Measurement of differential production cross section for high-p_T top quarks in proton-proton collisions at 13 TeV

(l+jets & hadronic ttbar decay channels)

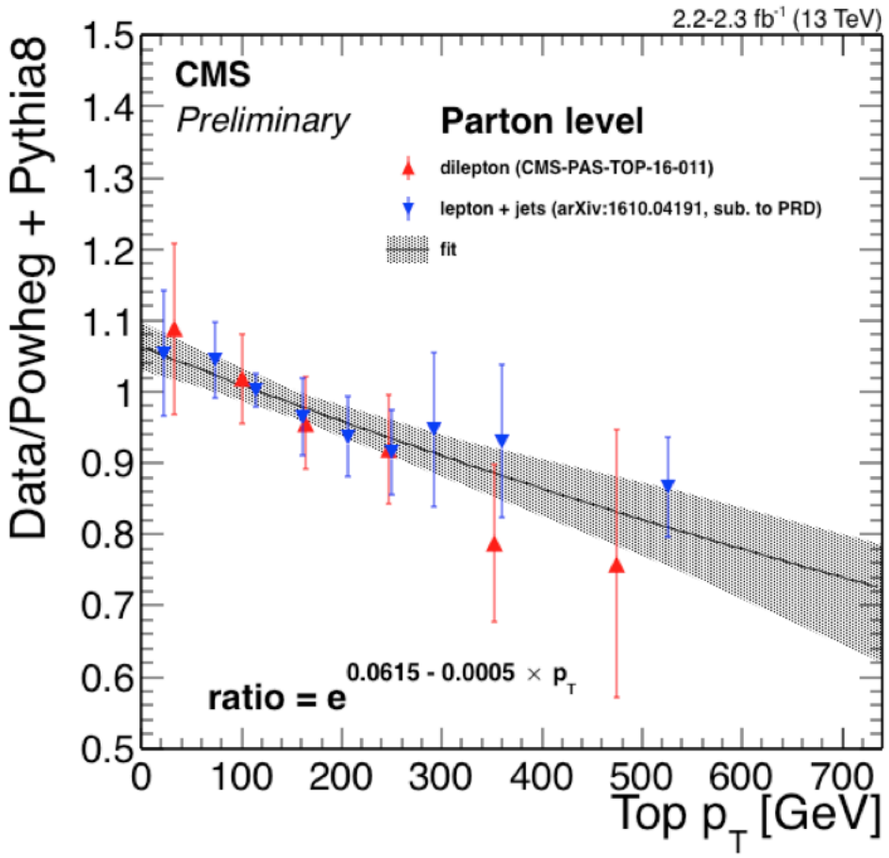
LHCtopWG

14 May 2020

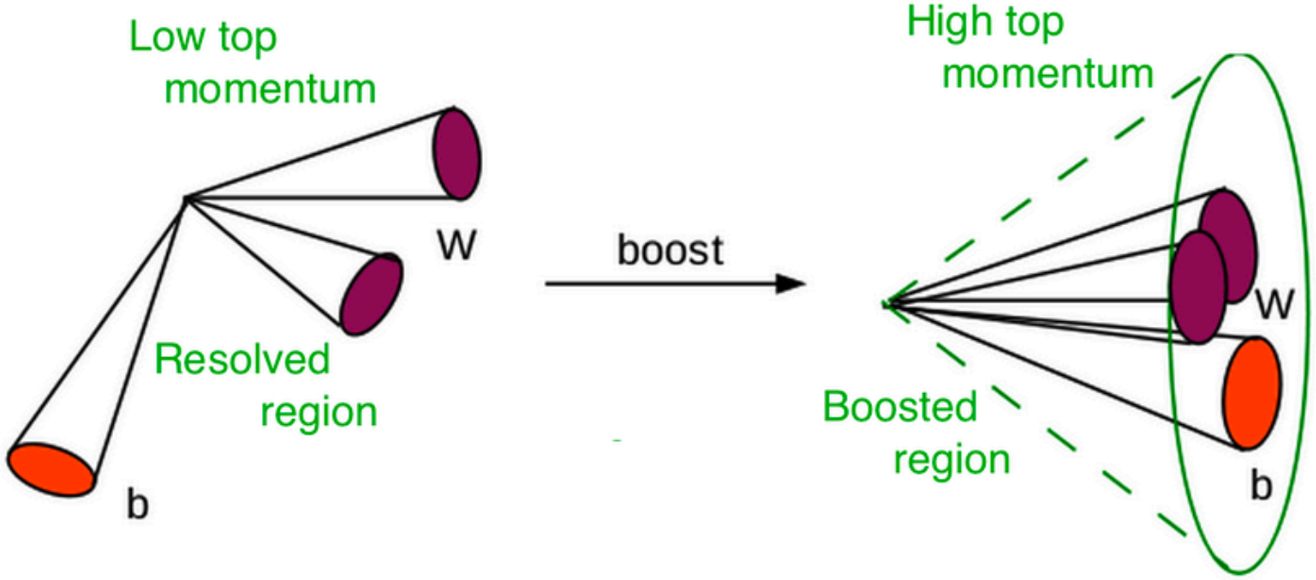
Georgios Bakas on behalf of the CMS collaboration

