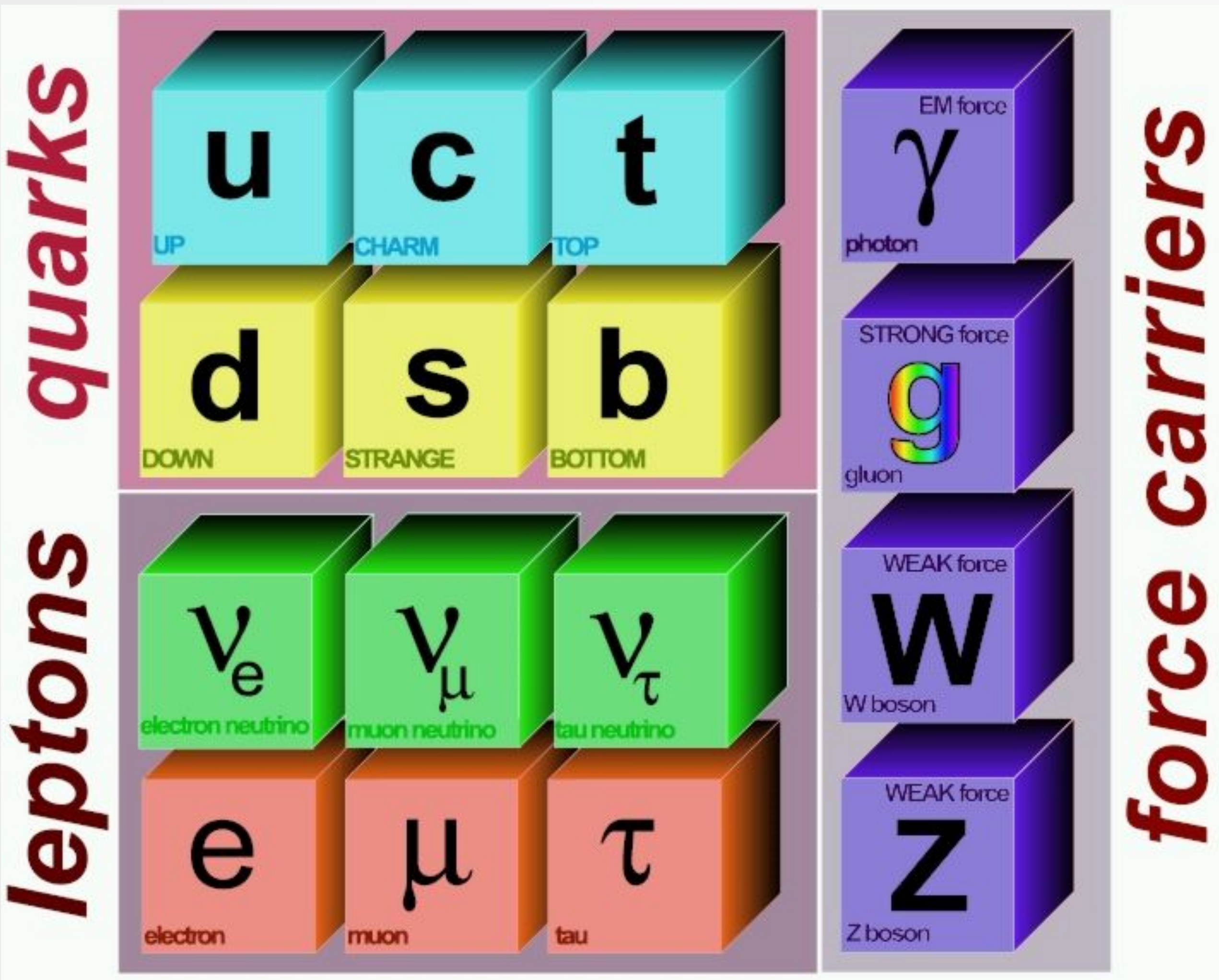
Measurement of the associated production of a W boson and a charm quark at $\sqrt{s} = 13$ TeV with the CMS detector at the LHC

Sergio Sánchez Navas

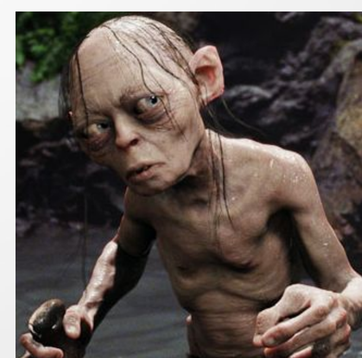
24 February 2022

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
cfp CIEMAT física de partículas

Standard Model of Elementary Particles

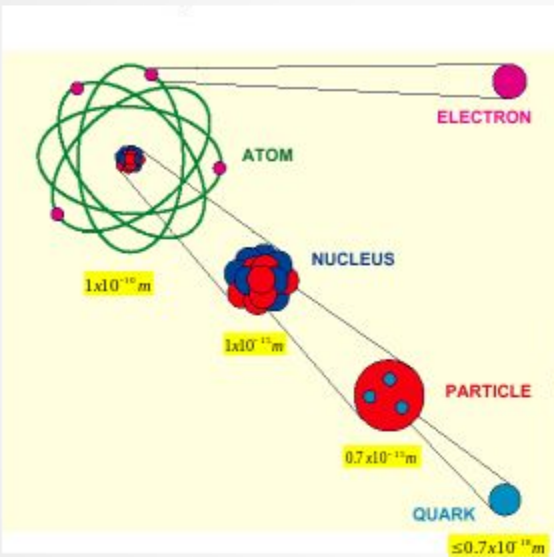


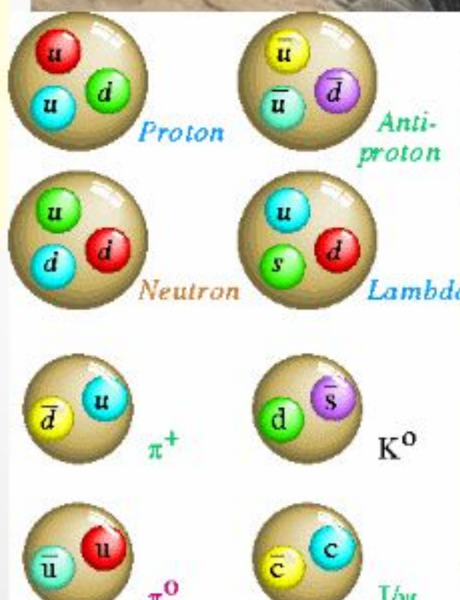
$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$
$$+ i\bar{\psi} \not{D} \psi + h.c.$$
$$+ \bar{\psi}_i y_{ij} \psi_j \phi + h.c.$$
$$+ |D_\mu \phi|^2 - V(\phi)$$



Quarks

Proposed in 1964. Elementary particles with fractionary electrical charge.

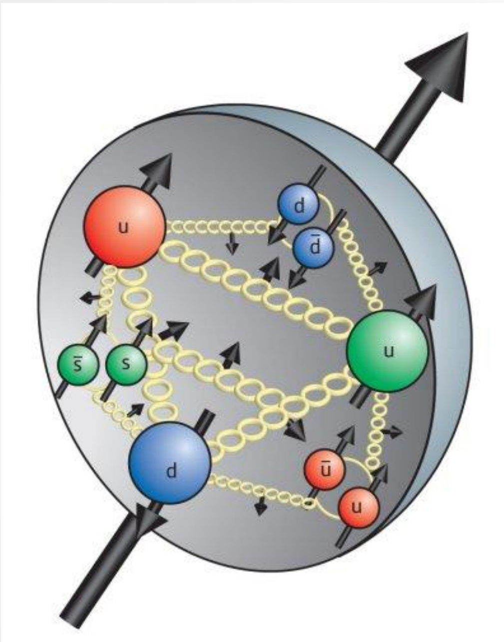




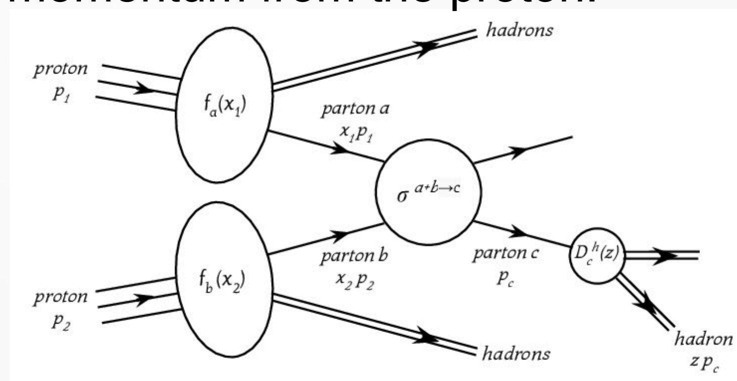
Mesons: 2 quarks
Baryons: 3 quarks

Quarks are never observed on their own but in groups forming particles called Hadrons.

Proton Structure



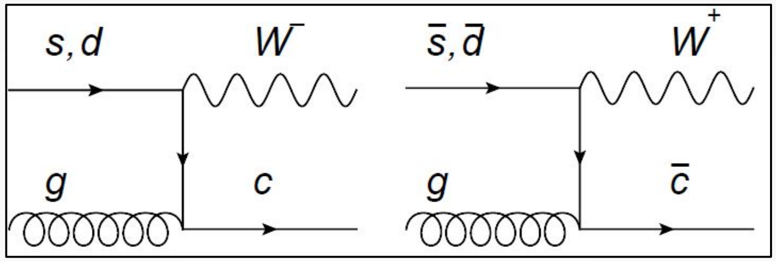
- Three valence quarks.
- Quark and gluon sea visible at high energies.
- Factorization of cross sections.
- Parton Distribution Functions (PDFs).
- Probability of finding a parton with x fraction of momentum from the proton.



a factorization scale μ_F a renormalization scale μ_R

$$\sigma_{2\to n} = \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b \, f_{a/h_1}(x_a, \mu_F)\, f_{b/h_2}(x_b, \mu_F)\, \hat{\sigma}_{ab\to n}(\mu_F, \mu_R)$$

Strange Content of the Proton



Weak interaction

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix}$$

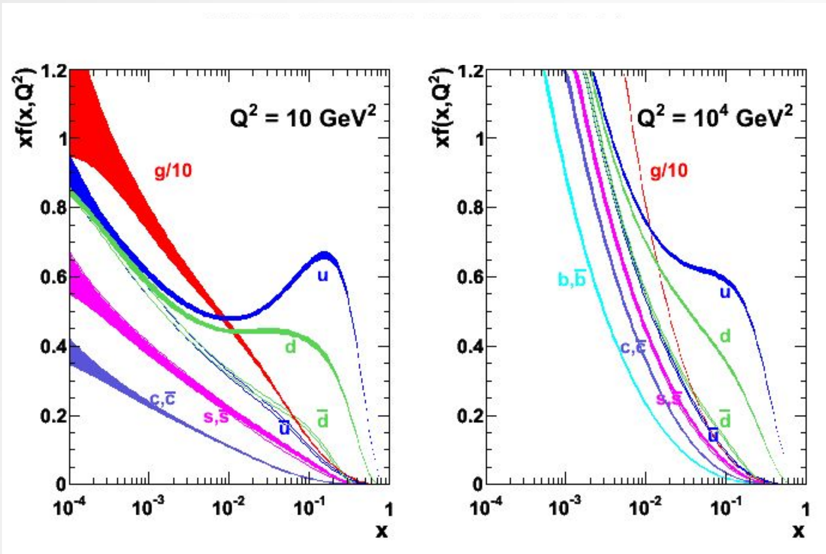
Cabibbo-Kobayashi-Maskawa (CKM) Matrix

$$\begin{bmatrix} 0.97370 \pm 0.00014 & 0.2245 \pm 0.0008 & 0.00382 \pm 0.00024 \\ 0.221 \pm 0.004 & 0.987 \pm 0.011 & 0.0410 \pm 0.0014 \\ 0.0080 \pm 0.0003 & 0.0388 \pm 0.0011 & 1.013 \pm 0.030 \end{bmatrix}$$

CKM matrix enhances the decay of s to W+c and suppresses the other channels.

This is the way to study the strange contribution to the proton.

Strange PDF

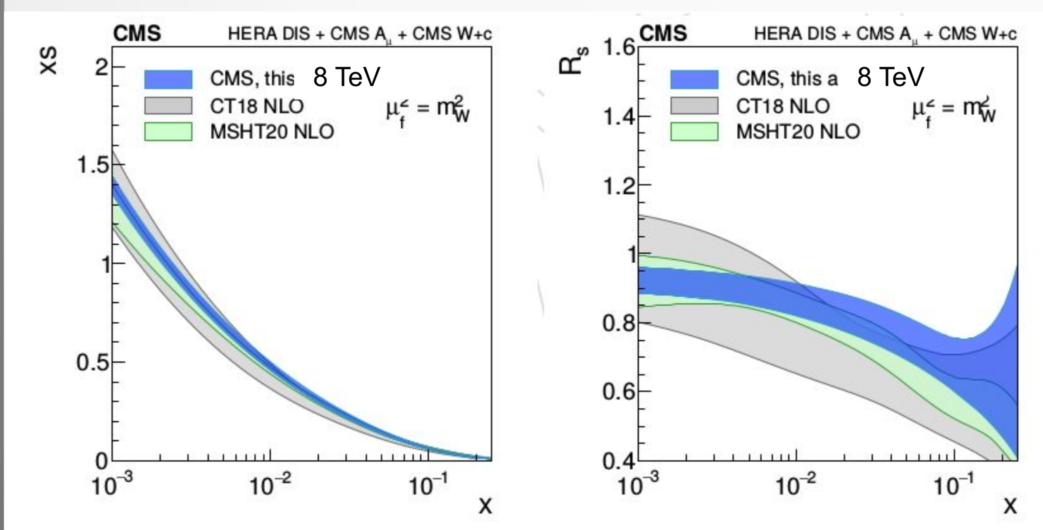


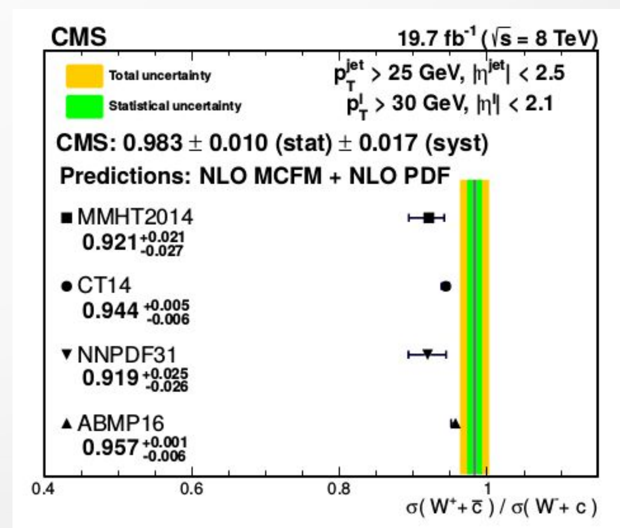
Probe on asymmetry of strange anti-strange

