Measurement of Top Quark Mass and Production Cross-Section using tt-bar pairs with Semi-Leptonic Decay Channel

Qamar-ul-Hassan
11-FBAS/PHDPHY/S11

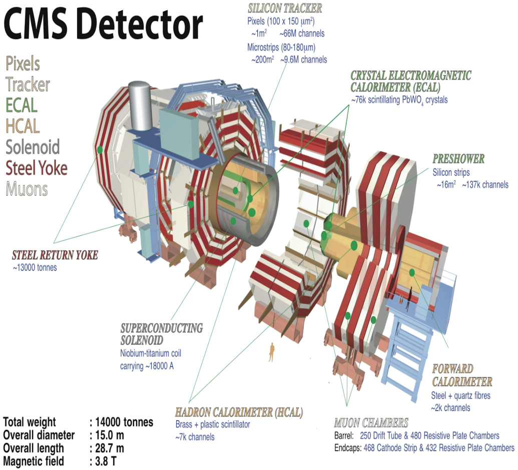
Supervisor: **Prof. Dr. Hafeez R. Hoorani** (NCP, Islamabad)
Co-Supervisor: **Dr. Waqar Adil Syed** (IIU, Islamabad)

October 25, 2017

Large Hadron Collider (LHC)

- the LHC is a discovery machine
- the LHC will determine the future course of high energy physics
- to smash protons moving at 99.999999% to the speed of light
- to create conditions a fraction of a second after the big bang
- High energies allows us:
  - to look deeper into nature (E α 1/size), ("powerful microscopes")
  - to discover new particles with high(er) mass (E=mc^2)
  - to study the early universe (E=kT)
  - revisit the earlier, hotter, history of our universe, searching for a new simplicity ("powerful telescopes") by observing phenomena and particles no longer observable in our everyday experience

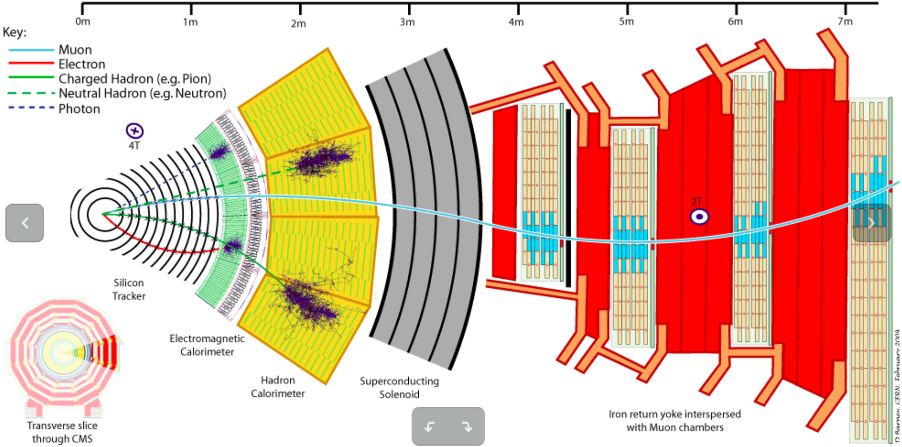
Compact Muon Solenoid (CMS)



- Distinct features are compact size and superconducting solenoid (3.8 T)
- Tracker and calorimeters are located inside the solenoid
- Muon detectors are out side the solenoid
- All these sub-detectors are arranged in layers around the interaction point

- Multi-purpose detector

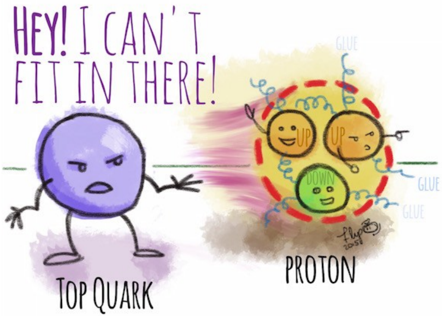
Compact Muon Solenoid (CMS)



- Particle identification: Various types of particles interact with matter differently, leaving a typical or no signal in specific subdetectors.

TOP quark physics

- The heaviest quark in the SM, discovered in 1995 at Tevatron.
- Charge of $+2/3$, weak isospin of $+1/2$ and life-time is 5×10^{-25} sec, so decays before hadronization.
- The spin information is kept by its decay products and its only quark which provides opportunity to study the properties of a bare quark.
- Major backgrounds to many important searches beyond the Standard Model (BSM) processes.

