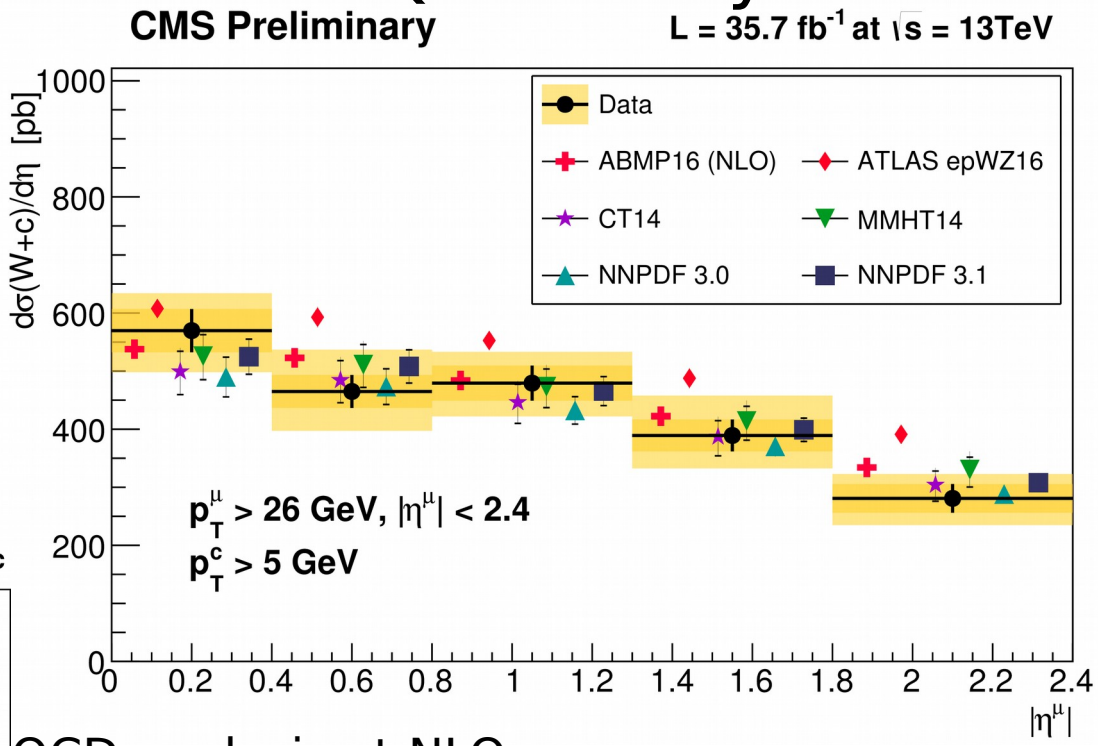
$$\frac{\sigma(\mathrm{W}^{+}+\bar{\mathrm{c}})}{\sigma(\mathrm{W}^{-}+\mathrm{c})} = 0.968 \pm 0.055\ (\mathrm{stat})\ ^{+0.015}_{-0.028}\ (\mathrm{syst})$$

In good agreement with the theoretical predictions at NLO using different PDF sets except for ATLASepWZ16nnlo PDF→ results do not support high strange quark contribution

Differential cross section and QCD analysis

- W+c vs η(μ) [W^+ and W^- also]
- In good agreement with NLO using different PDF sets except for ATLASepWZ16nnlo PDF
- Impact on strange PDF :





QCD analysis at NLO:

- This analysis W+c @ 13 TeV, $W^+ + \bar{c}$ & $W^- + c$
- inclusive DIS measurements
- CMS W+c @ 7 TeV
- W lepton charge asymmetry @ 7 & 8 TeV

The strange quark distribution and strangeness suppression factor agree with results from neutrino-scattering experiments.

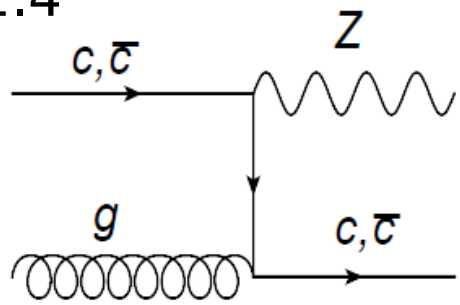
Z+c analysis

Measurement of associated Z + charm production in proton–proton collisions at $\sqrt{s}$ = 8 TeV

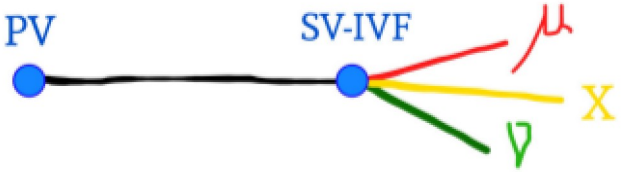
Eur. Phys. J. C (2018) 78:287

Selection:

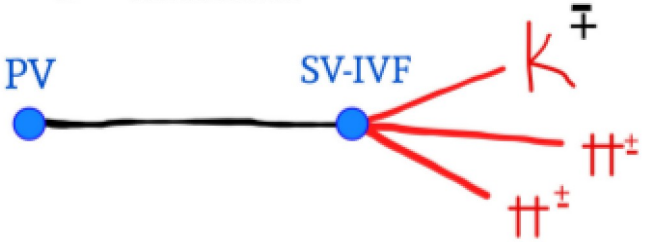
- 19.7 fb^{-1} of data @ 8 TeV (2012)
- $Z \rightarrow l^+l^-$ selection ($l = e,\mu$)
  - Isolated leptons with $p_T(l)>20$ GeV and $|\eta(l)|<2.4$
  - Dilepton invariant mass: [71,111] GeV
- anti-k_T jet R=0.5: $p_T^{jet} >25$ GeV & $|\eta^{jet}|<2.5$