$R_c = \sigma(W^+ + \bar{c})/\sigma(W^- + c)$

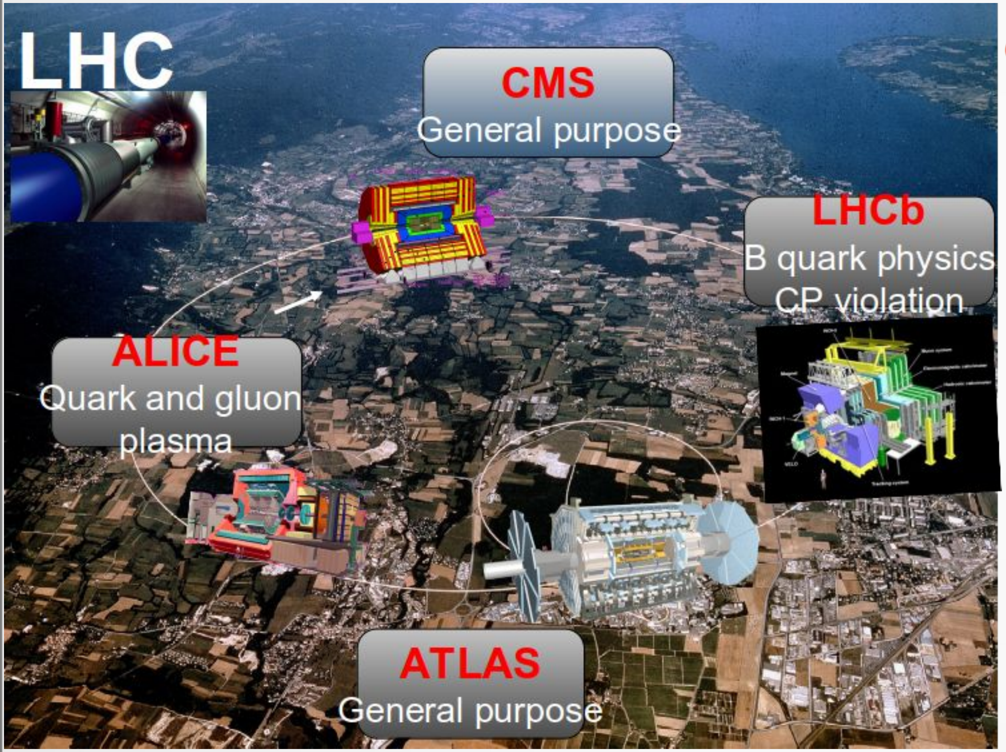
Probe on strange supression in quark sea

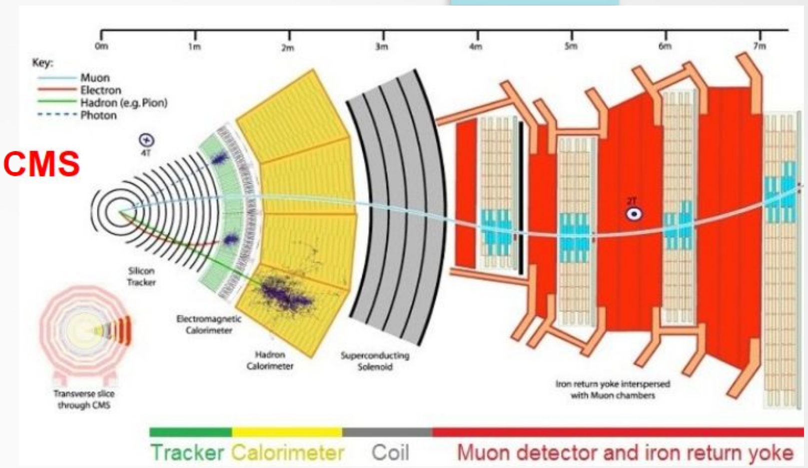
$R_s(x) = (pdf_s(x) - pdf_{\bar{s}}(x)) / (pdf_{\bar{u}}(x) + pdf_{\bar{d}}(x))$

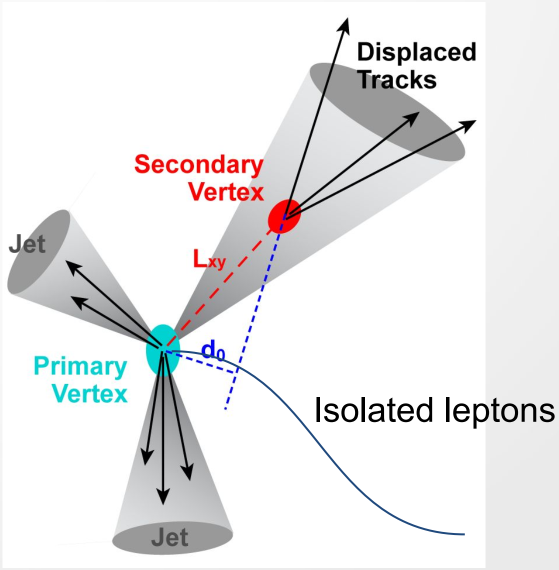




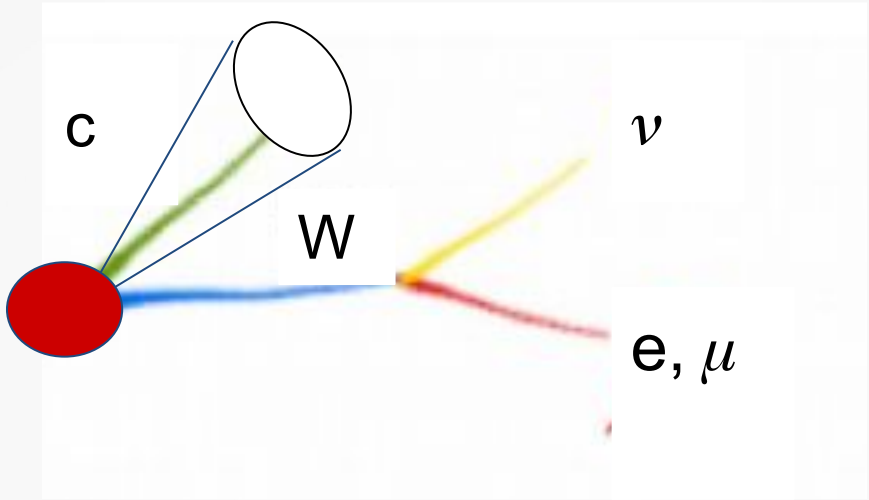
Experimental Setup







W+c Signature Signal



- One isolated high-pt electron or muon.
- Missing transverse energy due to neutrino.
- Charm jet.

Charm Identification

- Reconstruction of exclusive charm hadron decays.

  D^*, $D^{\pm}$, etc.

- Charm tagging techniques.

  b vs c and c vs light mmultivariate discriminants.

Limited statistics. Complex.
No charge information.

- Inclusive decays:

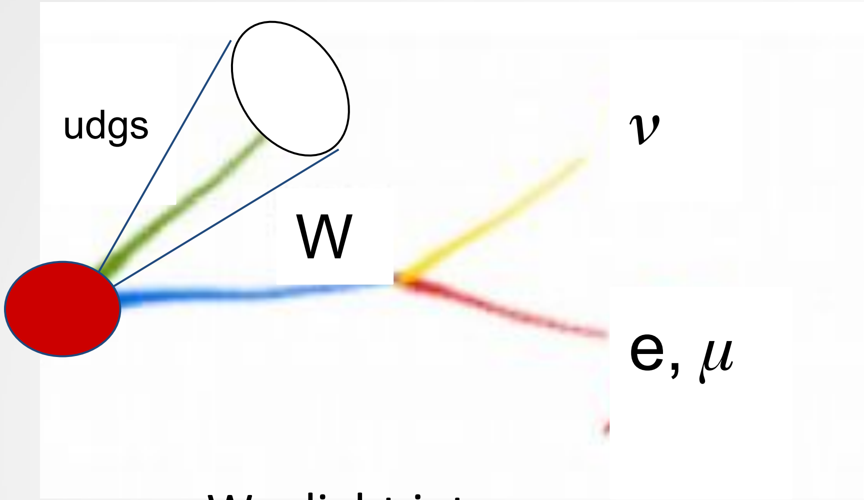
This analysis

  - Charm identification (**SV channel)**: Hadronic decays of charm hadrons. Secondary vertex in jet, electric charge different from zero and displaced from the primary vertex.
  - Charm identification (**SL channel)**: Semileptonic decays of charm hadrons. Muon inside a jet.

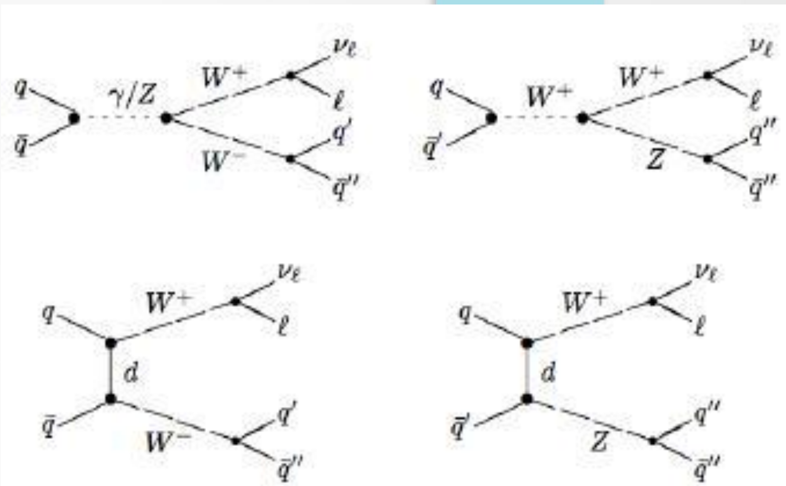
Increased statistics.
Charge identified.

➔ Events with muon in jet and SV fall into the SL channel.

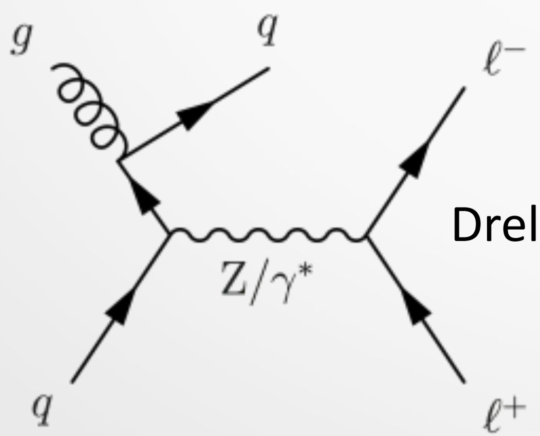
Main Backgrounds I



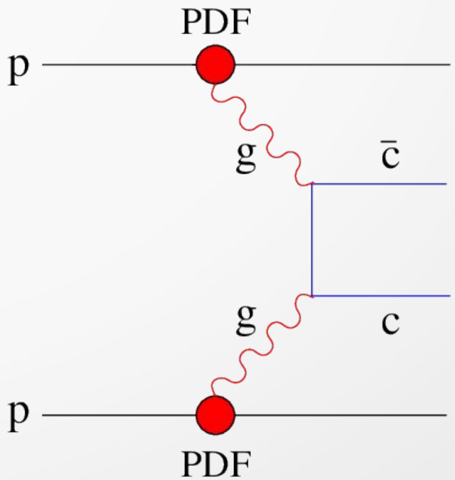
W + light jets



dibosons

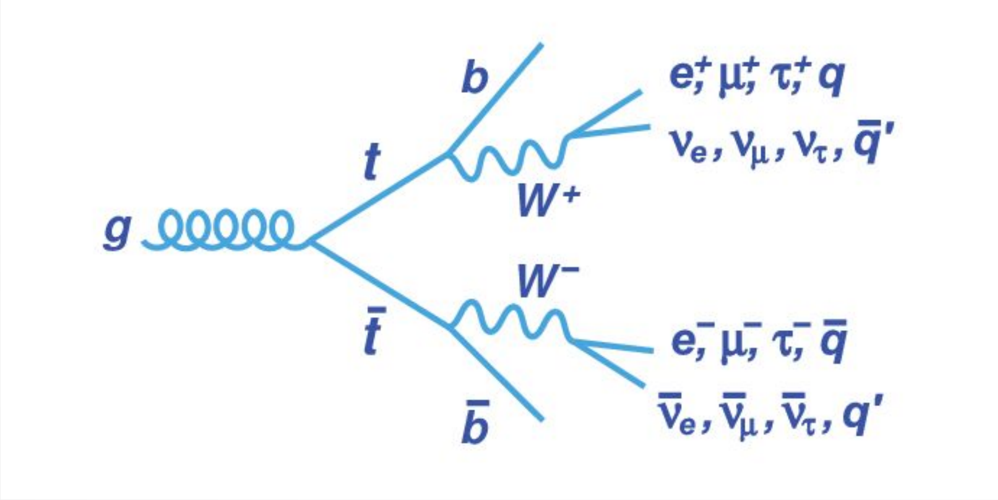


Drell-Yan + jets

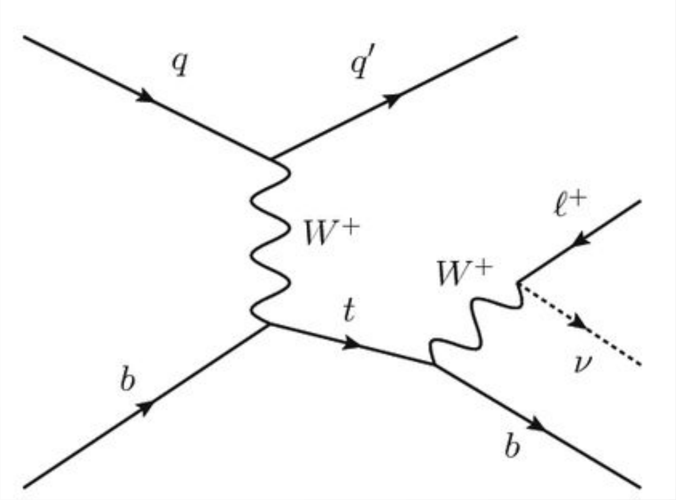


multijet QCD production

Main Backgrounds II



top-antitop



single top

Event selection

Run II (2016, 2017, 2018) 137 fb^{-1}

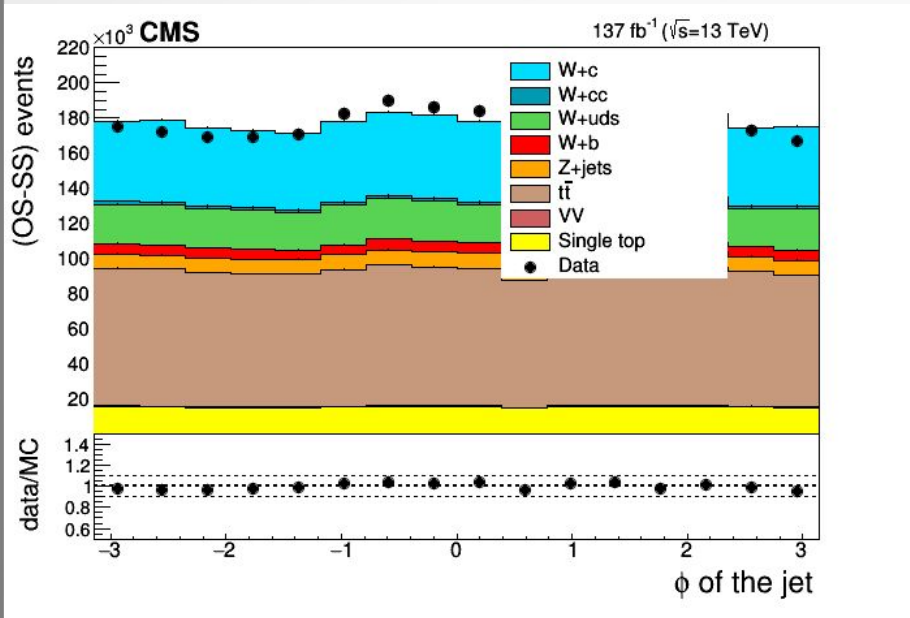
Analysis done with $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ plus c-jets:

- $W \rightarrow l+\nu$:
  - $p_T(l) > 35$ GeV,
  - $|\eta(l)| < 2.4$
  - Isolated
    - Exclude events with two isolated leptons with opposite sign to reduce ttbar background.
  - $M_T > 55$ GeV

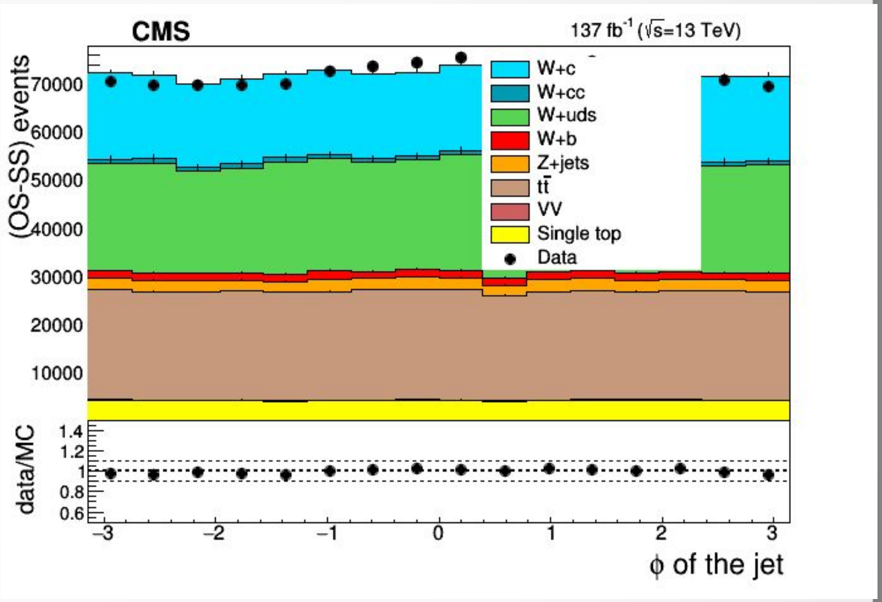
- Jets:
  - $p_T(jet) > 30$ GeV, $|\eta(jet)| < 2.4$
  - SL: $p_T(\mu) < 25$ GeV, $|\eta(\mu)| < 2.4$ non isolated.