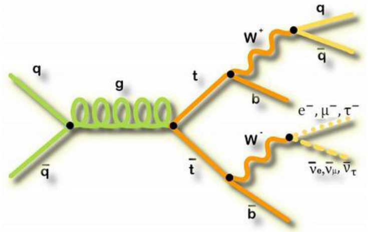


Scope

- We aim to measure the cross section at 13 TeV
  - how precise can we measure it in the l+jets final state
  - precise measurement of the cross section opens the possibility to perform
    - precision measurements (PDFs, α_S)
    - extrapolating the pole mass http://arxiv.org/abs/1511.00841
  - search for physics beyond the standard model is particularly relevant at Run 2
    - A deviation in the ratio of cross section at 13 TeV/ 8TeV http://arxiv.org/abs/1206.3557
    - by setting limit on stop production http://arxiv.org/abs/1407.1043
    - A deviation in the branching ratios of different final states (test lepton universality in ttbar decays)
    - For this preliminary result, we shall measure σ only

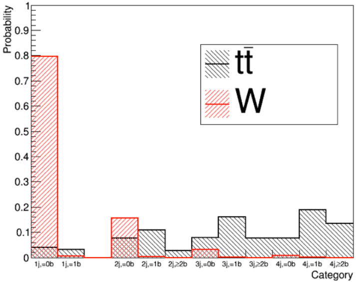
L+Jets Final State

- The l+jets final state is chosen
- High statistics, moderate background
- Branching ratio ($t\bar{t} \to l\nu q\bar{q}b\bar{b}$) = 30% + 5.3% from $tau \to l\nu\nu$ decays
- Expect 1 lepton, $\not{E}_T$ and 4 jets including at least 2 b-tags
- We start the analysis with 1 jet (next slide)



Analysis Strategy

- We divide the analysis in different categories by counting the jets and b-tags
- Low jet/btag categories to constrain backgrounds while high jet/btag to fit the signal
- Analysis is also divided according to lepton charge
  - charge asymmetry is expected for W+Jets
- We don't subdivide the simulation per heavy flavor
  - As we are not having the enough statistics
- similar approach of TOP-11-004/TOP-13-004 ($SH_Y FT$, legacy Run-I)



Physics object selection

Muon Selection

- IsoMu20||IsoTkMu20
- Exactly one muon
  - p_T >30 GeV, $|\eta|$ <2.1, tight ID
  - dB <0.2, dZ <0.5, relIso <0.15
  - SIP3D <4
- Loose Muon Veto
  - p_T >15 GeV, $|\eta|$ <2.4, relIso <0.25

Electron Selection

- Ele22_eta2p1_WPLoose_Gsf
- Exactly one electron
  - Using official electron tight ID
  - p_T >30 GeV, $|\eta|$ <2.1
  - SIP3D <4
- Using recommended electron veto ID

Jet Selection

- We require at least one jet
  - LOOSE jet ID is applied
  - p_T >30 GeV, $|\eta|$ <2.4, Fall15_25nsV2
- Count number of jets identified as b's
  - CSV Discriminator value >0.800 (medium working point)

Data and MC samples

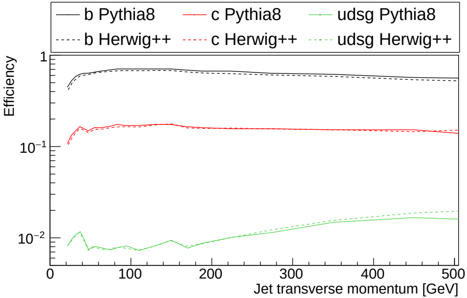
Data samples	Integrated Luminosity
/SingleMuon/Run2015C_25ns-16Dec2015-v1/MINIAOD /SingleElectron/Run2015C_25ns-16Dec2015-v1/MINIAOD /SingleMuon/Run2015D-16Dec2015-v1/MINIAOD /SingleElectron/Run2015D-16Dec2015-v1/MINIAOD	2.2 fb^{-1}

MC samples	Process	Cross section[pb]	Events
/TT_TuneCUETP8M1_13TeV-powheg-pythia8	$t\bar{t}$	831.76	97994442
/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	$t\bar{t}$+V	0.2043	250307
/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	$t\bar{t}$+V	0.4062	831847
/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	$t\bar{t}$+V	0.5297	747000
/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	$t\bar{t}$+V	0.2529	394200
/ZZ_TuneCUETP8M1_13TeV-pythia8	Multiboson	16.523	985600
/WWToLNuQQ_13TeV-powheg	Multiboson	49.497	6996000
/WWTo2L2Nu_13TeV-powheg	Multiboson	12.178	1979988
/WZ_TuneCUETP8M1_13TeV-pythia8	Multiboson	47.13	100000
/WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	W	61526.7	199037280
/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	tW	35.6	1000000
/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	tW	35.6	999400
/ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	t-ch	44.33	3299200
/ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	t-ch	26.38	1630900
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	DY	6025.0	247512446
/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	DY	18610	76558711
/Multijets (data)			

- MC samples are normalized by: $\hat{N} = \mathcal{L} \cdot \sigma \cdot \frac{\sum_{i=1}^{\mathrm{N_{sel}}} w_i}{\sum_{i=1}^{\mathrm{N_{gen}}} w_i}$

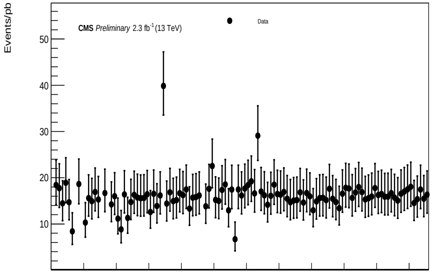
Correction Factors applied to Simulation