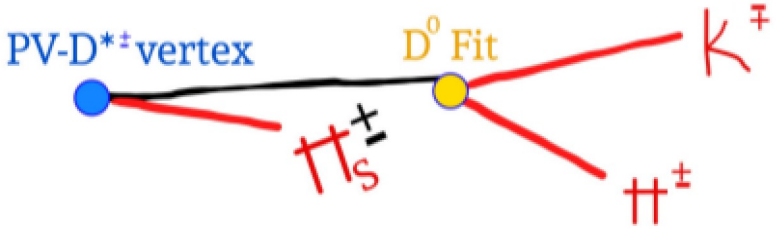
Semileptonic channel



$D^{\pm}$ channel



$D^{*\pm}(2010)$ channel



- Heavy flavor selection:

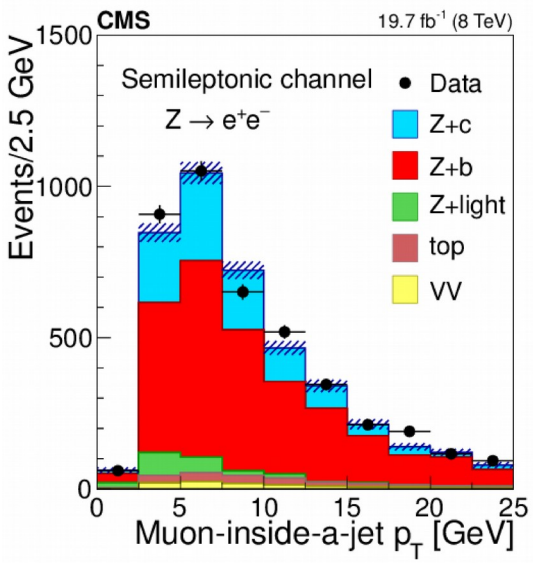
Identify jets from heavy flavor quarks through the decay of charm hadrons in 3 final states:

- Semileptonic decay of c/b hadrons: muon in a jet
- $D^{*\pm}$ and $D^{\pm}$ exclusive decays in jet

- Signal MC: DY+jets generated w. MADGRAPH5@LO+PYTHIA6 (PDF:CTEQ6L)

σ(pp→Z+X) calculated at NNLO with FEWZ (PDF set MSTW2008NNLO)

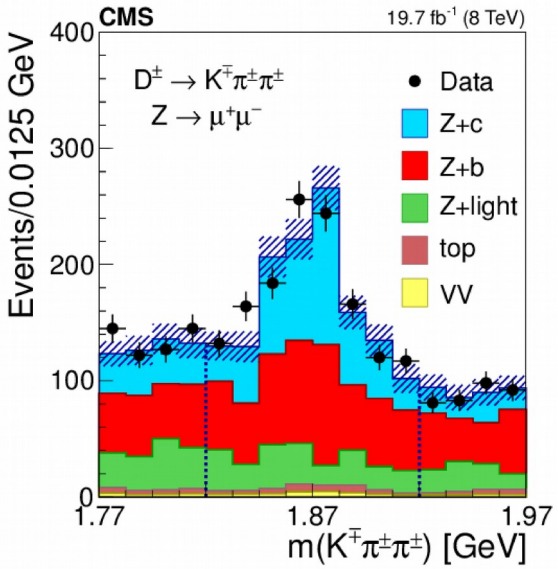
Semileptonic channel



- μ inside a jet and taking part of a secondary vertex (SV). This reduces the light contribution more than standard b-tagging algorithms.
- p_T^{μ}<25 GeV, with p_T^{μ}/p_T^{jet}<0.6, $|\eta^{\mu}|$<2.5
- non-isolated, I_{comb}/p_T^{μ}>0.2

Z+c: ~25%, Z+b: ~65%, Z+light:~5%,others:~5%

$D^{\pm}$channel



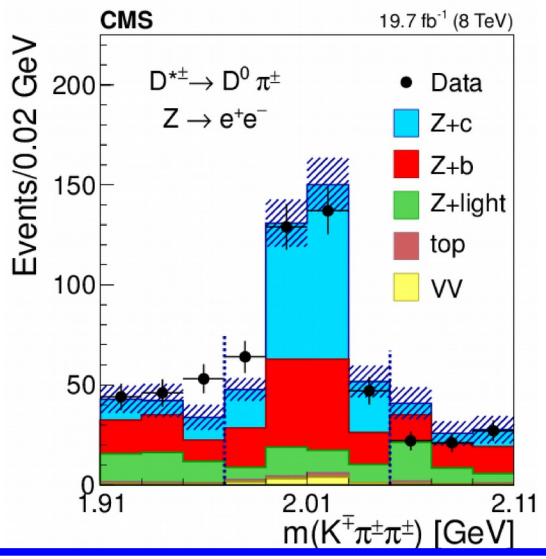
- Jet with a 3 track SV
- Search for D± → K∓π±π± resonant peak.
- Define signal region (SR):

  $|m(D^{\pm}) - 1.87| < 0.05$ GeV
- Sideband (SB) region for non resonant bck. subtraction:

  $0.05<|m(D^{\pm})-1.87|<0.1$ GeV

Z+c: ~60% Z+b: ~35%
Z+light: <1% Others:<4%

$D^{*\pm}(2010)$ channel



- $D^{*\pm} \rightarrow D^0 \pi_s^{\pm}$ [$D^0 \rightarrow K^-\pi^+$] +c.c.
- Jet tracks combinations
- Kaon:track with sign opposite to π_s
- D^0 vertex reconstruction
- $p_T(K, \pi, \pi_s)$ > 1.75, 0.75, 0.5 GeV
- $|\Delta R(D^*,jet)|<0.5, |\Delta R(D^0,\pi_s)|<0.1$
- $|m(D^0)-1.865|<100$ MeV,
- $|\Delta m-145|<5$ MeV
- SR: 1.97 < m(D*) < 2.05 GeV
- SB:$0.06<|m(D^{*\pm})-2.01|<0.12$ GeV