Require minimum angular separation between jet and lepton ($\Delta R > 0.4$)

OS-SS technique. Before substraction



SV

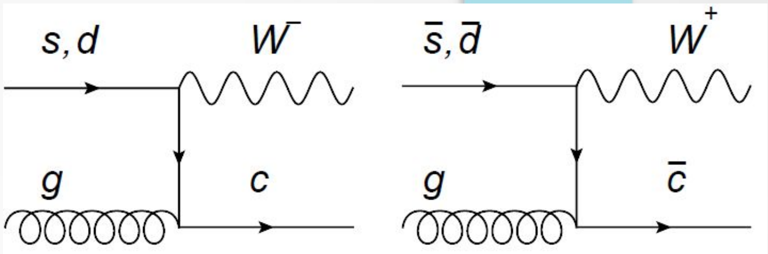


SL

We want to reduce the backgrounds to measure a precise cross section.

OS-SS technique

- Key ingredient to substract most of the background

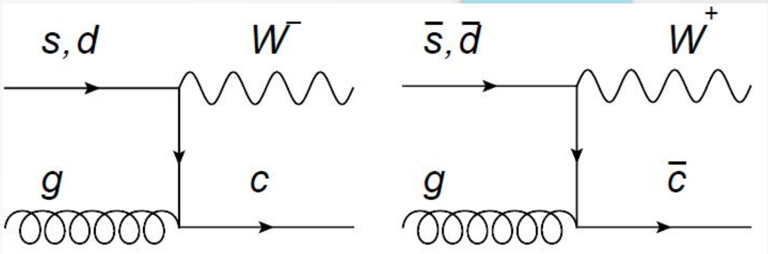


- In W+c events, the electric charge of the W is opposite to the charge of the c quark.
  - Charge of the W reconstructed from the charge of the isolated lepton.
  - Charge of the charm quark from the charge of the muon inside the jet or the SV charge.

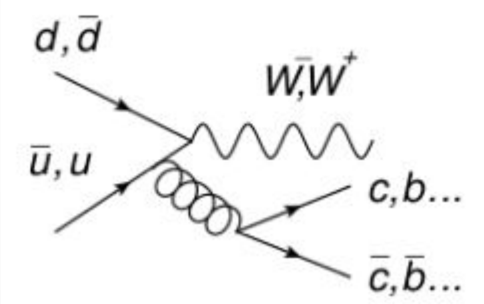
- In most of the background, there are equal amounts of SS and OS events with **identical kinematic properties**.

OS-SS technique

- Key ingredient to substract most of the background



- Some of the backgrounds difficult to simulate precisely, e.g., gluon splitting to c-cbar => completely removed with OS-SS subtraction.

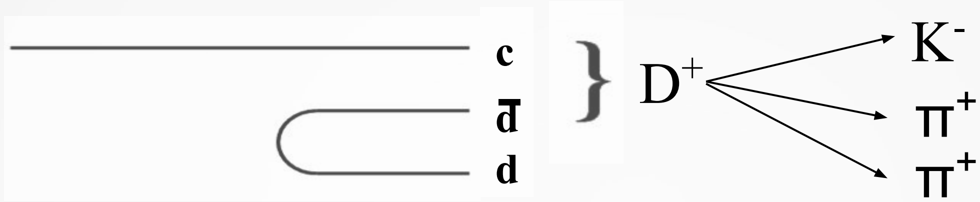


gluon splitting

- OS-SS selects charm produced in the hard interaction coming from the strange quark in the proton => enhanced sensitivity.

OS-SS technique. SV charge determination

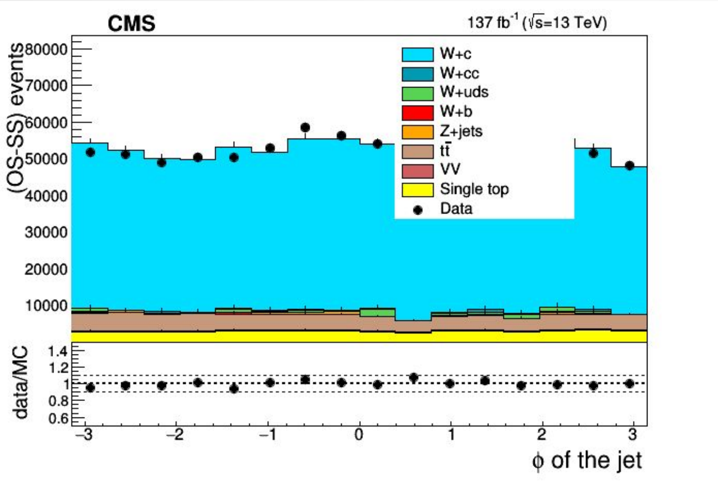
- Charge of the SV vertex : $\sum^{\text{Vertex-tracks}} q_{\text{tracks}}$
  - **c**

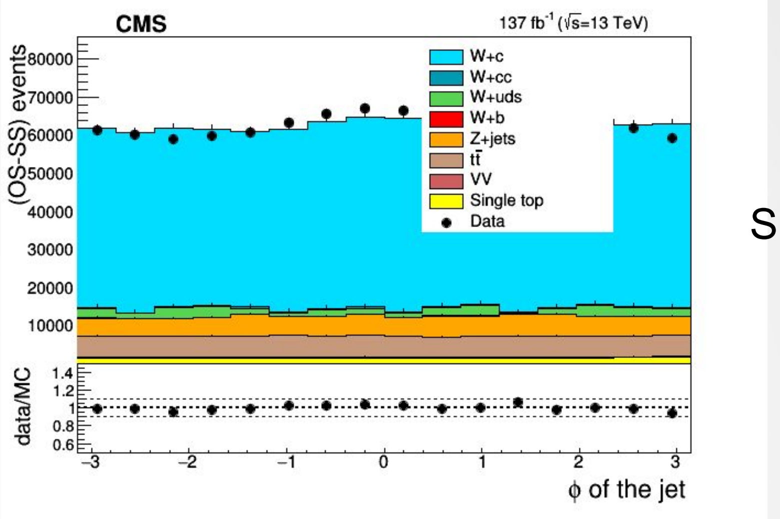


- Charge definition: OS if charges of the SV and the lepton from the W decay are opposite.
  - OS : $Q_{W \to e(\mu)} \neq Q_{SV}$ SS : $Q_{W \to e(\mu)} = Q_{SV}$
  - If no charge identified → event is removed

OS-SS technique. After substraction

SV



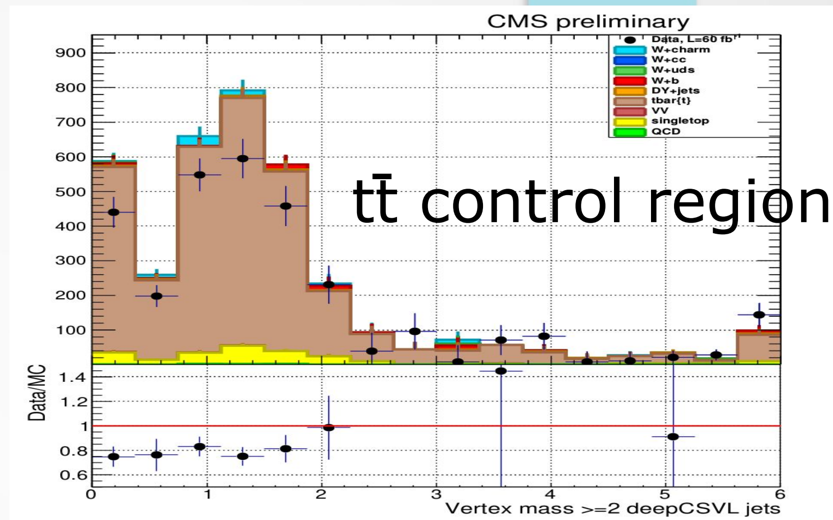


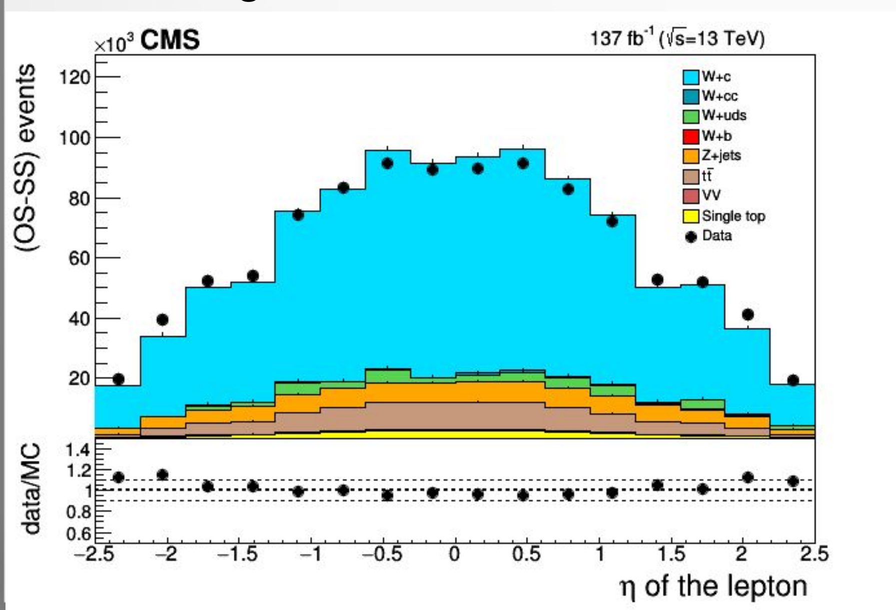
SL

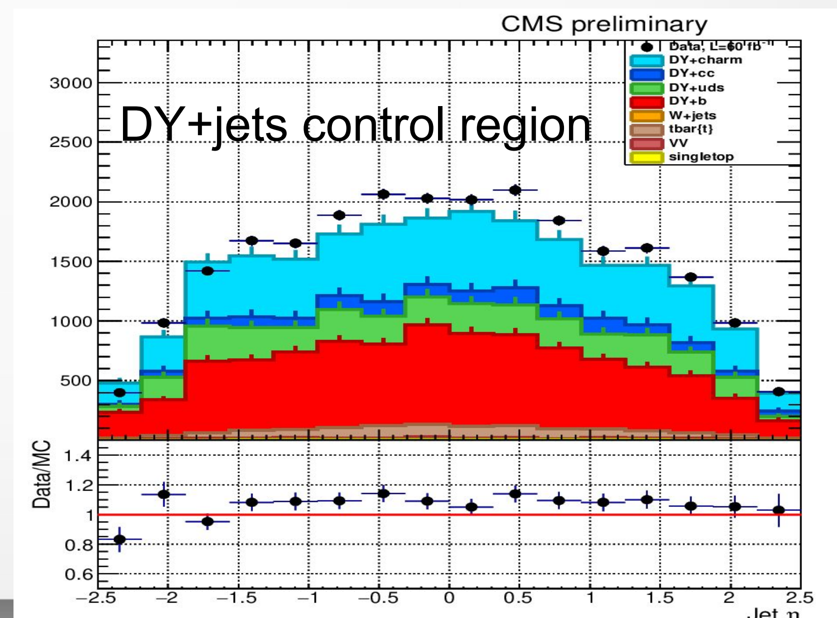
- Apply a statistical weight of -1 to SS events.
- The remaining backgrounds are truly OS processes:
  - ttbar: one W from top decays semileptonically and the other into charm-strange quarks.
  - DY: Z+jets (only in W->mu SL channel)
  - Single top
  - W+udsg: ug $\rightarrow W^+$ +d, dg $\rightarrow W^-$ +u

Evaluation and normalization of dominant backgrounds

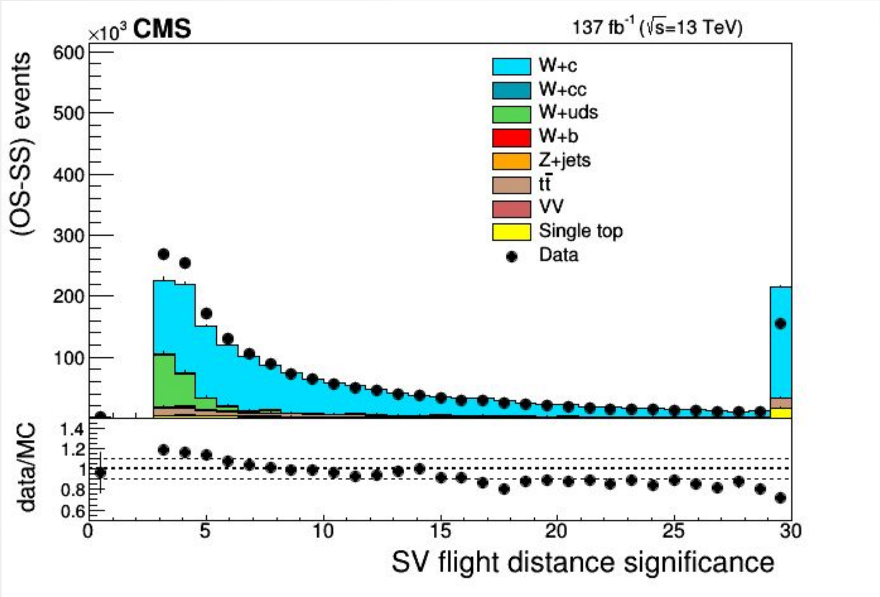
<u>ttbar</u>: We require 3 or more jets satisfying selection cuts and at least two of them with btagging. ~20%
<u>Drell-Yan</u>: We require two isolated leptons of opposite charge and same flavour and invariant mass cut around the Z peak (71-110GeV). ~13%
<u>W+jets</u>: Last we normalize to data and find W+light.