- PU reweighting based on 69 minimum bias cross section
- b-Tag scale factors (76X)
  - we scale Herwig++ efficiency to match pythia8 and apply the BTV scale factors to that
- Lepton SF (Official POG), cross checked in this analysis
- JEC (Fall15_25nsV2)
- JES sources
  - uncertainties broken in 29 sources
- JER
  - JER twiki

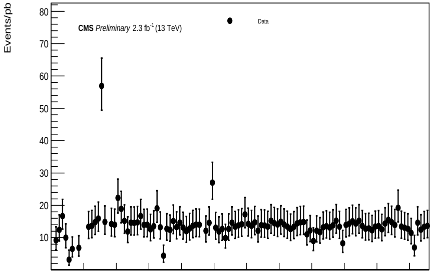


Stability of the Selection

- The rate of events is very stable as function of run number
- A couple of runs identified with anomalous rates have very low luminosity ($\sim 1pb^{-1}$)

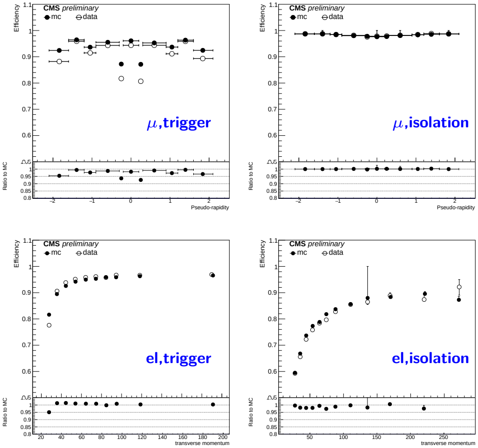


4J,μ^+



4J, el^+

Trigger Efficiency Measurements



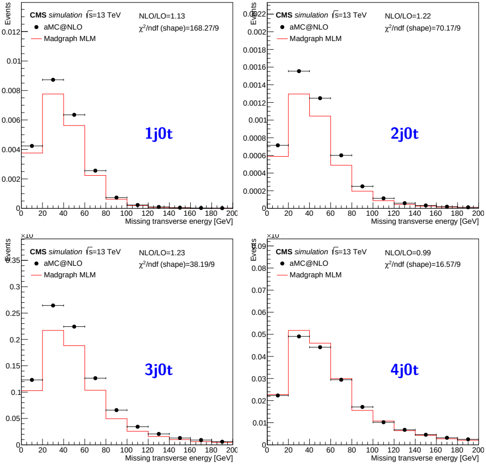
- Using official TnP trees adopted for our selection (Different fitting models tried)
- Results are used as a cross check of the official SFs

Background estimation

Background Estimation: W+Jets

- Modelled from simulation
- Analysis in each lepton channel is sub-divided according to lepton charge
- $t\bar{t}$ production is expected to be symmetric
- W production is expected to asymmetric due to PDF composition of protons
- Assign the corresponding uncertainties from the choice of the QCD scales and PDFs.
  - The templates are currently derived from the MG5_aMC@NLO
  - Large fraction of negative weights, sufficient statistics to provide positive defined and fairly smooth templates
- Compared MG5_aMC@NLO with MADGRAPH in different categories
- Difference in normalization is observed but the shape of the distributions is fairly similar

MET Distributions



W+Jets Normalization and Charge Asymmetry

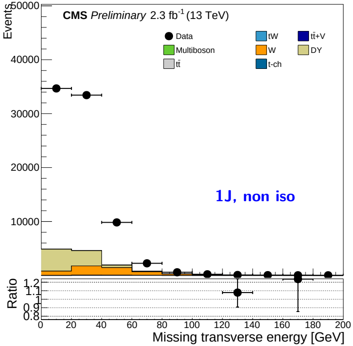
Category	=1 jet	=2 jets	=3 jets	≥4 jets
μ^+	1.112±0.002	1.184±0.006	1.20±0.01	0.96±0.02
μ^-	1.095±0.003	1.170±0.006	1.19±0.01	0.97±0.02
e^+	1.102±0.003	1.196±0.007	1.21±0.02	0.96±0.02
e^-	1.288±0.004	1.38±0.008	1.40±0.02	1.09±0.03

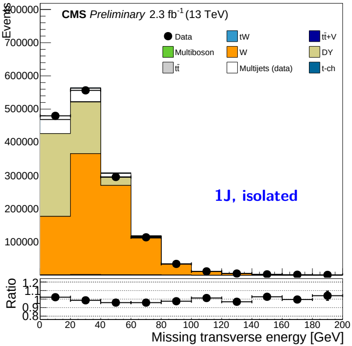
- Ratio of the events predicted by MG5_aMC@NLO wrt MADGRAPH
- Predicts $\sim$ 20% more events wrt to madgraph for Njets $<$4
- Predicts $\sim$ 5% less events for Njets $>$4
- As a consequence of this observation, different charge asymmetry is predicted

Background Estimation: QCD Multijets

- Normalization is determined from events with low $\not{E}_T$ (<20 GeV)
  - Expect to be dominated by QCD
  - Extract QCD shape from events with relIso >0.4
  - Normalized from:

$$N_{\rm SR}(QCD) = [N_{\rm CR}(obs) - N_{\rm CR}(non-QCD)] \cdot \frac{N_{\rm SR}^{E_T^{miss}<20}(obs) - N_{\rm SR}^{E_T^{miss}<20}(non-QCD)}{N_{\rm CR}^{E_T^{miss}<20}(obs) - N_{\rm CR}^{E_T^{miss}<20}(non-QCD)}$$