Z+c:~65% Z+b:~30%
Z+light:<1% Others:<4%

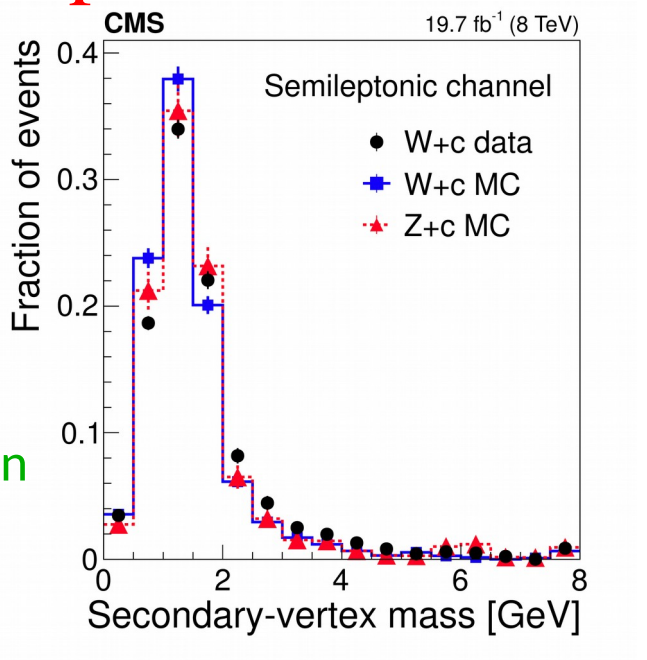
Z+b/Z+c separation for the Semileptonic channel

(using the corrected vertex mass discriminant)

$$M_{\text{vertex}}^{\text{corr}} = \sqrt{M_{\text{vertex}}^2 + p_{\text{vertex}}^2 \sin^2\theta} + p_{\text{vertex}} \sin\theta$$

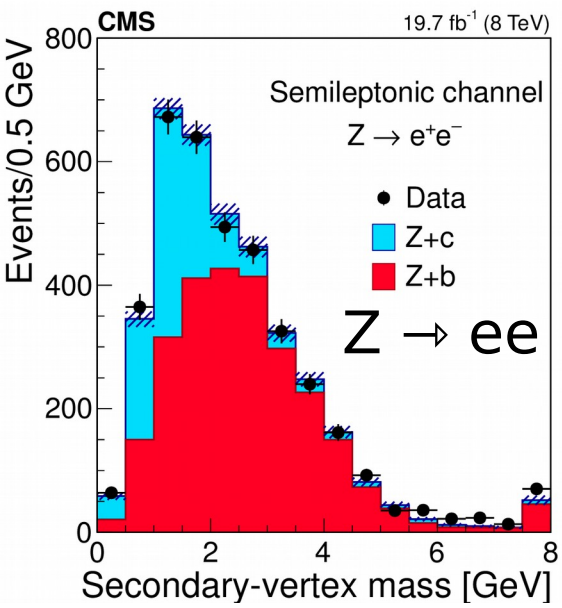
Correction included to account for unidentified neutral decay products

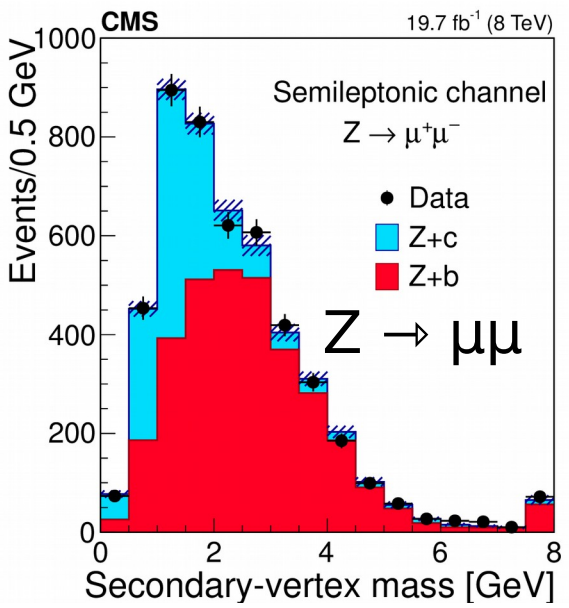
- Profit from W+c jet events to model the shape of the secondary vertex mass discriminant (c-jet template)
  - W→ lν plus jets with same HF selection as Z+HF : μ in jet (or D-hadron exclusive decays [next slide])



- Shape and normalization of templates from rest of contributions taken from MC (excpet ttbar from data) and subtracted from experimental distributions

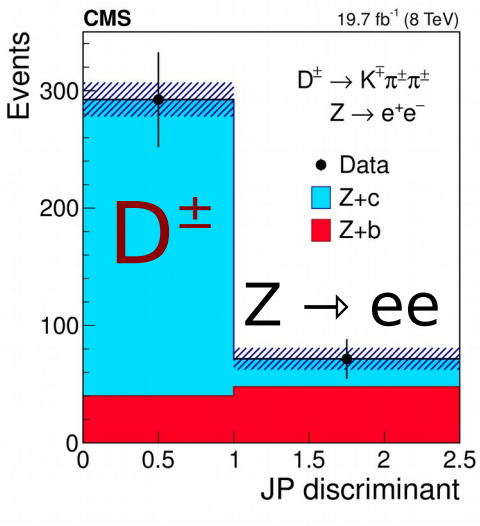
- # Z+c and Z+b extracted from a χ^2 minimization fit of the Z+c/Z+b templates to the background subtracted data distribution