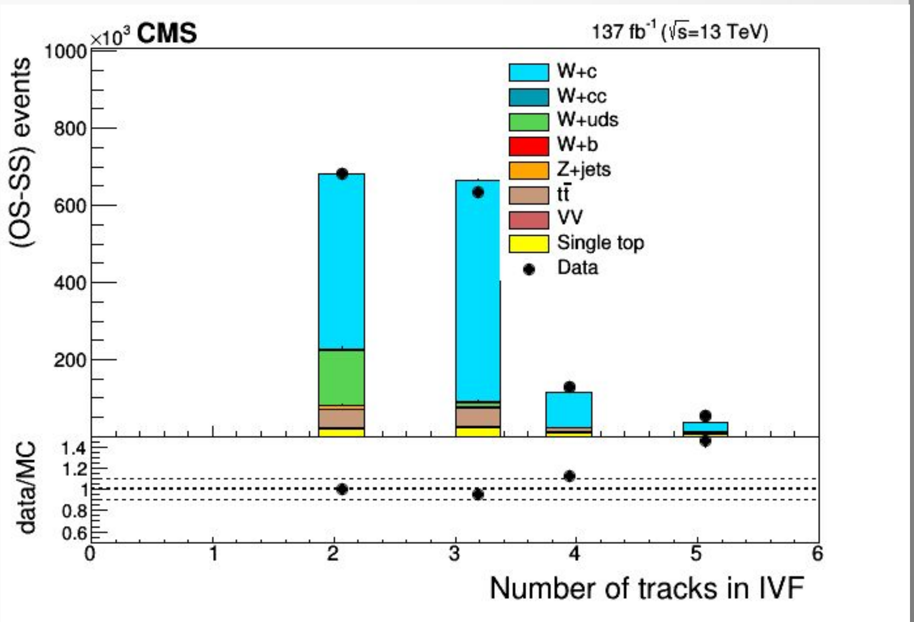






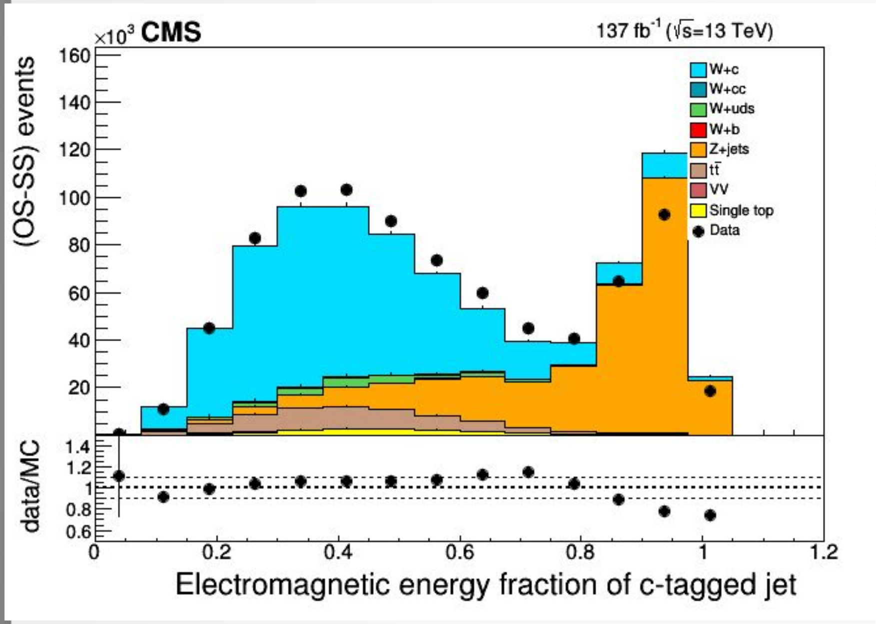
Removing W+light (usdg) background in SV channel



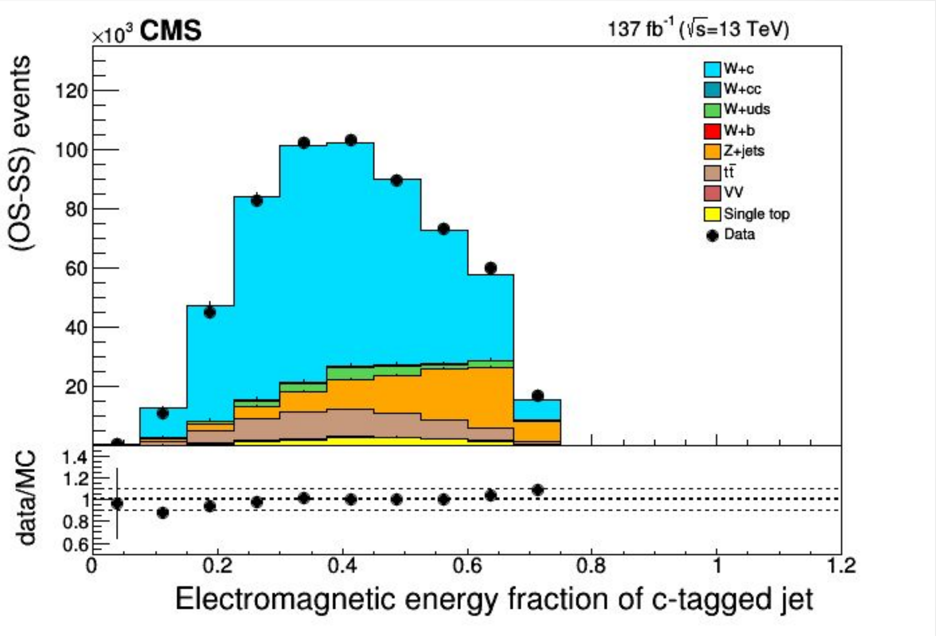


Light background is concentrated at low SV flight distance significance and in 2-track vertices.

Reducing DY background in SL channel (W→ mu)