Background Estimation: QCD Multijets

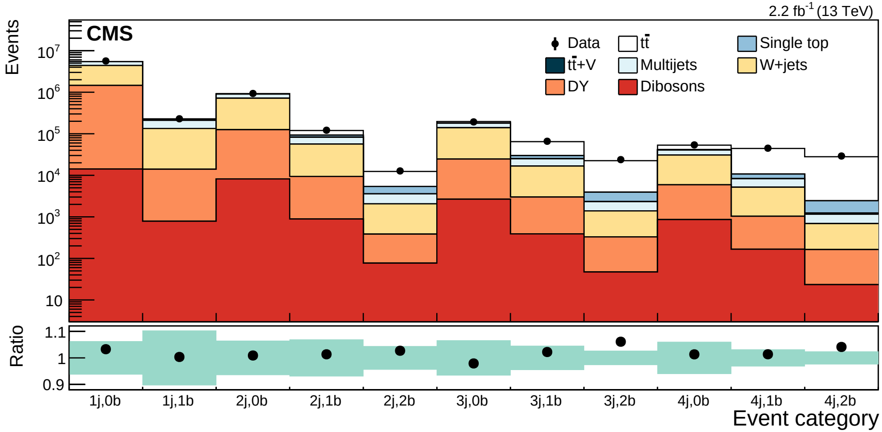
Category	=1 jet	=2 jets	=3 jets	≥4 jets
μ^+	1.89 ± 0.09	1.63 ± 0.14	2.77 ± 0.24	2.17 ± 0.23
μ^-	1.59 ± 0.05	1.19 ± 0.05	2.64 ± 0.22	2.01 ± 0.15
e^+	2.71 ± 0.09	2.29 ± 0.10	2.60 ± 0.13	2.14 ± 0.30
e^-	2.50 ± 0.12	2.06 ± 0.05	2.56 ± 0.13	2.03 ± 0.10

- Scale factors to be applied to the control region distributions for QCD to normalize them in the signal region
- Overall the scale factors are approximately constant across the different jet multiplicity categories
- tend to increase for higher jet multiplicities
  - Due to NON QCD contamination in control region
- Uncertainty is higher in the positively charged lepton sample (as a consequence of charge asymmetry)
- All scale factors are close to two, data is triggered by isolated lepton triggers
- We capture only a portion of non-isolated leptons

Control distributions

(pre-fit, stats only uncertainty)

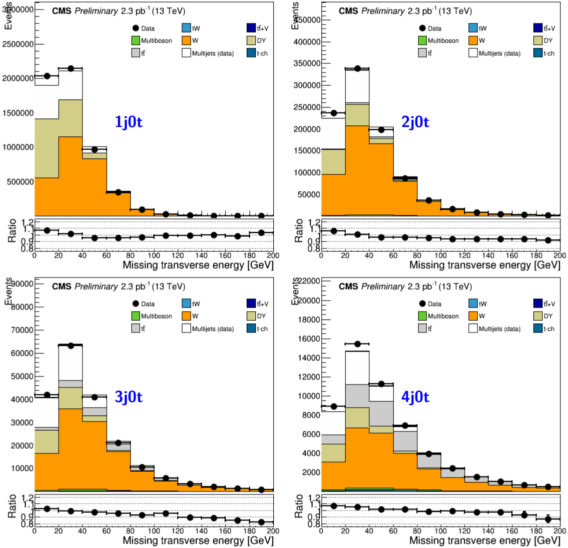
Events count



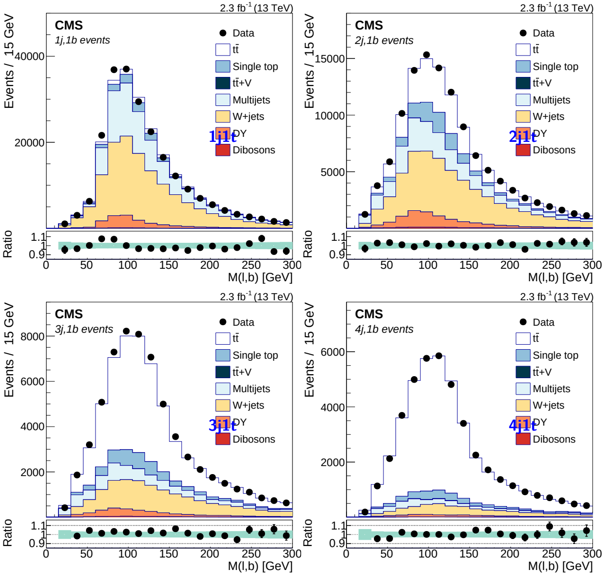
- e/mu+jets combined for positive/negative charged leptons
- Statistical uncertainty only

MET distributions

- At pre-approval (74X), we were using these distributions to constrain backgrounds
- Not so well modelled in 76X (electron channel)
    - we count simply the number of events with 0-btags