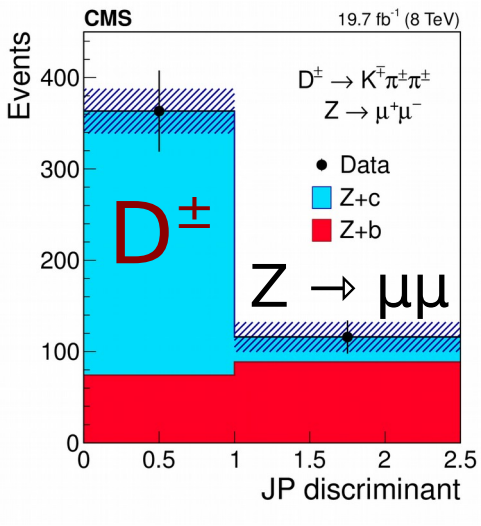


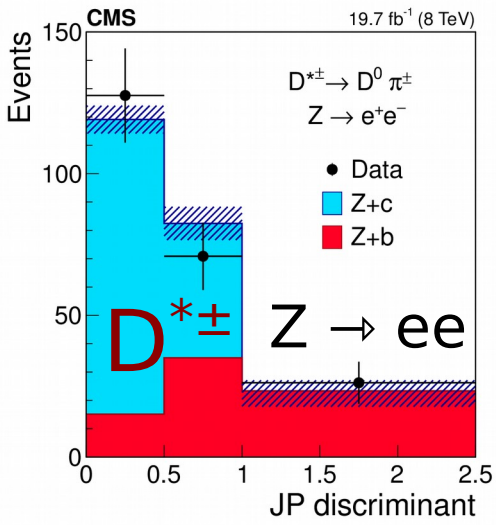


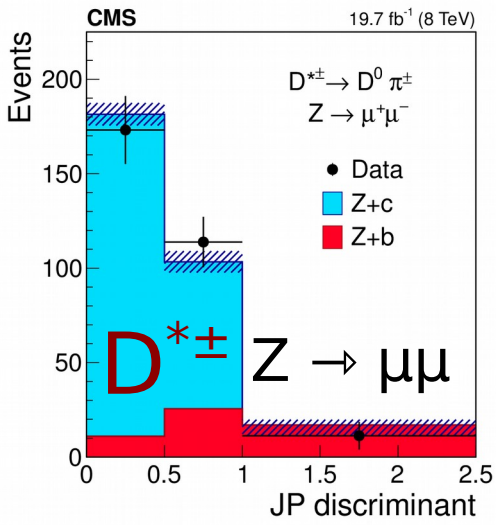
Z+b/Z+c separation for the $D^{\pm}$ and $D^{*\pm}(2010)$ channels

(using the JP discriminant)





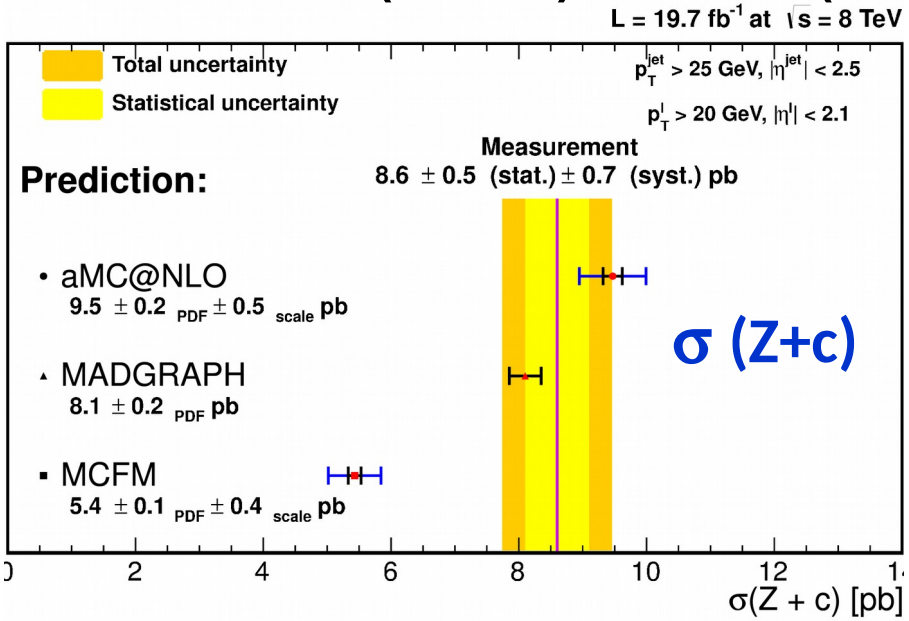


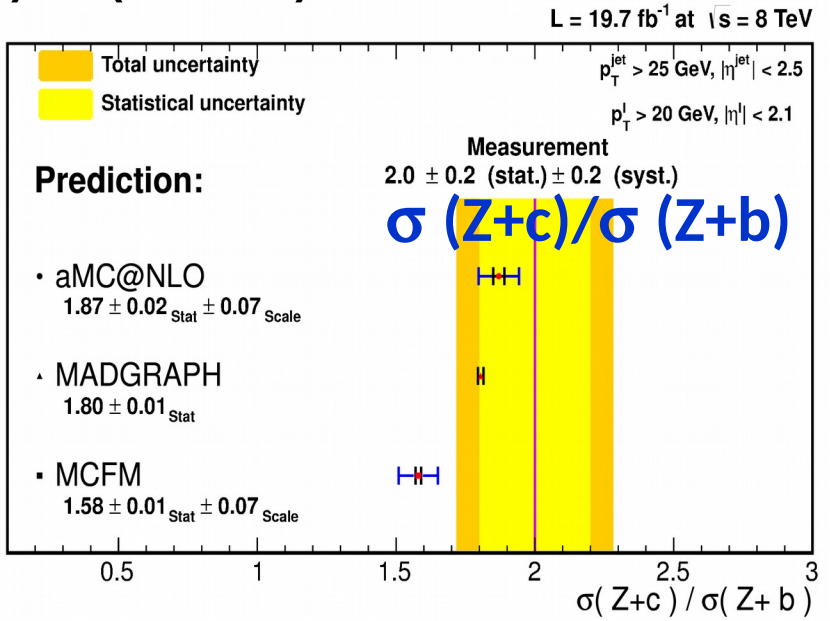


- JP : likelihood estimate of prob. of jet tracks to come from primary vertex
- c-jet templates from W+c data
- #Z+c extracted from a χ^2 minimization fit of the Z+c/Z+b templates to the experimental distributions
- c/b separation clearer in the D* mode (the soft pion comes from the PV for c→ D* and not for b→ B→D*)