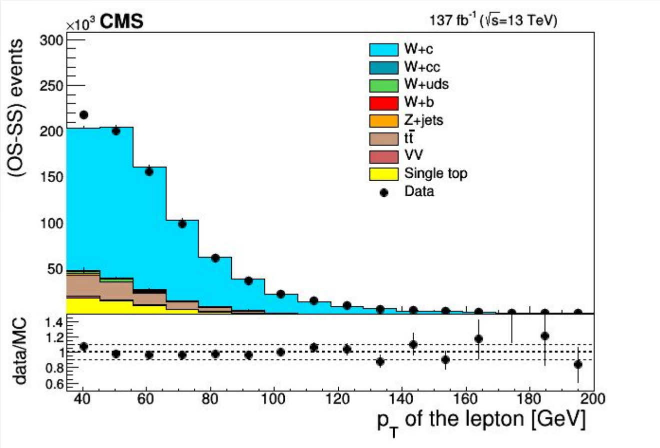
Before cut

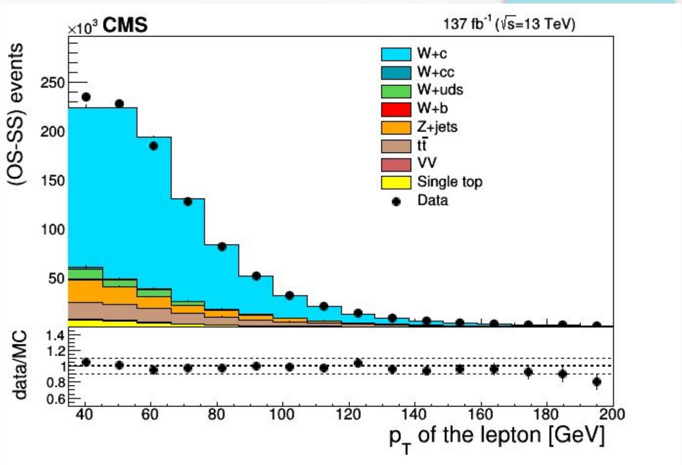


After cut

(Muon+Neutral EM) energy fraction of jet < 0.7

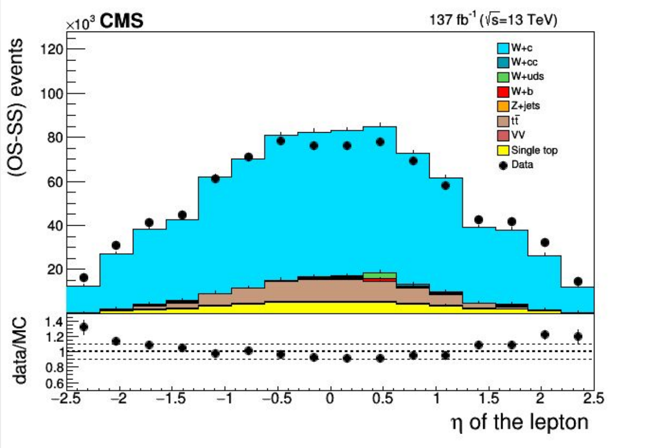
Control Variables. Isolated Lepton

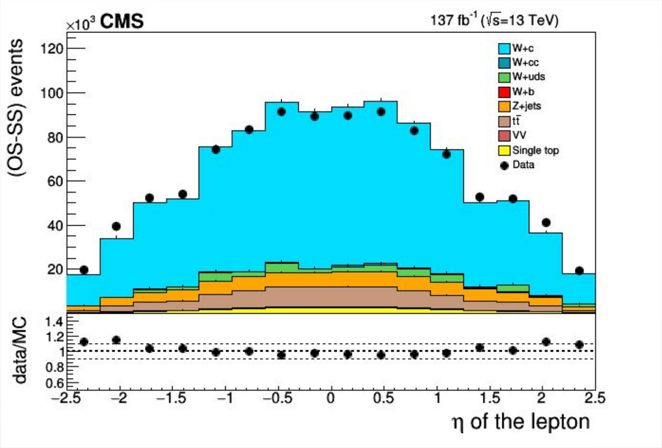




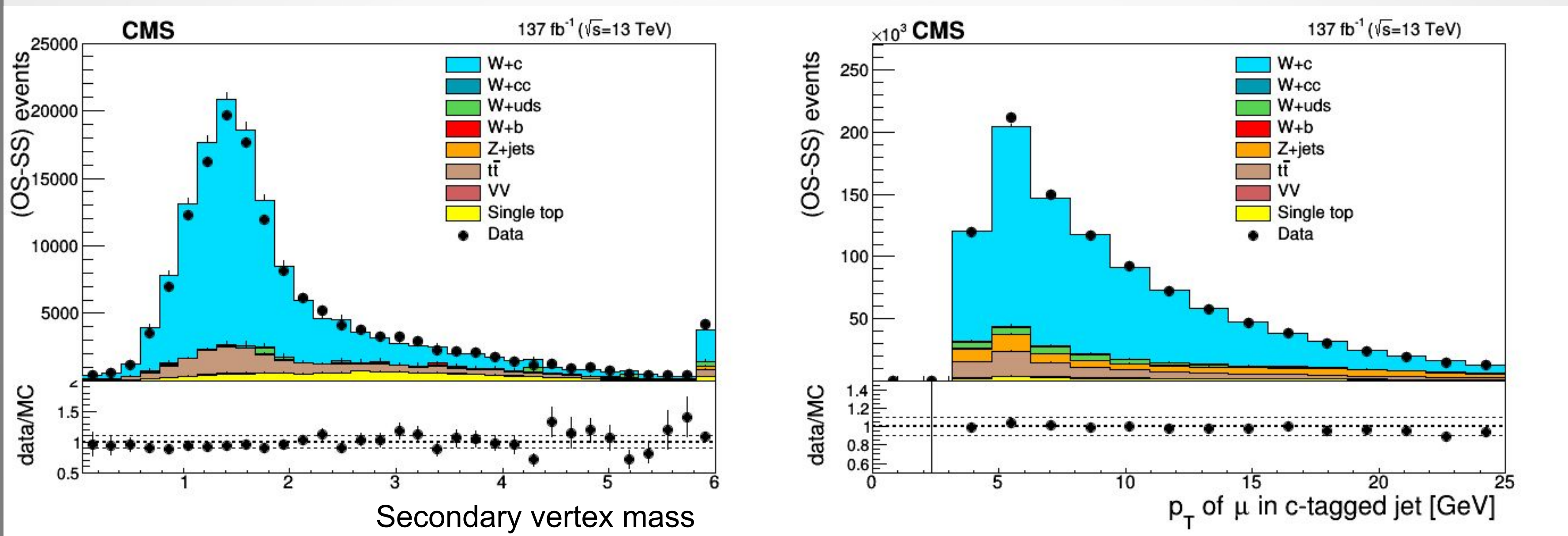
SV channel

SL channel





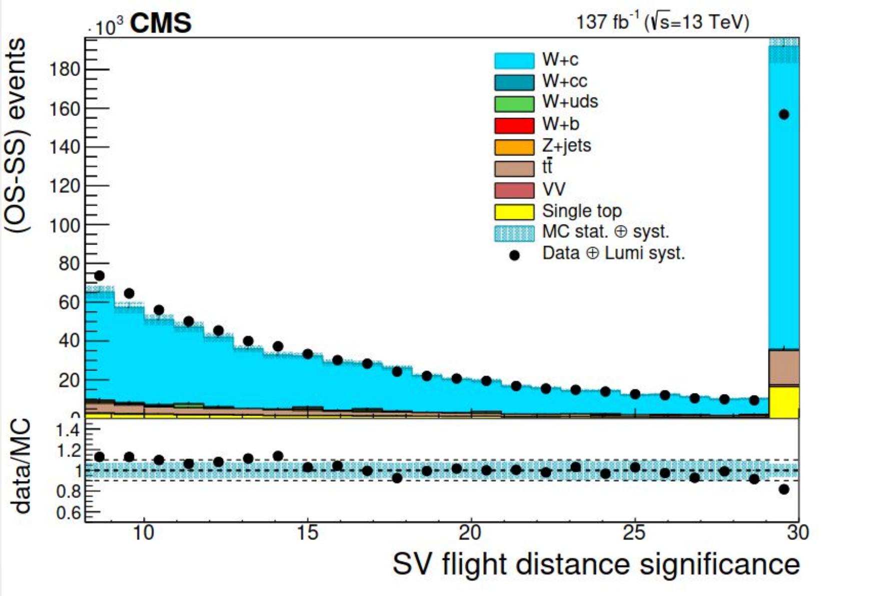
Control Variables. Charm Jet

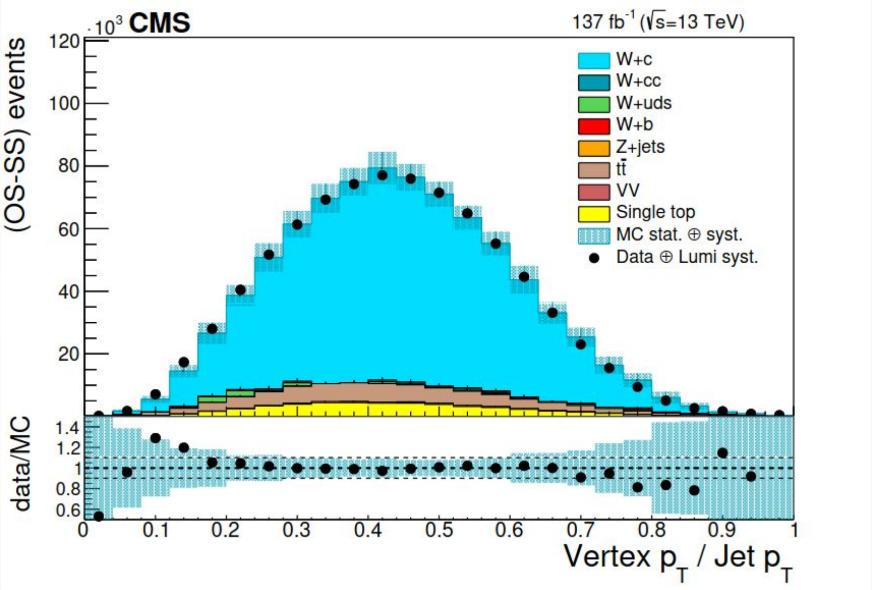


SV channel

SL channel

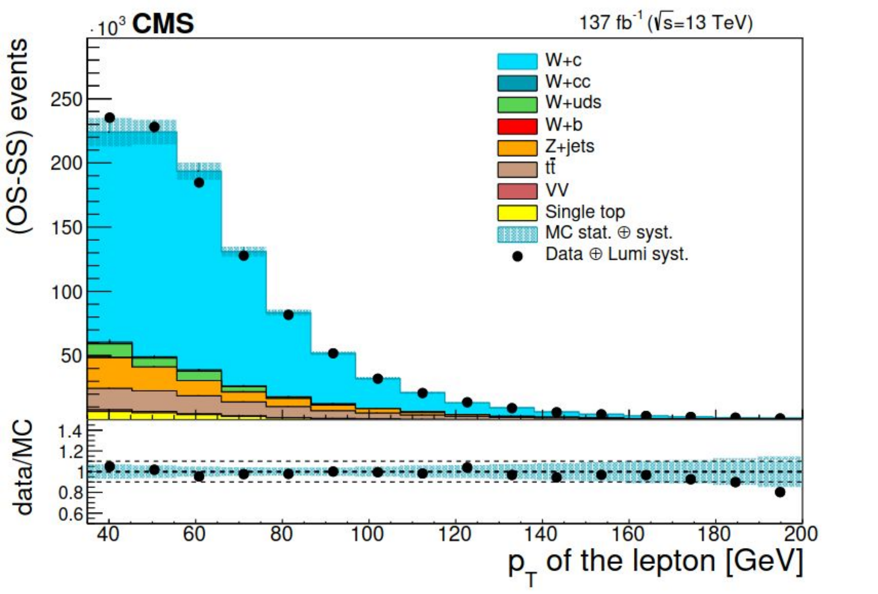
SV Channel Signal Fractions

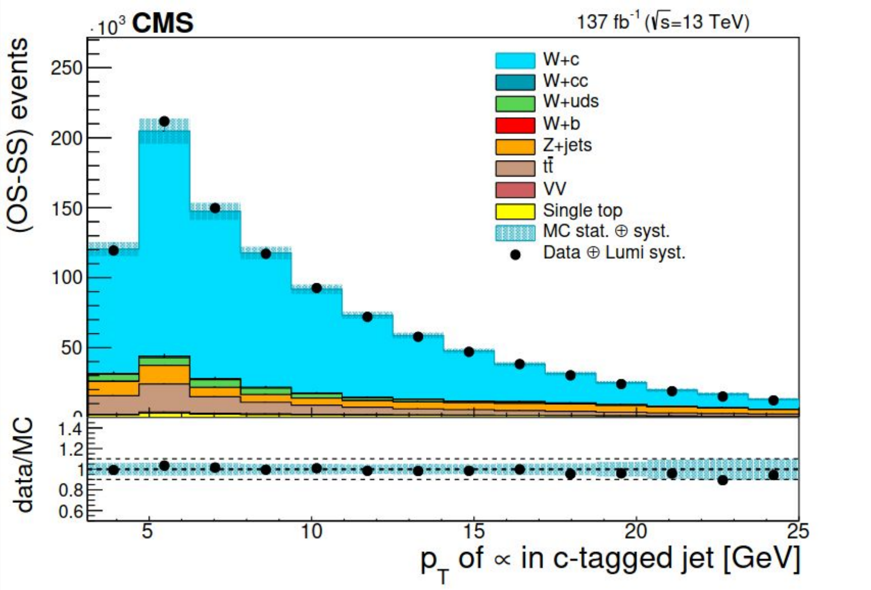




W+c	$W+Q\bar{Q}$	W + udsg	DY	$t\bar{t}$	single top	VV
83.9 ± 0.5	0.6 ± 0.2	0.9 ± 0.4	0.3 ± 0.1	8.0 ± 0.1	5.7 ± 0.1	0.6 ± 0.1

SL Channel Signal Fractions





W+c	$W+Q\bar{Q}$	W + udsg	DY	$t\bar{t}$	single top	VV
78.7 ± 0.4	-0.2 ± 0.1	4.7 ± 0.3	6.2 ± 0.1	8.0 ± 0.1	2.2 ± 0.1	0.4 ± 0.1

C-tagging efficiencies in SV determination

- Evaluated using W+c MC simulation events
- Determination of SV efficiency data/MC scale factor (SFc) using a data control sample: SL sample with an additional SV. The muon inside of a jet identifies the charm and we compare with the reconstructed SV.

$$\epsilon_c^{data} = \frac{N_{W+c}^{tag,\, data}}{N_{W+c}^{data}} \quad \epsilon_c^{sim} = \frac{N_{W+c}^{tag,\, sim}}{N_{W+c}^{sim}} \quad SF_c = \frac{\epsilon_c^{data}}{\epsilon_c^{sim}}$$

- Dedicated SFc for vertices with 3 or more tracks and flight distance significance > 8.

- SF_c depending on year and reconstructed vertex type. [0.94 – 1.0]± 0.03

Cross Section Measurement.

$$\sigma(\mathrm{W+c}) = \frac{N_{\mathrm{sel}}(1 - f_{\mathrm{bkg}})}{\mathcal{C}\,\mathcal{L}}$$

$$\mathcal{A} \times \epsilon_c^{data} = \frac{N_{\mathrm{W+c}}^{sel}}{N_{\mathrm{W+c}}^{gen}}$$

Fiducial region. Phase space that corresponds reasonably closely to what is experimentally accessible. Defined by pt-η of isolated lepton and pt-η of c-jet.

Cross section measurements at two levels:

- Particle or Hadronic level. Most similar to detection, after hadronization and fragmentation. Comparison with MC generators with hadronic jets.
- Parton level. Parton jets, have to trust simulated fragmentation and hadronization. Comparison with QCD calculations with different PDFs.

Cross Section Measurements.

We are presenting 6 measurements for each level:

- Integrated cross section:
  - $\sigma(W+c)$
  - $R_c = \sigma(W^+ + \bar{c})/\sigma(W^- + c)$
- Differential cross section (wrt p_T or η of lepton from W decay):
  - $d\sigma/dp_T$
  - dR_c/dp_T
  - $d\sigma/d\eta$
  - $dR_c/d\eta$

Systematic uncertainties

- **Isolated Lepton** trigger, identification and reconstruction efficiency uncertainties. Soft muon in SL channel as well. ≈ 1%
- **Jet Energy Scale and Resolution.** (<1 – 2%)
- **Missing transverse energy.** < 1%
- **Charm hadron branching ratios and fragmentation fractions.** ≈2% SV ≈1% SL
- **Fragmentation function.** < 1%
- **PDF.**≈ 1%
- **Pile Up** profile. < 1%
- **Flight distance significance of the SV.** ≈ 1%
- **c-tagging.** 3%
- **Monte Carlo** statistics. (Dominant for R_c) ≈ 1%
- **Luminosity.** 1.6%

Particle or Hadronic Level cross section measurements

- Identify charm genjets in the signal MC generator inside the fiducial region.

  pt_genjet>30 GeV, |η_genjet|<2.4, pt_leptonW>35 GeV, |η_leptonW|<2.4

  dR(charm_genjet, leptonW)>0.4

- Cross section measurement is restricted to W+c processes where the genjet is matched to the c quark coming from a strange quark in the proton.

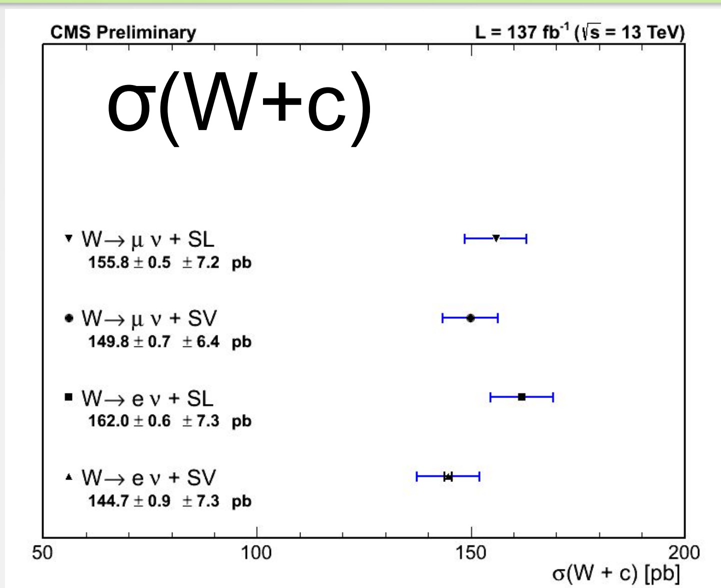
Acceptance x Efficiency (%)

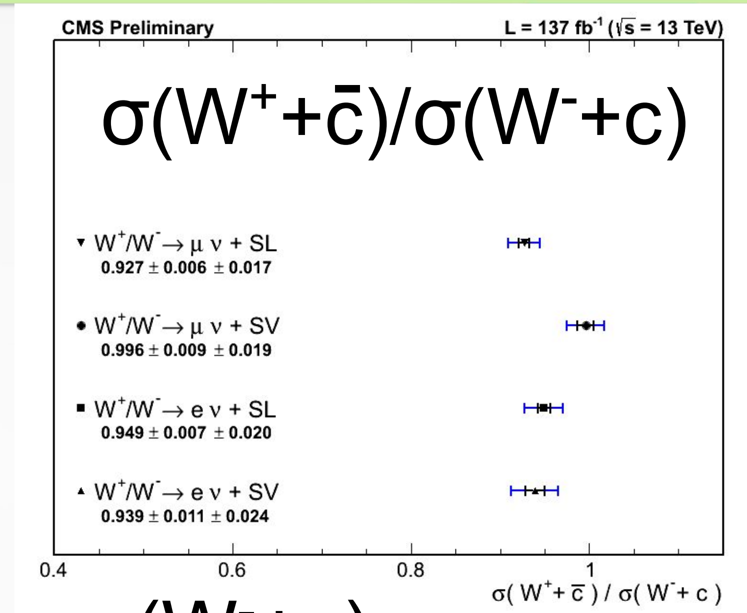
$$\mathcal{A} \times \epsilon_c^{data} = \frac{N_{W+c}^{sel}}{N_{W+c}^{gen}}$$

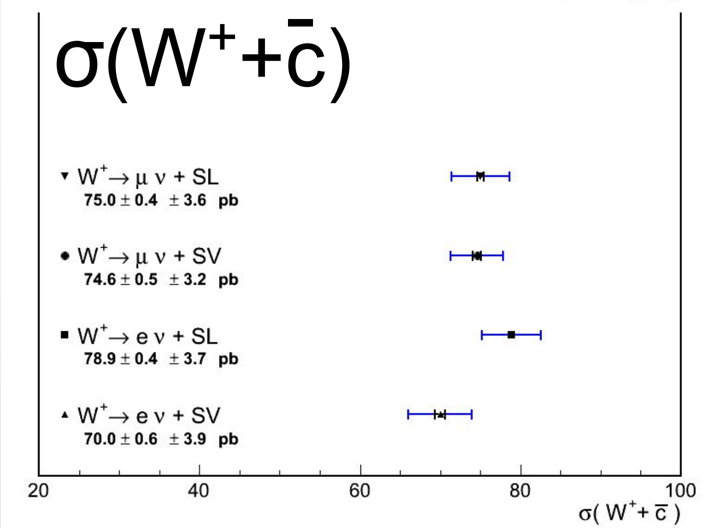
Run II	SL	SV
ELECTRON	1.563±0.028	1.430±0.027
MUON	0.946±0.021	1.966±0.031

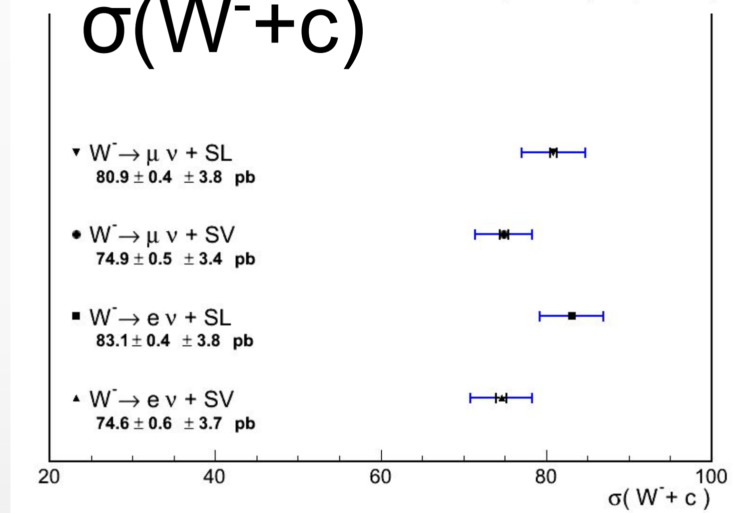
Run II obtained here using datasets from the three years together and weighting by the corresponding luminosity. Statistical uncertainty only.

Cross section σ(W+c) (pb) Run II









Combined cross section and cross section ratio RunII

$$\sigma(\mathrm{W}+\mathrm{c}) = \frac{N_{\mathrm{sel}}(1-f_{\mathrm{bkg}})}{\mathcal{C}\,\mathcal{L}}$$

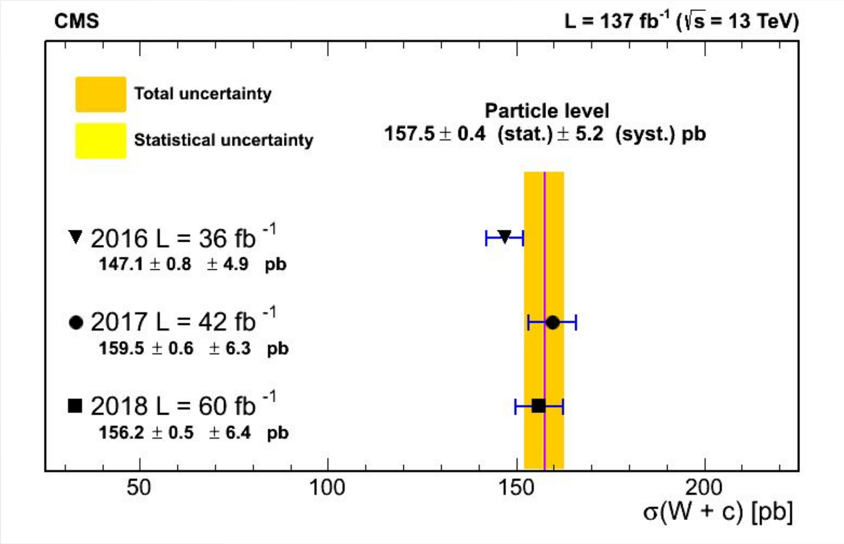
$\sigma(W+c) = 157.5 \pm 0.4$ (stat.) ± 5.2 (syst.)

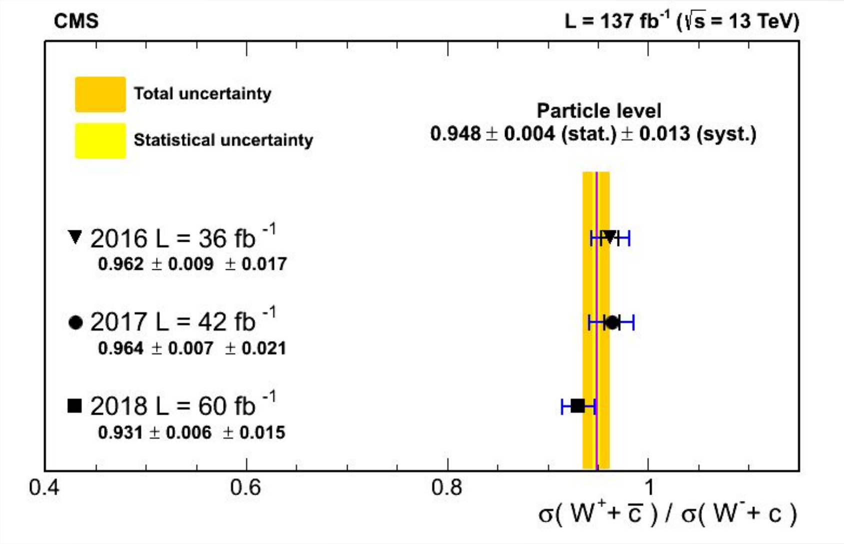
$\sigma(W^{+}+\bar{c})/\sigma(W^{-}+c) = 0.949 \pm 0.004$ (stat.) ± 0.013 (syst.)