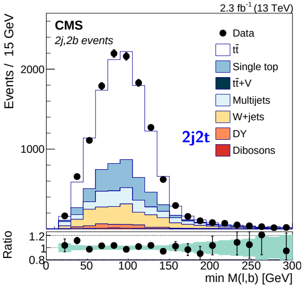


M(l,b)

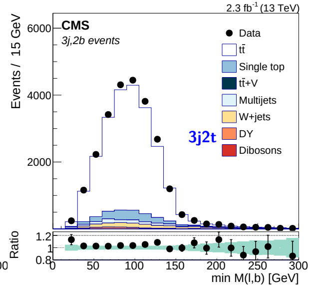


2.3 fb^{-1} (13 TeV)
CMS
1j,1b events
2j,1b events
3j,1b events
4j,1b events
Data
$t\bar{t}$
Single top
$t\bar{t}$+V
Multijets
W+jets
DY
Dibosons
1j1t
2j1t
3j1t
4j1t
Events / 15 GeV
Ratio
M(l,b) [GeV]

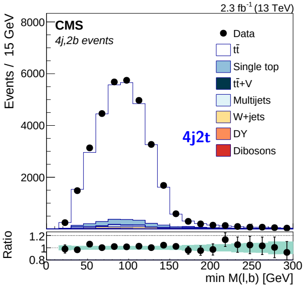
minM(l,b)



2.3 fb^{-1} (13 TeV)
CMS
2j,2b events
Data
$t\bar{t}$
Single top
$t\bar{t}$+V
Multijets
W+jets
DY
Dibosons
2j2t
Events / 15 GeV
Ratio
min M(l,b) [GeV]



2.3 fb^{-1} (13 TeV)
CMS
3j,2b events
Data
$t\bar{t}$
Single top
$t\bar{t}$+V
Multijets
W+jets
DY
Dibosons
3j2t
Events / 15 GeV
Ratio
min M(l,b) [GeV]



2.3 fb^{-1} (13 TeV)
CMS
4j,2b events
Data
$t\bar{t}$
Single top
$t\bar{t}$+V
Multijets
W+jets
DY
Dibosons
4j2t
Events / 15 GeV
Ratio
min M(l,b) [GeV]

Fit results and cross section measurement

Cross Section Measurement

- The $t\bar{t}$ production cross section $\sigma_{t\bar{t}}$ is extracted from the following expression

$$\sigma = \frac{\boxed{N_{obs} - N_{bkg}}}{A\varepsilon\mathcal{L}}$$

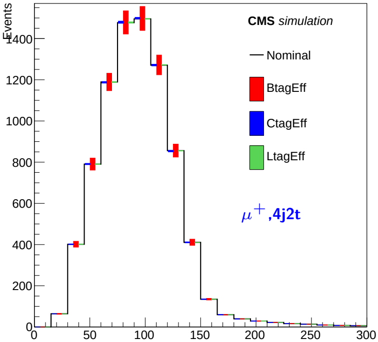
- We make use of Higgs Combination tool
- Use all categories to fit the signal and adjust initial background estimation
- In the fit, uncertainties are included

$$N^{bkg} \rightarrow N^{bkg}_{nom} \cdot (1 + \theta_{btag}) \cdot (1 + \theta_{jes}) \cdot .....$$

- Use the profile likelihood method (PLR) to fit the signal strength

$$\mu = \sigma_{obs}/\sigma_{th}$$

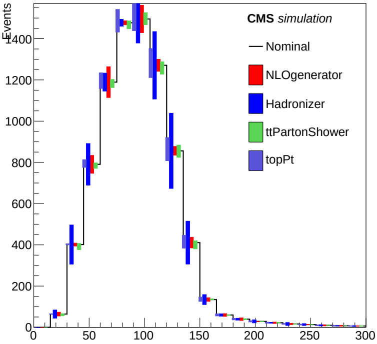
Experimental uncertainties



- b-tagging efficiency is expected to have the largest impact on the signal shapes.
- nuisances are constrained by log-normal distributions

Uncertainty group	Name	Type	Prior	Range
JES	AbsoluteFlavMap			1.0-1.14
	AbsoluteMPFBias			1.0-1.18
	AbsoluteScale			1.0-1.3
	AbsoluteStat			1.0-1.2
	FlavorPureBottom			1.0-1.084
	FlavorPureCharm			1.0-1.144
	FlavorPureGluon			1.0-1.156
	FlavorPureQuark			1.0-1.073
	Fragmentation			1.0-1.106
	PileUpDataMC			1.0-1.12
	PileUpPtBB			1.0-1.17
	PileUpPtEC1			1.0-1.15
	PileUpPtEC2			1.0-1.13
	PileUpPtHF			1.0-1.12
	PileUpPtRef	shape+rate	lnN	1.0-1.14
	RelativeFSR			1.0-1.2
	RelativeJEREC1			1.0-1.10
	RelativeJEREC2			1.0-1.14
	RelativeJERHF			1.0-1.17
	RelativePtBB			1.0-1.19
	RelativePtEC1			1.0-1.11
	RelativePtEC2			1.0-1.12
	RelativePtHF			1.0-1.13
	RelativeStatEC			1.0-1.07
	RelativeStatHF			1.0-1.12
	SinglePionECAL			1.0-1.1
	SinglePionHCAL			1.0-1.12
	TimeEta			1.0-1.12
	TimePt			1.0-1.20
JER	JER	shape+rate	lnN	1.0-1.16
	BtagEff			1.0-1.17
b-tagging efficiency	CtagEff	shape+rate	lnN	1.0-1.13
	LtagEff			1.0-1.16
Luminosity	lumi_13TeV	rate	lnN	1.027
Pileup	Pileup	shape	lnN	
	EleEfficiency			0.98-1.027
	EleScale			1.0-1.016
	EleTrigger			0.97-1.032
Lepton efficiencies	MuEfficiency	shape+rate	lnN	1.0-1.003
	MuScale			1.0-1.045
	MuTrigger			1.0-1.003
QCD multijets	MultijetsNorm*	rate	lnN	1.0-1.5
	MultijetsShape*	shape	lnN	