Signal strength w.r.t. MC prediction in the 0.9-1.1 range

σ(Z+c) and σ(Z+c)/σ(Z+b)ratio





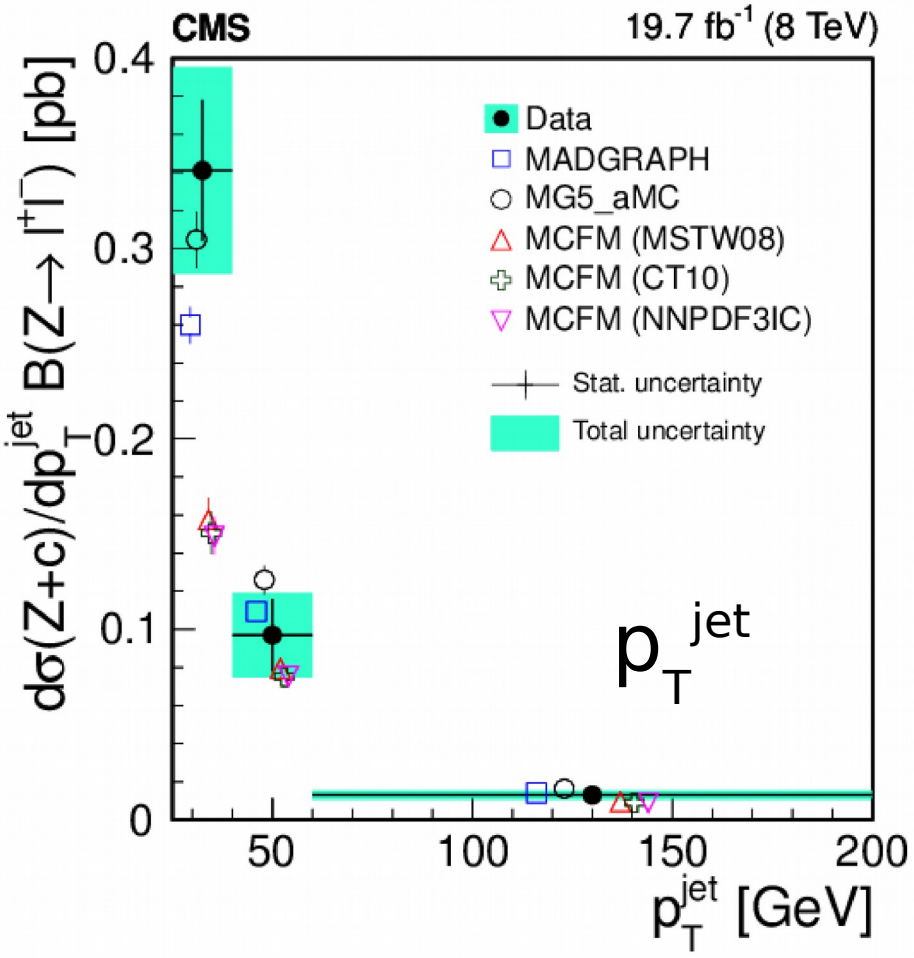
σ(Z+c)= 8.6 ± 0.5(stat) ± 0.7(syst) σ(Z+c)/σ(Z+b)= 2.0± 0.2 ±0.2

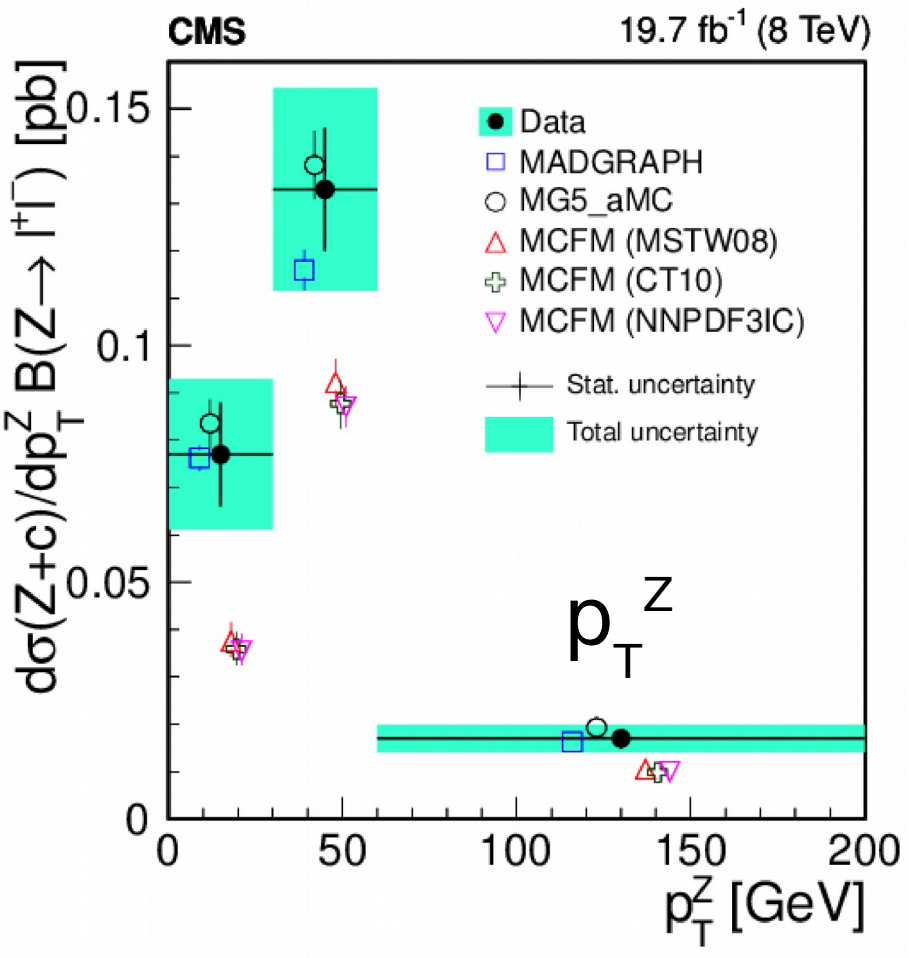
μ,e(p_T)>25 GeV & |η|<2.1, **jet(p_T)>25 GeV** &|η|<2.5