Many systematic uncertainties cancel out in cross section ratios and are significantly reduced in integrated measurement but the limited MC sample size dominates the result.

Cross section integrated over fiducial region

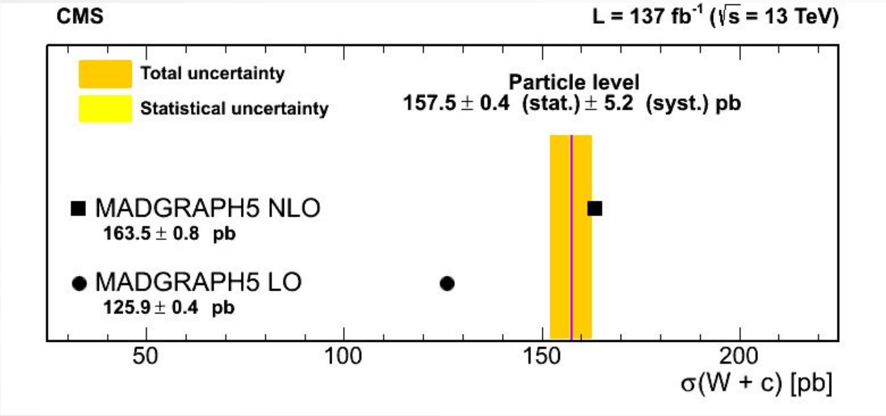
Comparison with measurements with 2016, 2017 and 2018 data

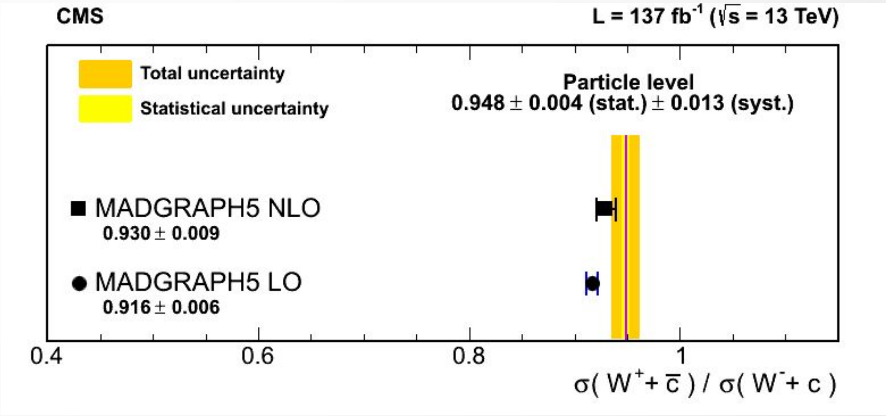




Cross section integrated over fiducial region

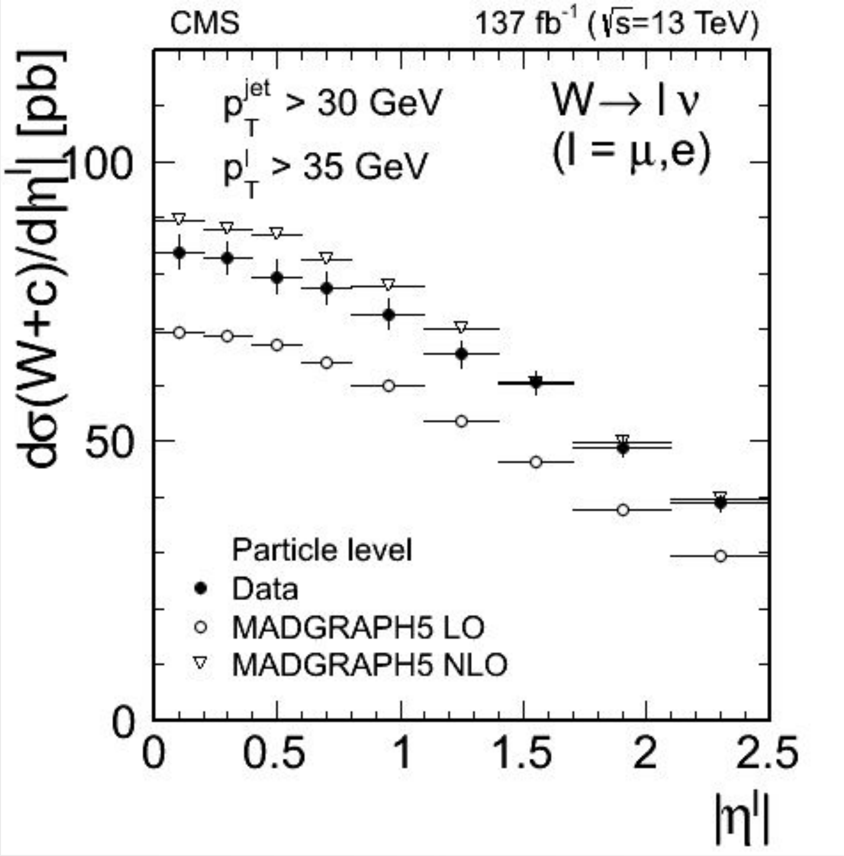
Comparison with MADGRAPH prediction at LO and NLO accuracies

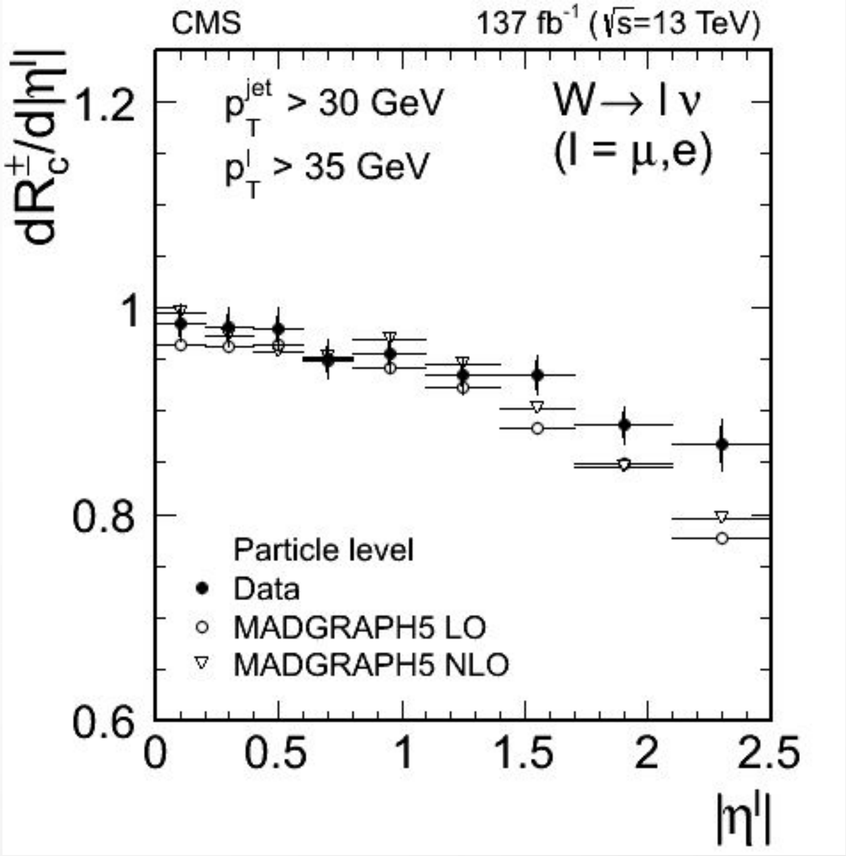




Combination of differential cross section vs eta W-lepton RunII

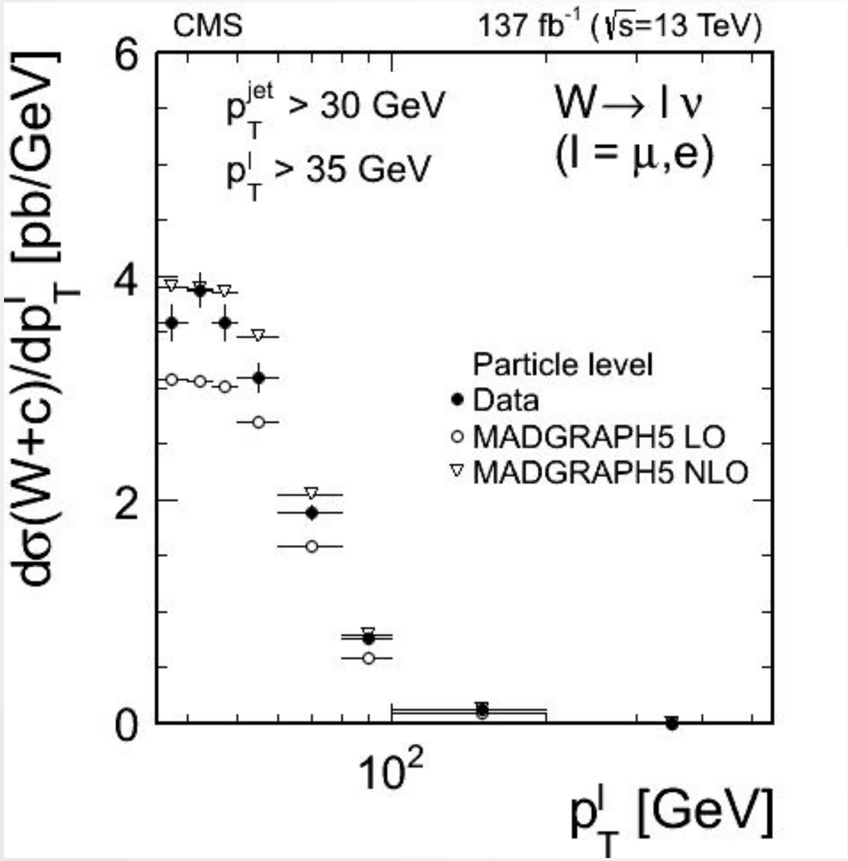
Comparison with MADGRAPH predictions at LO and NLO accuracies

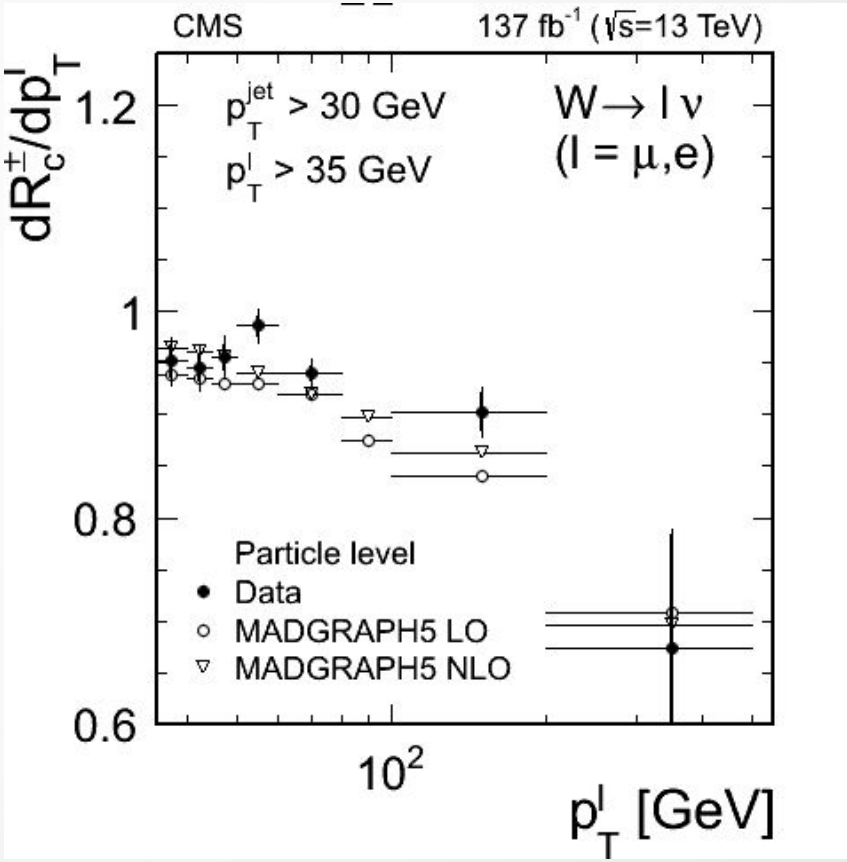




Combination of differential cross section vs pt W-lepton RunII

Comparison with MADGRAPH predictions at LO and NLO accuracies





Parton Level cross section measurements

- Identify charm parton jets instead bare partons for a more meaningful comparison with theoretical calculations.
- Then require the c-parton jet to be inside the fiducial region.

  pt_charm>30 GeV , |η_charm|<2.4, pt_leptonW>35 GeV, |η_leptonW|<2.4, dR(charm, leptonW)>0.4

- We perform this measurement to compare with NLO theoretical calculations (MCFM) and potentially proceed with a QCD analysis to extract the strange PDF, as done previously in the W+c @ 8 TeV analysis.

Acceptance x Efficiency (%)

$$\mathcal{A} \times \epsilon_c^{data} = \frac{N_{W+c}^{sel}}{N_{W+c}^{gen}}$$

Run II	SL	SV
ELECTRON	1.419±0.025	1.261±0.024
MUON	0.856±0.019	1.786±0.028

Run II obtained here using datasets from the three years together and weighting by the corresponding luminosity. Statistical uncertainty only.

Combined cross section and cross section ratio RunII

$$\sigma(\mathrm{W}+\mathrm{c}) = \frac{N_{\mathrm{sel}}(1 - f_{\mathrm{bkg}})}{\mathcal{C}\,\mathcal{L}}$$

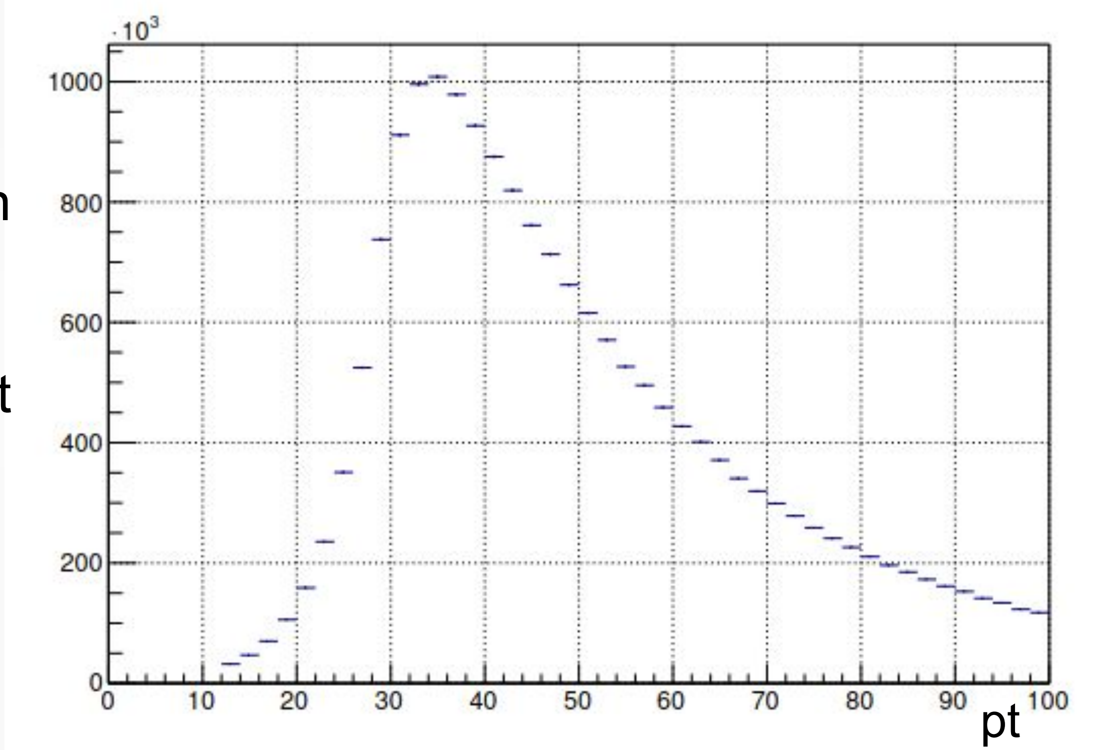
$\sigma(W+c) = 163.5 \pm 0.5$ (stat.) $\pm\ 6.1$ (syst.)