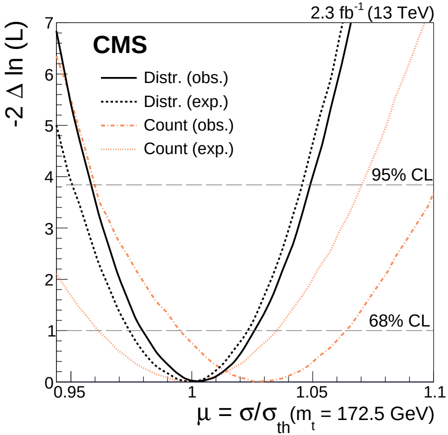
Theory uncertainties



Uncertainty group	Name	Type	Prior	Range
Theoretical cross sections	DYnorm_th	rate	lnN	1.038
	VVnorm_th			1.2
	Wnorm_th			1.037
	tWnorm_th			1.054
	tnorm_th			1.044
$t\bar{t}$ modelling	Hadronizer	shape+rate	lnN	1.0-1.344
	NLOgenerator	shape+rate	lnN	1.0-1.20
	topPt	shape+rate	lnN	1.0-1.055
	Mtop	shape	lnN	
	ttCombScale	shape	lnN	
	ttFactScale	shape	lnN	
	ttPartonShower	shape	lnN	
	ttRenScale	shape	lnN	
tW modelling	tWscale	shape+rate	lnN	1.0-1.14
	tWttinterf	shape+rate	lnN	1.00-1.30
W+jets modelling	wCombScale		lnN	1.05-1.13
	wFactScale (shape)	shape+rate		1.0-1.039
	wRenScale (shape)			1.0-1.15

- Distributions for the signal shapes
- Hadronizer is expected to have the largest impact on the signal shapes

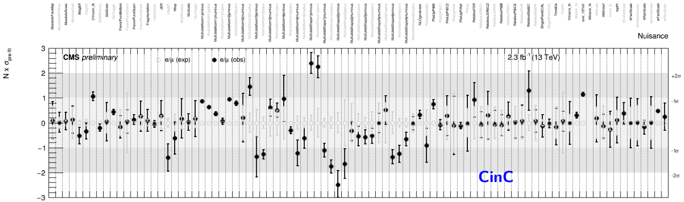
Fit result

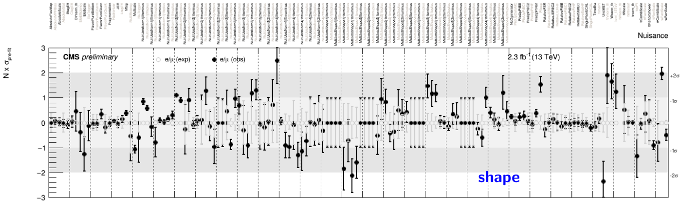


- Expected/observed for the cut-in-categories and and the shape analysis

$\mu = 1.003 \pm 0.003(stat) \pm 0.023(syst)$
$\mu = 1.030 \pm 0.004(stat) \pm 0.034(syst)$

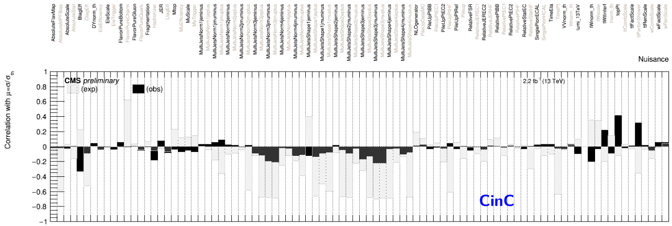
post-fit nuisances (expected & observed)

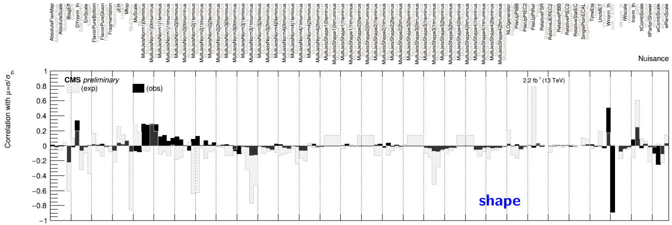




- Largest pulls are observed for the main backgrounds W and QCD multijets
- All the rest with in 1 σ
- Post fit uncertainties smaller in shapes as expected