MCFM do not include contributions from GS

- Systematics
  - production and decay charm fractions of c hadrons in the sim.samples
  - determination of the efficiency to identify charm in jets(ctagging eff.)
  - uncertainty in the jet energy scale and jet energy resolution and MET

Differential cross sections





Sensitivity to intrinsic charm at high p_T region:

- ~2% charm quark component (intrinsic + perturbative) → 20-25%increase in the production of Z+c events with a $p_T^Z \approx 100$ GeV expected
- No increase in the production rate in the highest p_T^{jet} bin is observed (in agreement with current upper limits on IC component)

Conclusions

- Measured sigma W+c & $W^+ + \bar{c}/W^- + c$ total and differential (with $\eta(\mu)$)
- Input to strange PDF QCD analysis : s PDF in agreement with previous determinations
- First measurement of Z+c in the central region : good description of results by MadGraph5 amc@nlo using PDF sets with no IC component
- Already collected ~ 100 fb-1 @ 13 TeV (2016-2018 so far) good prospects for new results in vector boson + HF with full statistics offered by Run II

These and previous results from CMS can be found:
http://cms-results.web.cern.ch/cms-results/public-results/publications/SMP/SMP.html

Back up

Cross section determination

$$\sigma(\mathrm{Z}+\mathrm{c}) = \frac{N_{\mathrm{Z+c}}^{fitted}}{\epsilon_c^{\mathrm{Z+c}} \mathcal{L}}$$

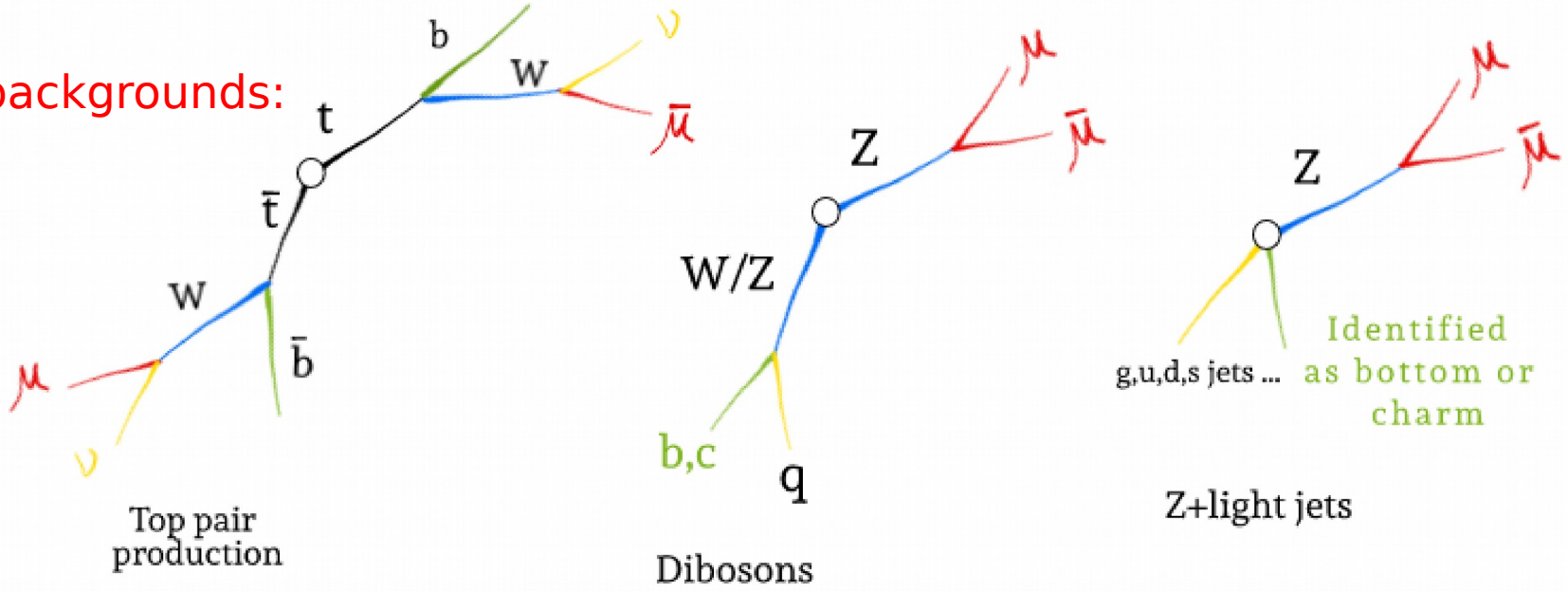
Semileptonic mode			
Channel	$N_{\mathrm{Z+c}}^{\mathrm{signal}}$	$\mathcal{C}_{\mathrm{Z+c}}$ (%)	$\sigma(\mathrm{Z+c})$ [pb]
$\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$	1066 ± 95	0.63 ± 0.03	$8.6 \pm 0.8 \pm 1.0$
$\mathrm{Z} \to \mu^+\mu^-$	1449 ± 144	0.81 ± 0.03	$9.1 \pm 0.9 \pm 1.0$
$\mathrm{Z} \to \ell^+\ell^-$	$\sigma(\mathrm{Z+c}) = 8.8 \pm 0.6\,(\mathrm{stat}) \pm 1.0\,(\mathrm{syst})\,\mathrm{pb}$		
Channel	$N_{\mathrm{Z+b}}^{\mathrm{signal}}$	$\mathcal{C}_{\mathrm{Z+b}}$ (%)	$\sigma(\mathrm{Z+c})/\sigma(\mathrm{Z+b})$
$\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$	2606 ± 114	2.90 ± 0.08	$1.9 \pm 0.2 \pm 0.2$
$\mathrm{Z} \to \mu^+\mu^-$	3240 ± 147	3.93 ± 0.10	$2.2 \pm 0.3 \pm 0.2$
$\mathrm{Z} \to \ell^+\ell^-$	$\sigma(\mathrm{Z+c})/\sigma(\mathrm{Z+b}) = 2.0 \pm 0.2\,(\mathrm{stat}) \pm 0.2\,(\mathrm{syst})$		
$\mathrm{D}^{\pm}$ mode			
Channel	$N_{\mathrm{Z+c}}^{\mathrm{signal}}$	$\mathcal{C}_{\mathrm{Z+c}}$ (%)	$\sigma(\mathrm{Z+c})$ [pb]
$\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$	276 ± 55	0.13 ± 0.02	$10.9 \pm 2.2 \pm 0.9$
$\mathrm{Z} \to \mu^+\mu^-$	316 ± 75	0.18 ± 0.02	$8.8 \pm 2.0 \pm 0.8$
$\mathrm{Z} \to \ell^+\ell^-$	$\sigma(\mathrm{Z+c}) = 9.7 \pm 1.5\,(\mathrm{stat}) \pm 0.8\,(\mathrm{syst})\,\mathrm{pb}$		
$\mathrm{D}^{*\pm}(2010)$ mode			
Channel	$N_{\mathrm{Z+c}}^{\mathrm{signal}}$	$\mathcal{C}_{\mathrm{Z+c}}$ (%)	$\sigma(\mathrm{Z+c})$ [pb]
$\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$	151 ± 31	0.11 ± 0.01	$7.3 \pm 1.5 \pm 0.5$
$\mathrm{Z} \to \mu^+\mu^-$	247 ± 28	0.14 ± 0.01	$9.3 \pm 1.1 \pm 0.7$
$\mathrm{Z} \to \ell^+\ell^-$	$\sigma(\mathrm{Z+c}) = 8.5 \pm 0.9\,(\mathrm{stat}) \pm 0.6\,(\mathrm{syst})\,\mathrm{pb}$		
Combination			
$\mathrm{Z} \to \ell^+\ell^-$	$\sigma(\mathrm{Z+c}) = 8.8 \pm 0.5\,(\mathrm{stat}) \pm 0.6\,(\mathrm{syst})\,\mathrm{pb}$		