$\sigma(W^{+}+\bar{c})/\sigma(W^{-}+c) = 0.953 \pm 0.005$ (stat.) $\pm\ 0.014$ (syst.)

Comparison between genjet and genparticle results

The cross section at parton level is slightly higher.

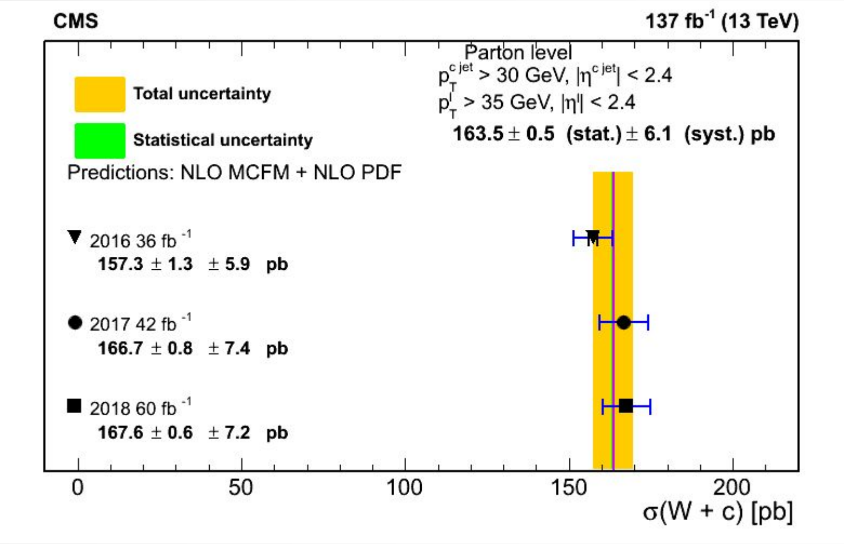
Consistent with the amount of genjets that do not fall into the acceptance region when the matching genparticle does.

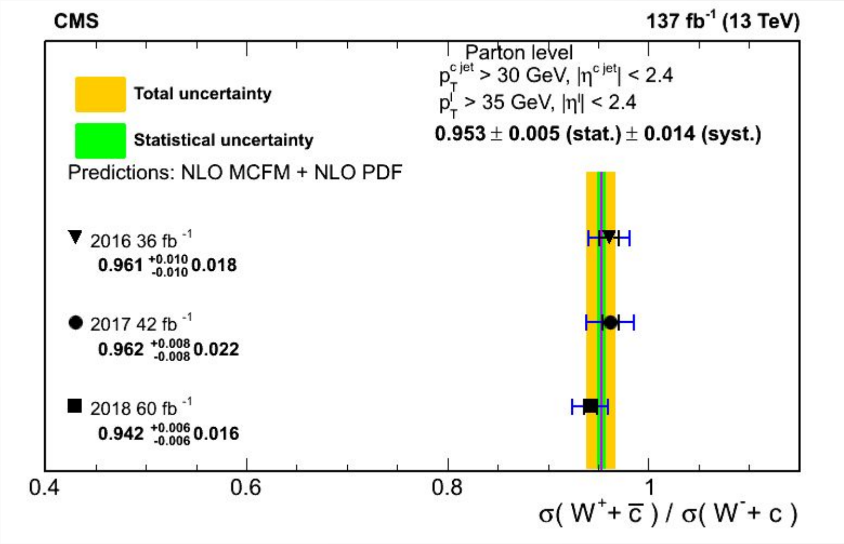


Transverse momentum of the genjet when the matched charm quark pt is > 30 GeV.

Cross section integrated over fiducial region

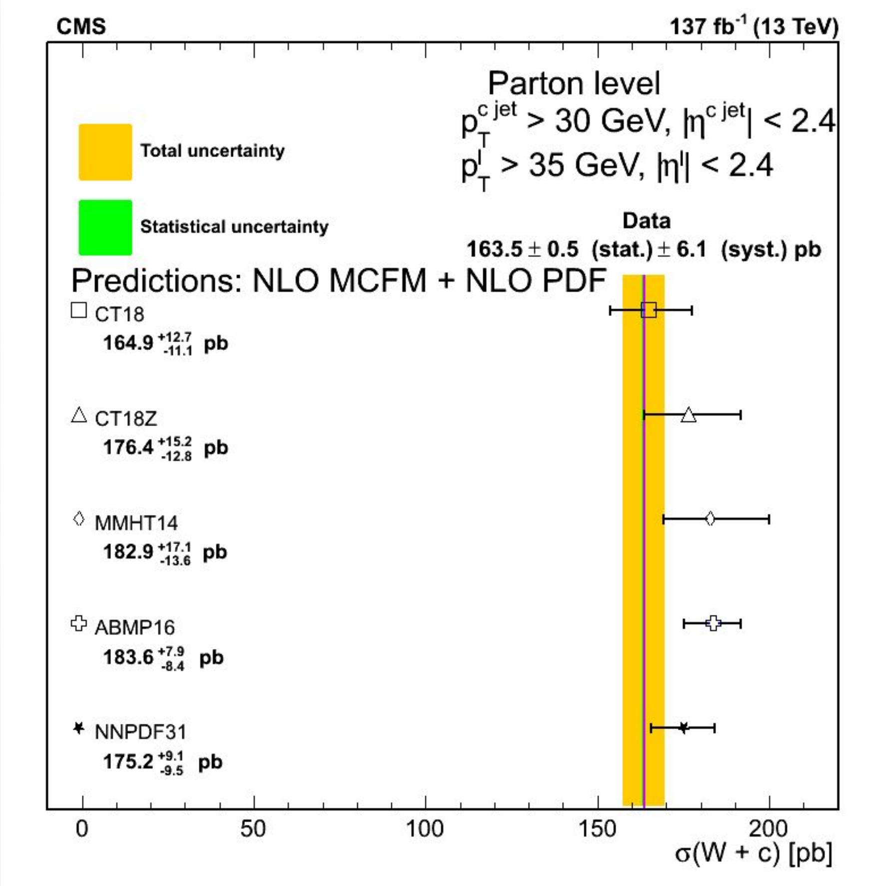
Comparison with measurements with 2016, 2017 and 2018 data

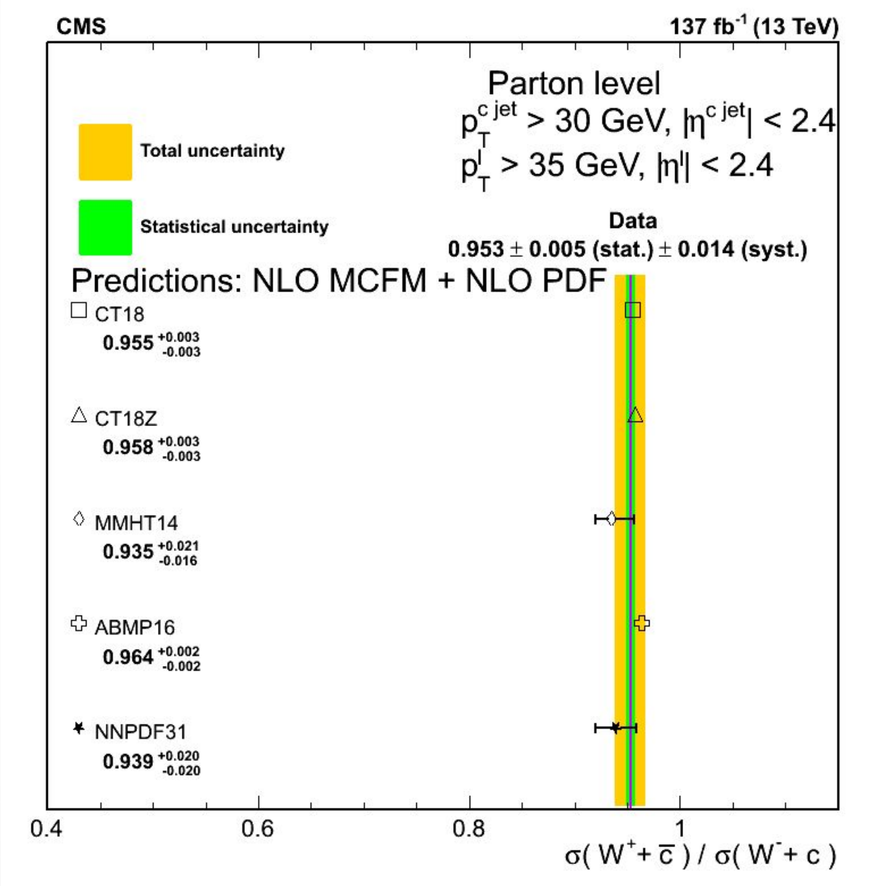




Cross section integrated over fiducial region

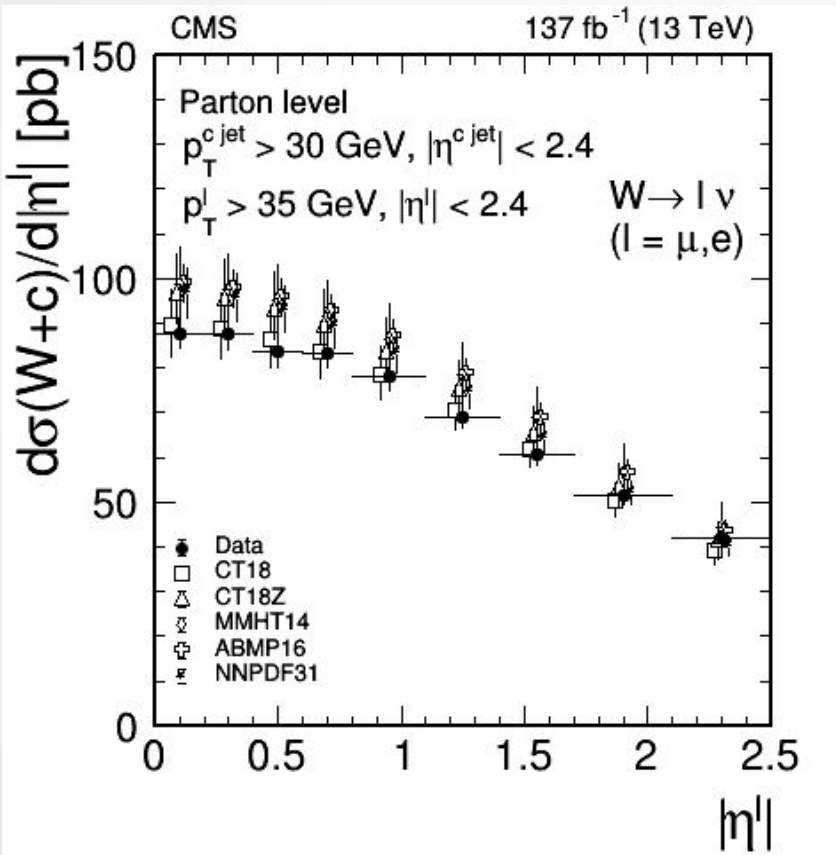
Comparison with MCFM calculations using different strange PDFs.

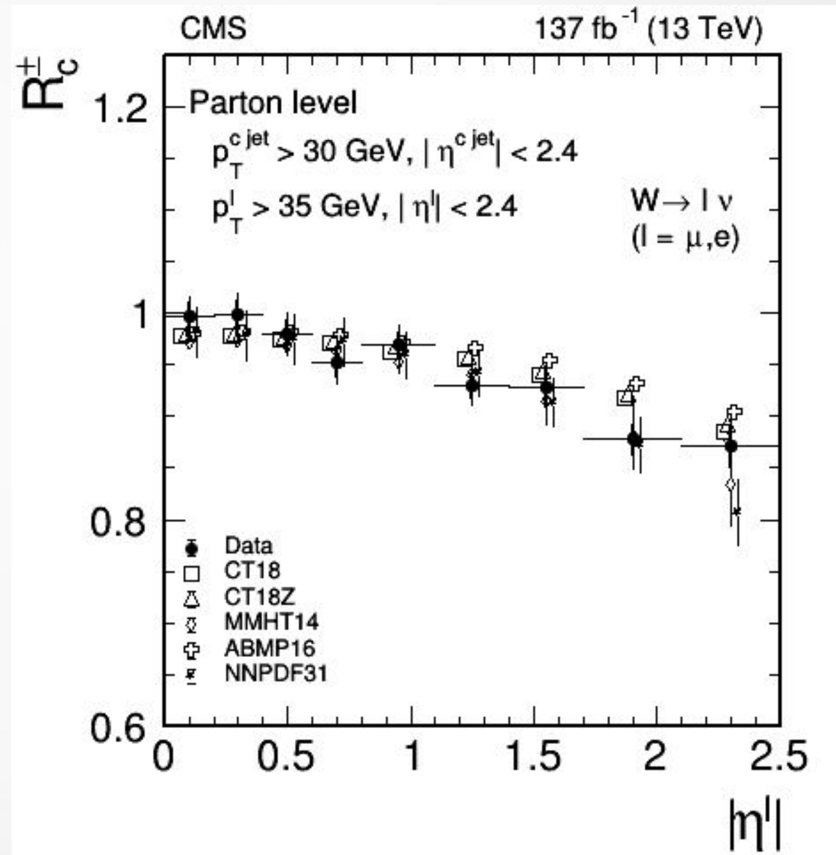




Combination of differential cross section vs eta W-lepton RunII

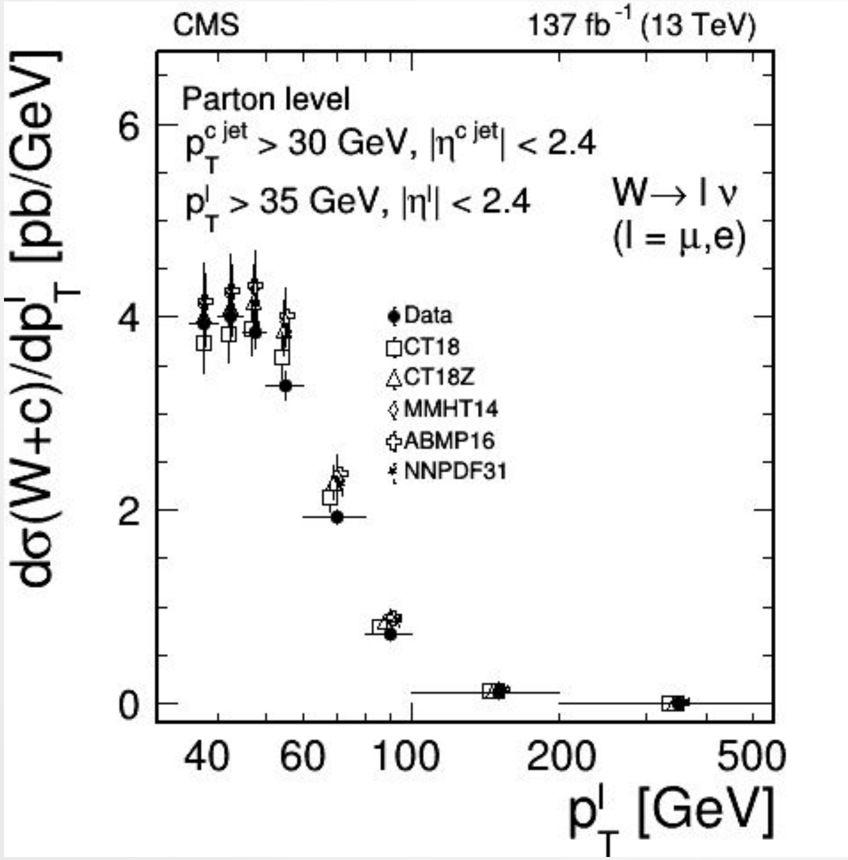
Comparison with MCFM calculations using different strange PDFs.