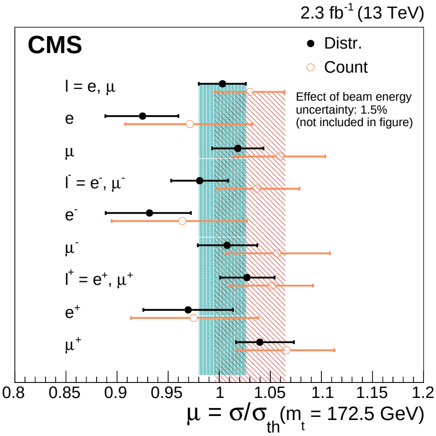
correlation factor between each nuisance and the signal strength





- Largest correlations in main backgrounds W, QCD multijets and luminosity

Fit in different categories



- The QCD contamination makes the W+jets background being pulled by the fit slightly differently in each channel.
- Difference observed is within uncertainties.
- Fair agreement between shape and cut-in-categories.

Systematics

Source	Distributions	Count
Statistical	0.003	0.004
Experimental uncertainties		
Jet energy scale/resolution	0.001	0.007
b tagging	0.005	0.011
Pileup	$<$ 0.001	$<$ 0.001
Lepton trigger/selection efficiency	0.002	0.001
Lepton energy scale	0.001	$<$ 0.001
W+jets model	0.019	0.022
multijet	0.011	0.021
Other backgrounds	0.001	0.001
Theoretical uncertainties		
$t\bar{t}$ model	0.002	0.001
Top quark p_T	0.004	0.003
Parton shower scales	0.001	0.005
$\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$	0.002	0.003
Single top quark model	0.002	0.003
Top quark mass	$<$ 0.001	0.001
Total	**0.023**	**0.034**

- The impacts of the systematic uncertainties on the likelihood are evaluated using the following prescription:
  - fix the nuisances to their post-fit values
  - repeat the fit freezing several times adding at a time a nuisance (or a group of nuisances)

Cross Section Measurement

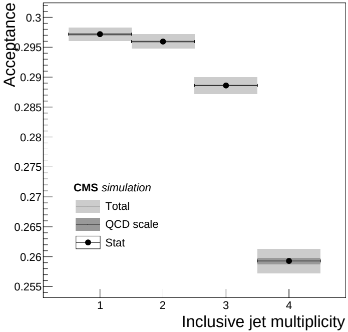
- The $t\bar{t}$ production cross section $\sigma_{t\bar{t}}$ is extracted from the following expression

$$\sigma = \frac{N_{obs} - N_{bkg}}{\boxed{A\varepsilon\mathcal{L}}}$$

- Acceptance is estimated from Powheg+Pythia8 $m_t = 172.5$ GeV
  - central value affected by the MC choice, hadronizer, parton shower scale
  - PDF uncertainties
  - QCD scale choice at ME level
- Uncertainty on the luminosity is intrinsic to the measurement 2.7 % (LUM-15-001)

Acceptance Measurement

Uncertainty	Procedure	Impact
PDF+α_s	RMS of the 100 NNPDF3.0 variations with $\Delta\alpha_s = \pm 0.001$	± 0.0006
QCD scale	Envelope of the variations (μ_R, μ_F)	± 0.0017
PS scale	Difference between nominal PS scale POWHEG+PYTHIA8 and 2,1/2 variations	
Hadronizer	Difference between PYTHIA8 and HERWIG++	± 0.0039
Integrated luminosity	2.7%	



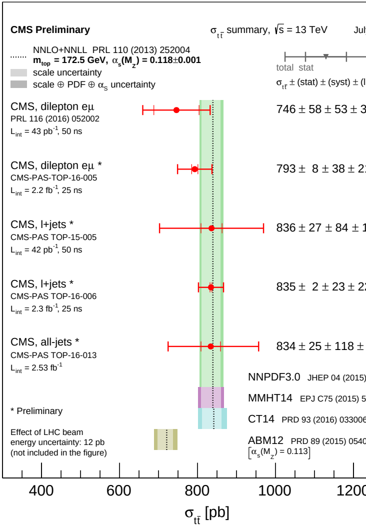
$$\boxed{Acceptance = 0.2972 \pm 0.0001(stat) \pm 0.0043(syst)}$$

- Table indicates the uncertainties considered in the extrapolation to full phase space
- Dependence on the mass enhanced by the reqquirement of extra jets