Samples:

- DATA: 2012 8 TeV (19.7 fb^{-1} ± 0.5)
- Signal MC: DY+jets generated w. MADGRAPH5@LO+PYTHIA6 (PDF:CTEQ6L)

σ(pp→Z+X) calculated at NNLO with FEWZ (PDF set MSTW2008NNLO)

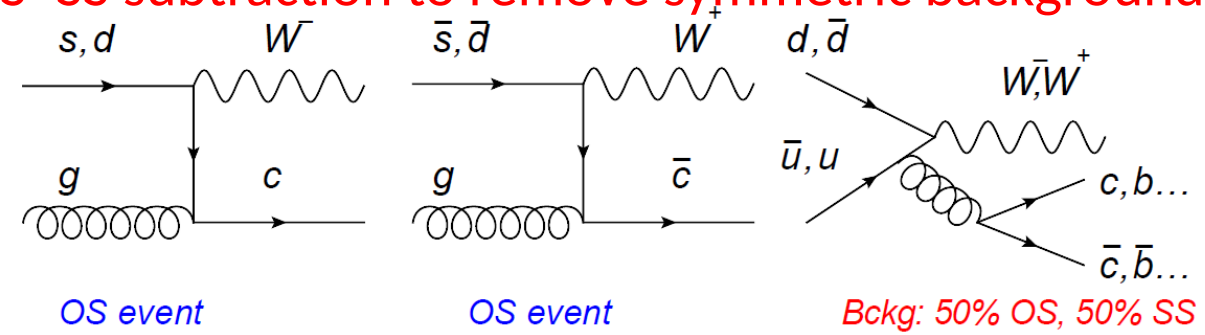
Main backgrounds:



- Contributions from $t\bar{t}$, diboson, Z+light processes (from simulations except $t\bar{t}$ from data).
- Missing transverse energy < 40 GeV (to reduce $t\bar{t}$ background).
- Data-MC differences in lepton trigger, identification and isolation efficiencies corrected (tag & probe method).Pileup events included in the MC.

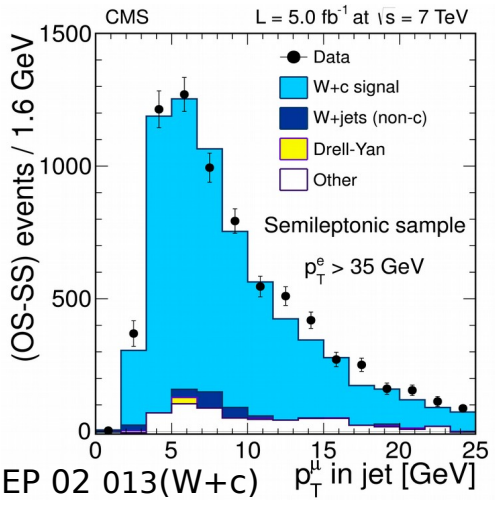
Selection of W+c sample

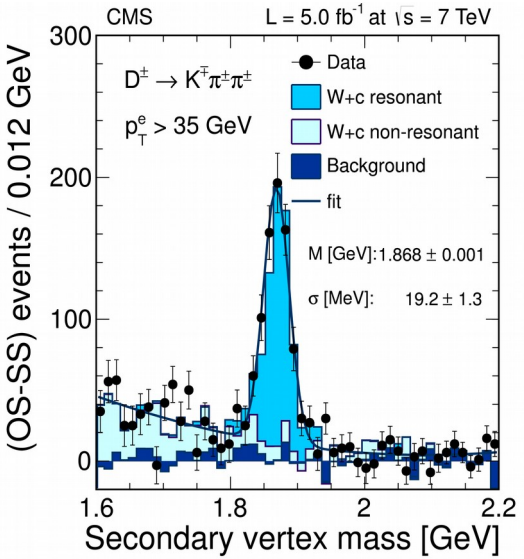
- W→ eν plus jets with similar selection to Z+HF
- Same identification of heavy flavor jet: **μ in jet or D-hadron exclusive decays**
- OS–SS subtraction to remove symmetric backgrounds