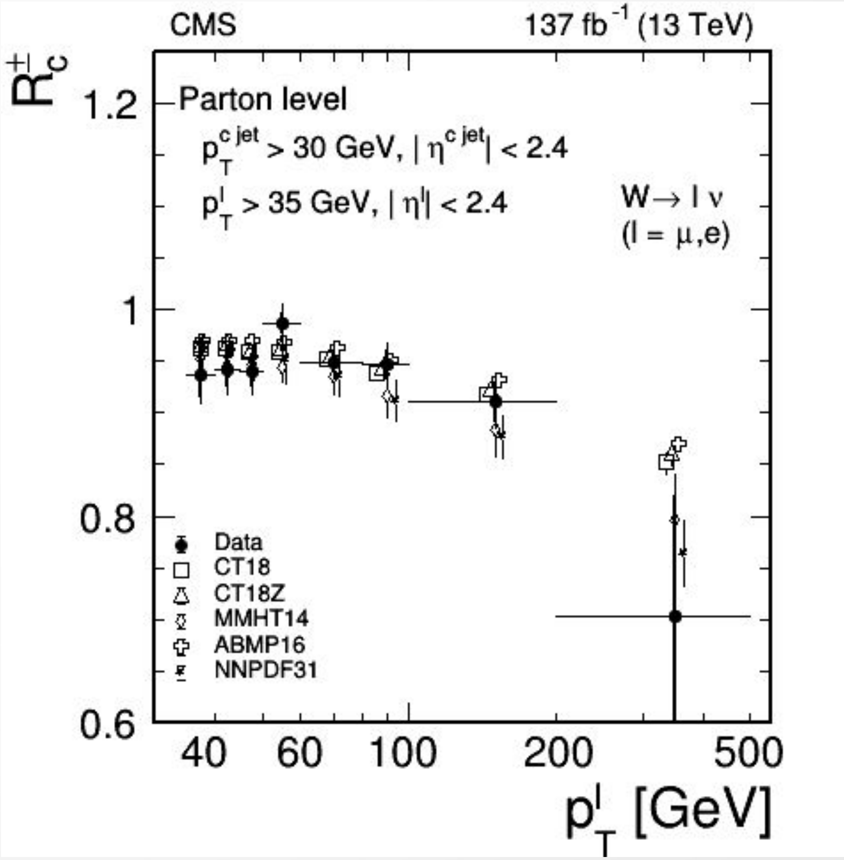




Combination of differential cross section vs pt W-lepton RunII

Comparison with MCFM calculations using different strange PDFs.





Conclusions

- Measurements of W+c and W+/W- integrated and differential (vs eta and pt of the isolated lepton) over a fiducial region for full Run II.
- Particle level. Comparison with the predictions from the madgraph MC generator at LO and NLO accuracies.
- Parton level. Comparison with the MCFM models using different PDF sets.
- This measurement can be used to improve the precision in the determination of the strange PDF.

- BACK UP

Simulation corrections

- Pileup Reweighting.
- Muon ID, Iso, Trigger Efficiency
- Electron ID, Reco and Trigger Efficiency
- Soft muon in SL
- Rochester Corrections for muon pt.
- JES Corrections.
- L1 PrefiringWeights 2016 and 2017.
- c-tagging efficiencies.
- Branching ratios and fragmentation fractions.

Cross section σ(W+c) (pb) Run II for each decay channel

SV channel

Channel	N_{rec}	σ_{W+c} (pb)
$W \to e\nu$	283871 ± 1717	$144.7 \pm 0.9 \pm 7.3$
$W^+ \to e\nu$	139476 ± 1227	$70.0 \pm 0.6 \pm 3.9$
$W^- \to e\nu$	144221 ± 1204	$74.6 \pm 0.6 \pm 3.7$
$W^+/W^-(\to e\nu)$		$0.939 \pm 0.011 \pm 0.024$
$W \to \mu\nu$	412175 ± 1876	$149.8 \pm 0.7 \pm 6.4$
$W^+ \to \mu\nu$	204586 ± 1342	$74.6 \pm 0.5 \pm 3.2$
$W^- \to \mu\nu$	206878 ± 1314	$74.9 \pm 0.5 \pm 3.4$
$W^+/W^-(\to \mu\nu)$		$0.996 \pm 0.009 \pm 0.019$

SL channel

Channel	N_{rec}	σ_{W+c} (pb)
$W \to e\nu$	347242 ± 1294	$162.0 \pm 0.6 \pm 7.3$
$W^+ \to e\nu$	168362 ± 922	$78.9 \pm 0.4 \pm 3.7$
$W^- \to e\nu$	178822 ± 907	$83.1 \pm 0.4 \pm 3.8$
$W^+/W^-(\to e\nu)$		$0.949 \pm 0.007 \pm 0.020$
$W \to \mu\nu$	374468 ± 1277	$155.8 \pm 0.5 \pm 7.2$
$W^+ \to \mu\nu$	180890 ± 912	$75.0 \pm 0.4 \pm 3.6$
$W^- \to \mu\nu$	193623 ± 893	$80.9 \pm 0.4 \pm 3.8$
$W^+/W^-(\to \mu\nu)$		$0.927 \pm 0.006 \pm 0.017$

Table 10: Measured production cross sections σ_{W+c} in the four channels for full Run II. Statistical (first error) and systematic (second error) uncertainties are also given.

JES systematics in the SL channel (vs SV channel)

First the selection of the jet identified as charm is different for the SL and SV . The selection is such that the shape of the pt of the jet is different for the SL and SV channels. The pt of the jet is in SL channel vs SV channel. The shape is harder in the case of the SV and JEC is more important at low pT.

They key of the different pt spectrum is SL : the jet identified as charm is the one with highest pt non isolated muon (regardless of this jet being the highest pt jet in the event or not). For the SV: the jet identified as charm is the highest pt-jet with a SV (priority to the pt of the jet NOT the highest pt-SV on any jet). [The reason why doing it differently is to reduce the ttbar background a bit more in this SV channel]

May be also the fact that we miss the neutrinos in the SL channel also helps/explains the differences in shape/systematics

Systematic uncertainties. Charm fragmentation fractions and charm hadron decay branching ratios

In the MC simulation we have reweighted the values used by PYTHIA8 for the charm fragmentation fractions into D hadrons (D0,D+,Ds,Λc) and their decay branching ratios to match more recent measured values (PDG).
To consider the systematic uncertainty, we have varied these quantities around their uncertainties following the procedure:

1. Change all the branching ratios and fragmentation fractions following a gaussian distribution with mean the reweighted values of the analysis and as sigma the error of such reweighted value.
2. For each change, repeat the analysis.
3. Repeat N-times.
4. Do a fit of the resultant distribution to a gaussian.
5. Take the Gauss-sigma/Gauss-mean as the systematic uncertainty.

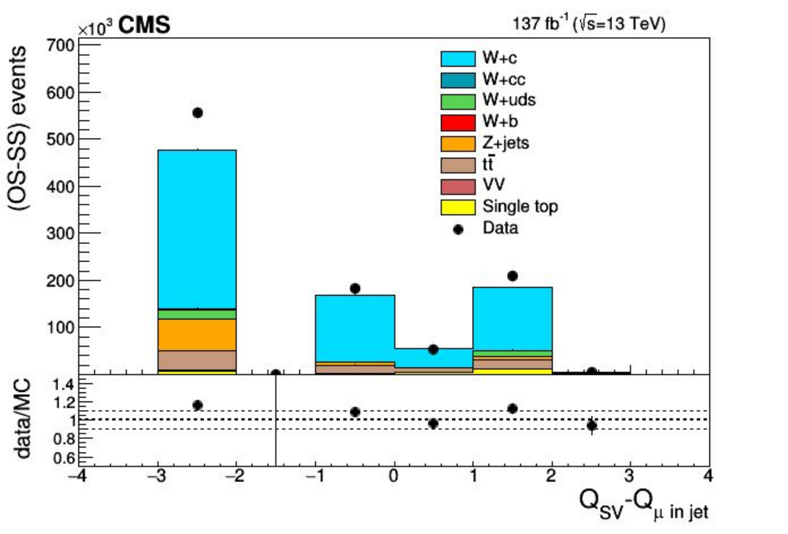
This study was performed in three different ways: changing branching ratios and fragmentation fractions at the same time, changing only branching ratios and changing only fragmentation fractions. Changing only the branching ratios has more impact than changing only fragmentation fractions. The change of both at the same time is consistent with the sum in quadrature of both effects alone. The uncertainty in the measured cross section due to branching ratios and fragmentation fractions of charm hadrons is then ≈ 2% (SL channel) and ≈ 1% (SV channel).

Systematic uncertainties (III). Correlations for the combination

Each systematic uncertainty is considered to be uncorrelated with the others but we consider correlations among the four channels (either 0 or 100) so that systematic uncertainties arising from a common source and affecting several measurements are considered to be fully (100) correlated among them.

- Electron efficiencies are considered correlated between the two $W \rightarrow e\nu$ channels and independent of the $W \rightarrow \mu\nu$ channels (viceversa for muons efficiencies).
- PU, JES, MET and luminosity are considered correlated among the four channels.
- Secondary vertex reconstruction efficiency and secondary vertex charge assignment are considered to be correlated in the two SV channels.
- Monte Carlo statistics is considered to be uncorrelated among the four channels.
- Semileptonic fragmentation fraction and branching ratio is taken as correlated in the two SL channels.
- SV fragmentation fraction and branching ratio is taken as correlated in the two SV channels.

Test of Q_{SV} algorithm



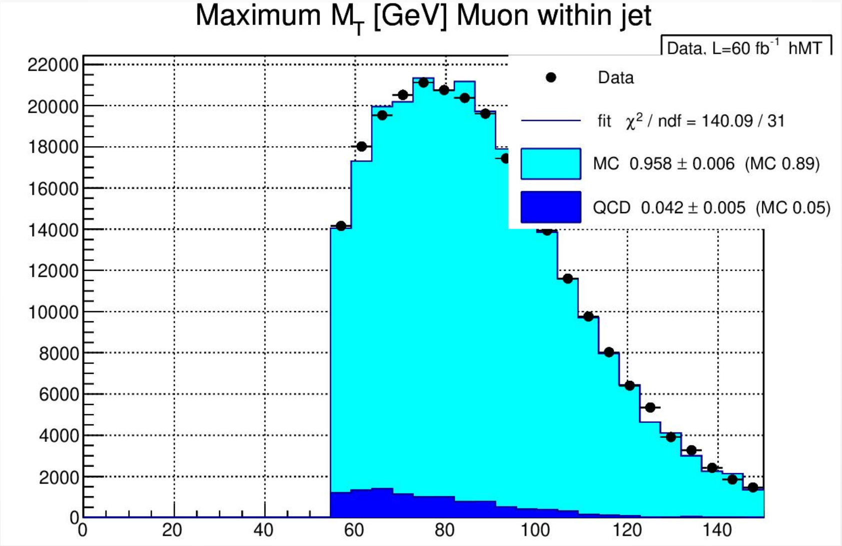
SL sample. The electric charge of the c quark is given by the sign of the track reconstructed as a muon inside the c-jet. The value of the charge of the muon is tested against the value obtained using our algorithm. Difference in performance between data and MC used as a source of systematics in the determination of the charge of the SV

- $Q_{\mu}-Q_{vertex} = \pm 1$:
  - D_{hadron} candidates made out of two tracks.
  - D_{hadron} candidates with flight distance significance of the secondary vertex<4.
  - Even when we have the charge of those vertices we are setting Q_{μ} to 0 in this plot so that we only see in the bin 0 the ones of our SV channel. These D_{hadron} candidates are not being used in the SV channel.
- $Q_{\mu}-Q_{vertex} = -3$: D_{hadron} candidates containing a muon in the jet but no secondary vertex is being reconstructed. Approximately half of our SL events contain a secondary vertex.
- $Q_{\mu}-Q_{vertex} = \pm 2$:D_{hadron} candidates with Q_{μ} and Q_{vertex} being different.
- $Q_{\mu}-Q_{vertex} = 0$: D_{hadron} candidates with Q_{μ} and Q_{vertex} being the same.

When we have a vertex with >= 3 tracks in > 90% of the cases both values agree, $Q_{\mu} = Q_{vertex}$.

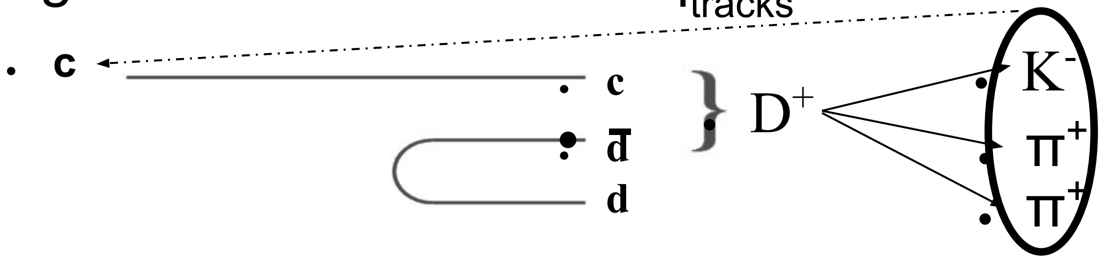
QCD background contamination

- We check a control region requiring that the isolated lepton be in this case non-isolated.
- Then look at transverse mass of the lepton-neutrino system and normalize data over MC to estimate the missing background.
- For the electron decay of the W the contribution is compatible with 0. For the muon decay it is very small and with big uncertainties.
- We perform a fit and add it to the cross section calculations.

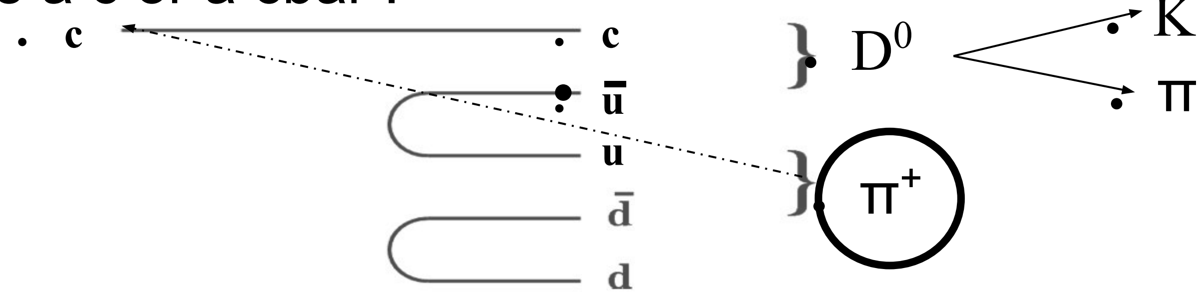


Charm charge determination

- Charge of the SV vertex : $\sum_{\text{Vertex-tracks}} q_{\text{tracks}}$



- If vertex-charge == 0 use charge of closest PV-track (p_T>0.3 && ΔR(track, IVF-vertex)<0.1). The sign of that track, the closest track to the $\text{Charm}_{\text{hadron}}$ in the process of fragmentation, tells you whether we have a c or a cbar :



- Charge definition: OS if charges of the SV and the lepton from the W decay are opposite.
  - OS : $Q_{W\rightarrow e(\mu)} \neq Q_{SV}$ SS : $Q_{W\rightarrow e(\mu)} = Q_{SV}$

- If no charge identified → event is removed