$t\bar{t}$ cross-section summary

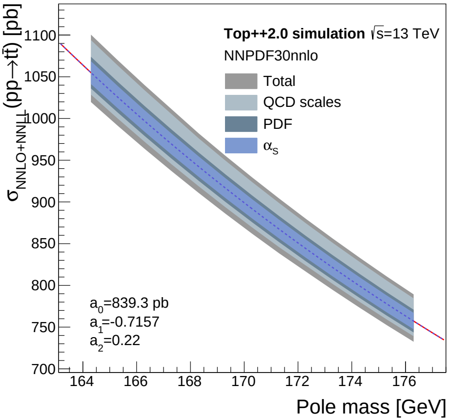
$$\boxed{\sigma \;=\; 834.7 \pm 2.5(stat) \pm 20.7(syst) \pm 22.6(lumi) \pm 12.5(extrapol)}$$

$$\boxed{\Delta\sigma/\sigma = 3.98\%}$$



Pole mass extraction

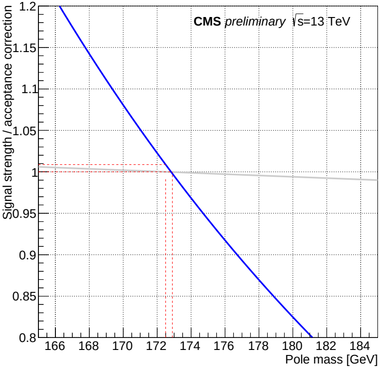
top^{++} prediction



$$\boxed{\sigma(m_t) = \sigma(m_{ref})(\frac{m_{ref}}{m_t})^4 \times [1 + a_1(\frac{m - m_{ref}}{m_{ref}}) + a_2(\frac{m - m_{ref}}{m_{ref}})^2]}$$

- We parameterize the cross section dependency on the pole mass using TOP^{++} v2.0
- Thanks J.Kieseler for cross-checking our numbers with top^{++}

Model for the Pole Mass Extraction



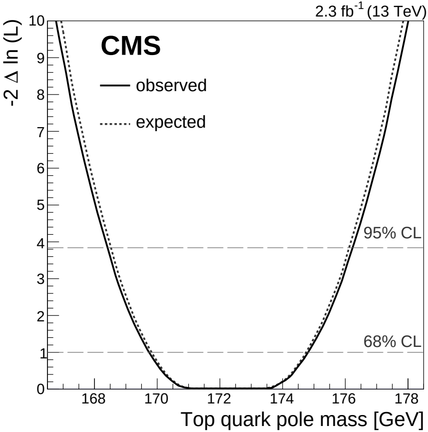
- The signal strength is re-parametrized as function of the top pole mass
  - model includes an acceptance correction estimated from MC simulation
  - extra correction to bring the LHCTOPWG reference value to NNPDF 3.0 central value
- Thanks to A. David for his help in cross checking our implementation in Higgs combine tool

Extra uncertainties in the extraction of pole mass

- In addition to the nuisances added to standard fit in the visible phase space we include:
- 1.5% extrapolation (lnN)
- 2.7% lumi (lnN)
- QCD scale (lnU)
- PDF+α_s (lnN)
- Beam energy (lnN)

Source	Δm_t [GeV]
Scales+α_S	+1.1/-0.9
Beam energy	+0.8/-0.5
Extrapolation to full phase space	+1.1/-0.7
Uncertainties from fit in the fiducial region	+2.1/-2.0
Total	+2.7/-2.3

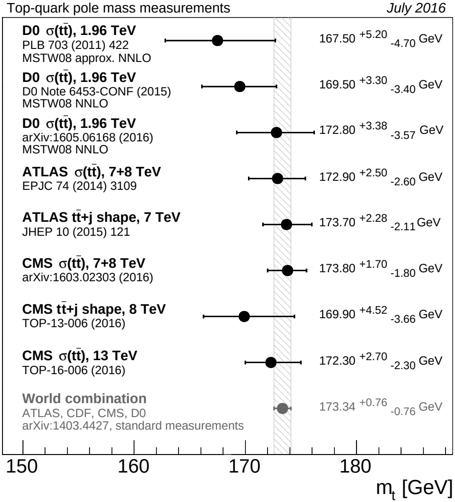
Measured top mass



- We measure 172.3 +2.7/-2.3 GeV while we expect an uncertainty +2.3/2.7 GeV notice the flatness of the likelihood around the minimum. That's induced by a flat prior on the QCD scale uncertainty. The QCD scale uncertainty is assymetric so it also induces an asymmetry in the uncertainty

$t\bar{t}$ mass summary

- Thanks Benjamin Stieger for this nice summary plot



Conclusion

- We have presented a measurement of the inclusive cross section in the l+jets channel
  - focus on optimizing the data to understand the signal in the visible region
  - uncertainty in the extrapolation reduced by requiring only one jet in the selection
- We measure:

$$\boxed{\sigma = 834.7 \pm 2.5(stat) \pm 20.7(syst) \pm 22.6(lumi) \pm 12.5(extrapol)}$$

$$\boxed{m_{\mathsf{t}} = 172.3 + 2.7/ - 2.3GeV}$$

Acknowledgments

We thank T. Arndt, M. Ivova and A. Anuar for their help with the usage of the tag and probe tool. The cross section and modelling and generators sub-conveners (M. Aldaya, M. Soares, J. Keveaney, E. Yazgan, M. Seidel) are acknowledged for their revision of the current analysis or specific parts of it. J. Kieseler is thanked for helping cross checking the values obtained with TOP++ for the top quark mass dependency of the cross section. A. David is thanked for cross checking our implementation of the model used to parametrize the signal strength as function of the top quark mass in the Higgs combination tool package.

THE END