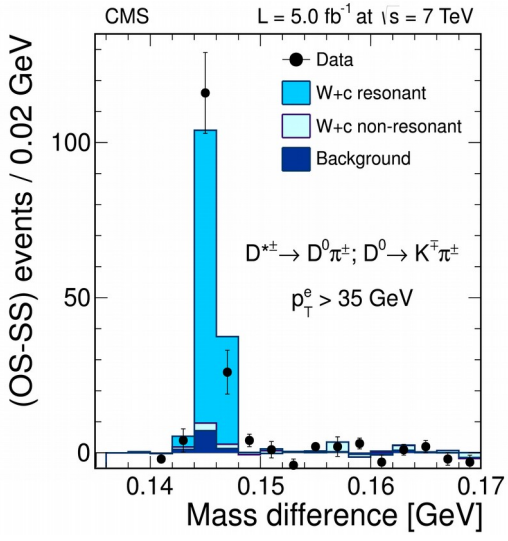


W: Total # of observed $N_{(W+D^*)data} = N_{data} - (N^{MC}_{DY} + N^{MC}_{b} + N^{MC}_{uds} + N^{MC}_{t\bar{t}})$

After OS-SS subtraction the purity in W+c of the resulting sample is > 90% (semil. channel) and >98% (D-hadron channel, D* + Wrong charge subtraction)

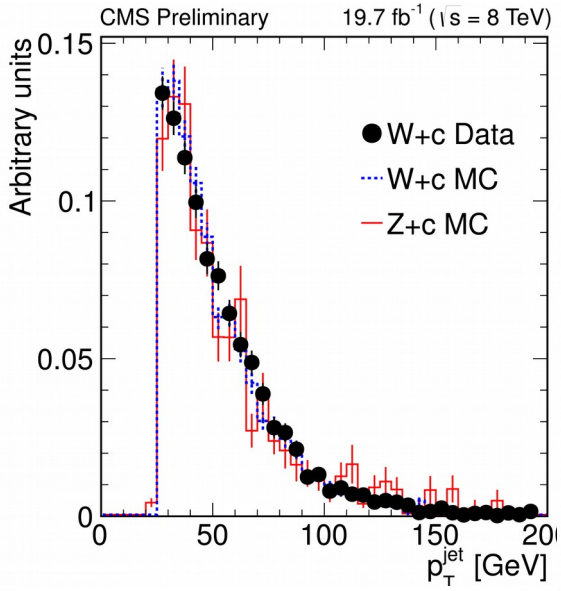


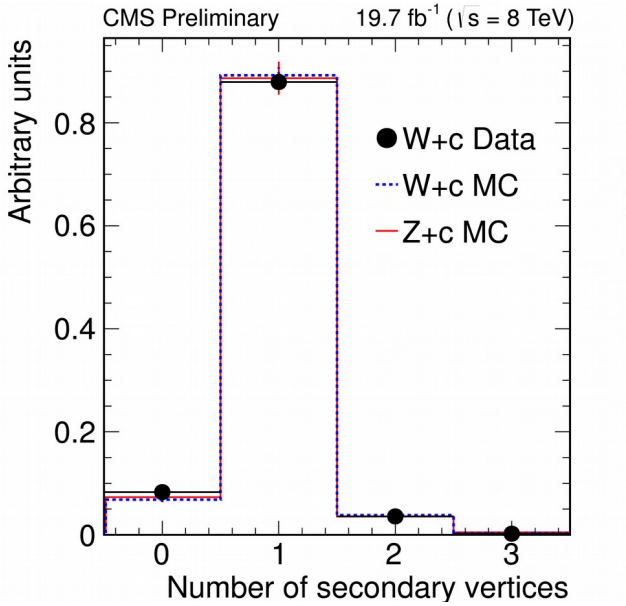


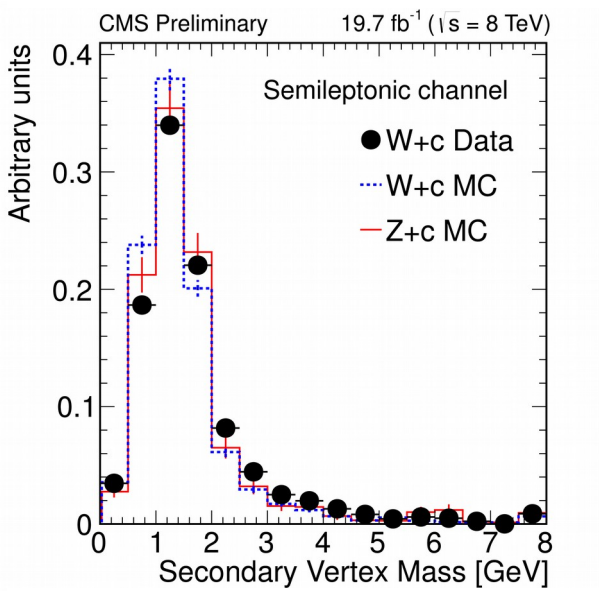


Template (shape) modeling for Z+c

Comparison of c-jets from Z+c and W+c processes
(data from W+c : after subtraction of remaining (little) background)







- *Agreement in general distributions (p_T^{jet} , N_{SV})*
- Discriminant distributions (SV-mass and JP) **W+c MC** and **Z+c MC** agree
- JP prob **W+c MC** and **W+c data** agree and validates the **Z+c MC** description
- SV-mass **W+c MC** and **W+c data** do not agree

The shape is not well modeled by the **W+c MC**. We take the shape of SV-mass from **W+c data** since there is agreement in the kinematic properties between **Z+c MC** and **W+c MC**