TOP-18-013

Towards accessing high p_T with precision



- ✦Explore the kinematic region beyond the reach of the resolved analyses ($p_T > 400$ GeV)
  - ➡Possible to become precise
  - ➡Sensitive to new physics
  - ➡Crucial test for perturbative QCD
- ✦Two distinct final states
  - ➡L + jets (boosted hadronically decaying top quark & resolved leptonically decaying top quark)
  - ➡Hadronic (both boosted top quarks decaying hadronically)

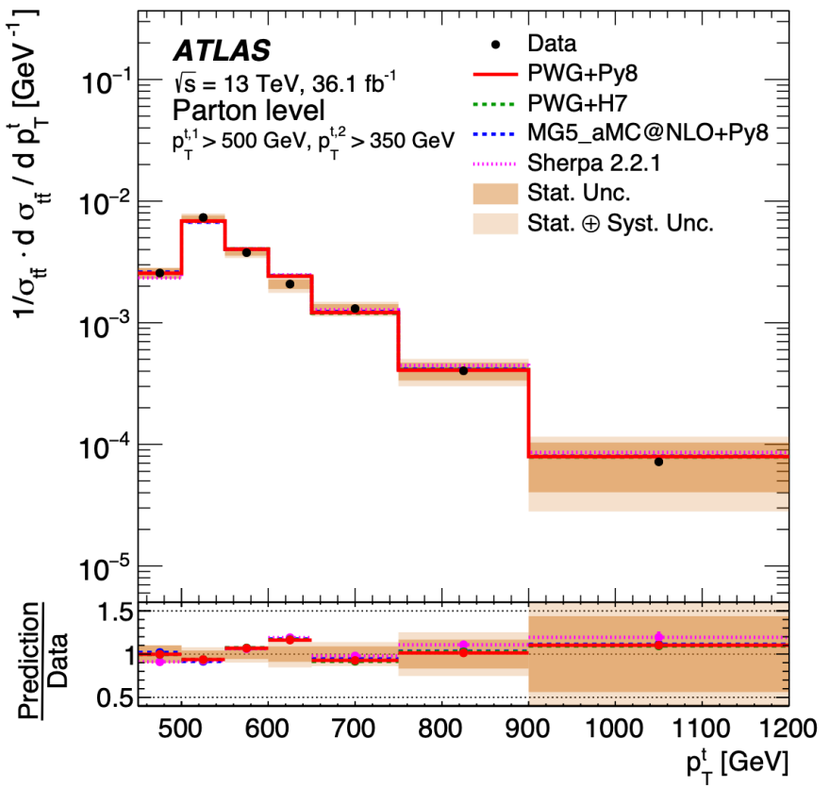


boosted top decaying hadronically reconstructed as a large-radius jet

$p_T \sim 2M_t/R \sim 400$ GeV

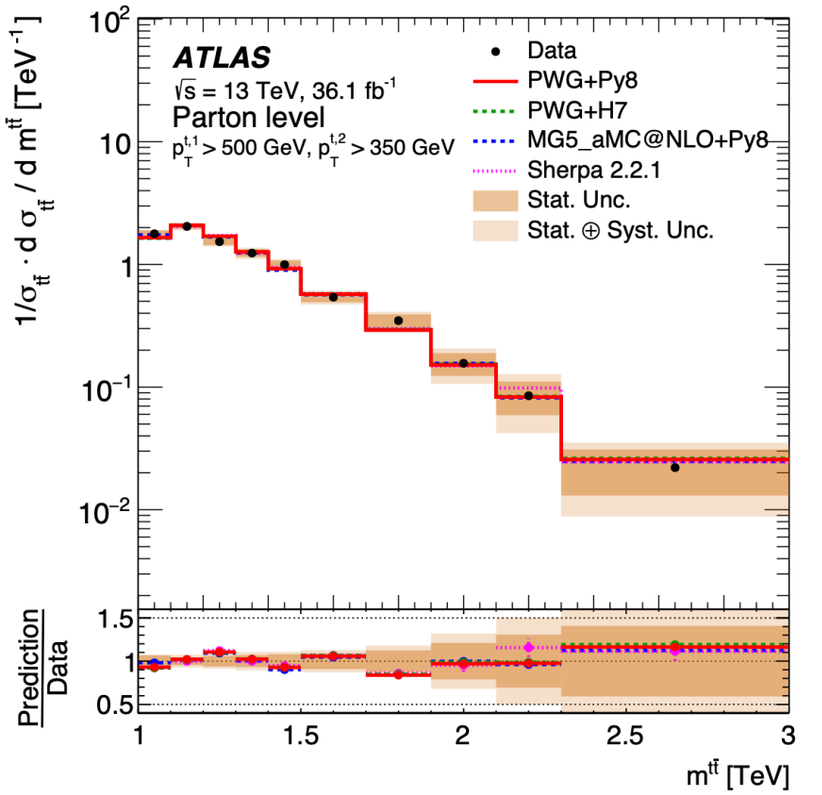
Towards accessing high p_T with precision

- ATLAS:
  - TOPQ-2016-09
  - $t\bar{t}$ pair to all-hadronic final states
  - 13 TeV
  - 2 large-R radius jets
    - Leading p_T >500 GeV
    - Subleading p_T > 350 GeV
  - Probes up to 1.2 TeV for leading jet p_T
  - Probes up to 3 TeV for m_{ttbar}



- CMS:
  - Probes up to 1.5 TeV for the leading jet p_T
  - probes up to 4 TeV for m_{ttbar}

Compatible results regarding the inclusive cross section



Analysis Overview

- 2016 dataset
  - 35.9 fb-1
- Differential cross section:
  - Absolute
  - Normalized
  - 2 final states: hadronic, l+jets

- L + jets channel Deliverables:
  - Hadronically decaying top p_T & |y|

- Hadronic channel Deliverables:
  - Leading & subleading top p_T & |y|
  - $t\bar{t}$ system mass, p_T & rapidity

Unfold to Parton & Particle Levels

Analysis overview (hadronic)

Hadronic

-Final State:

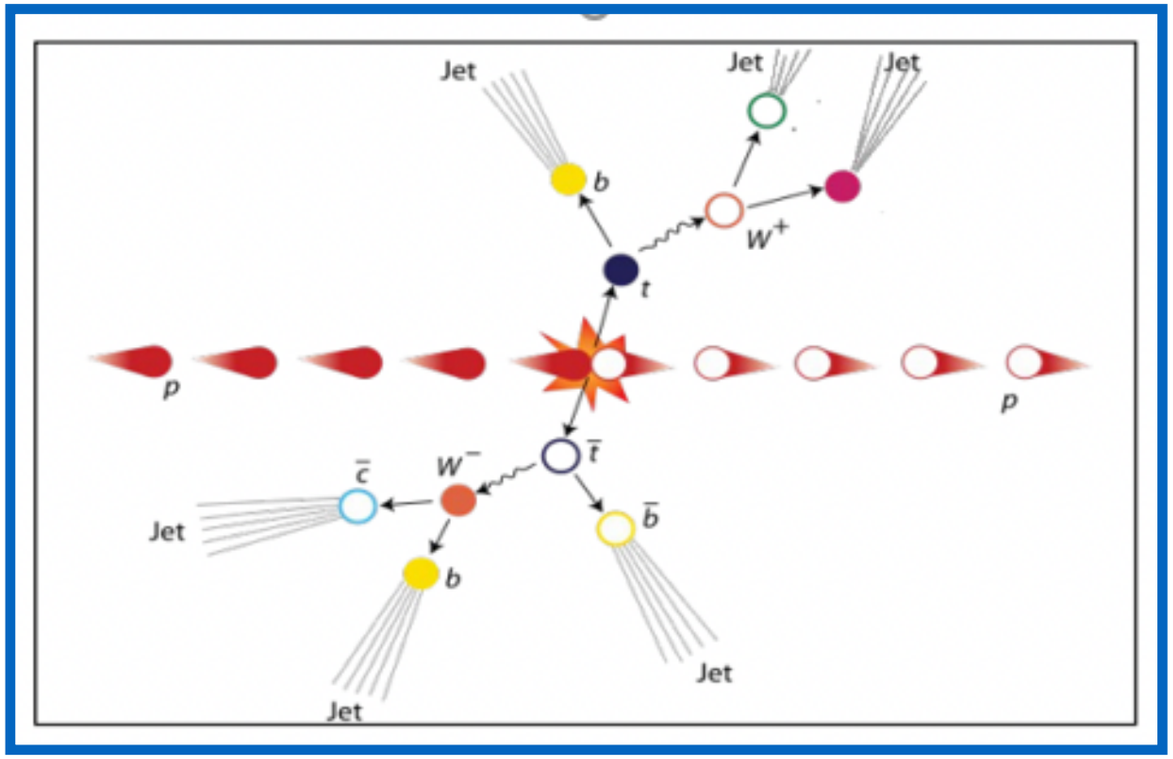
- Two large R jets containing the decay products of top

-Trigger:

- Two AK8 (anti-kt, R=0.8) jets @ HLT and b tagging

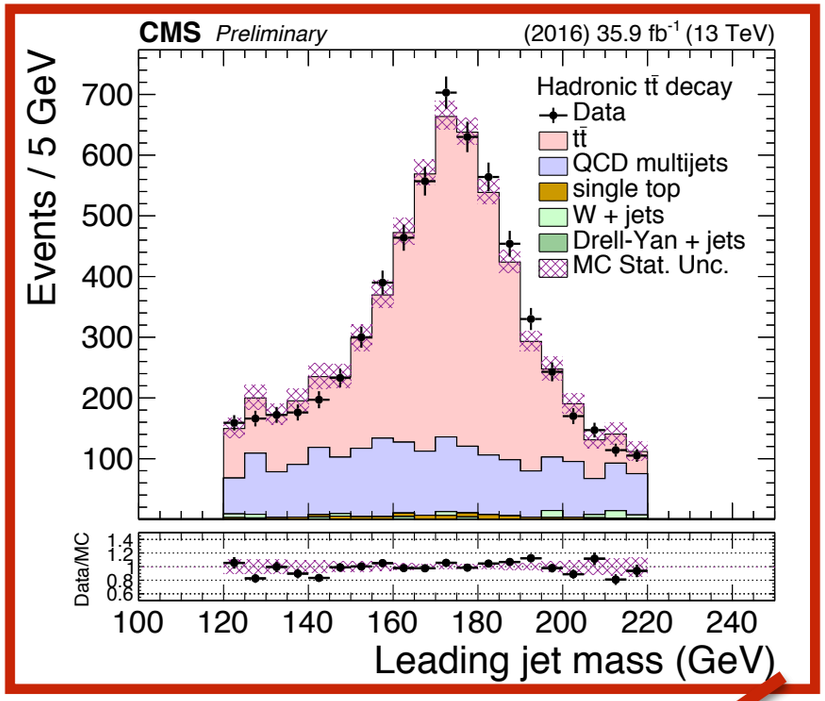
- Selection:

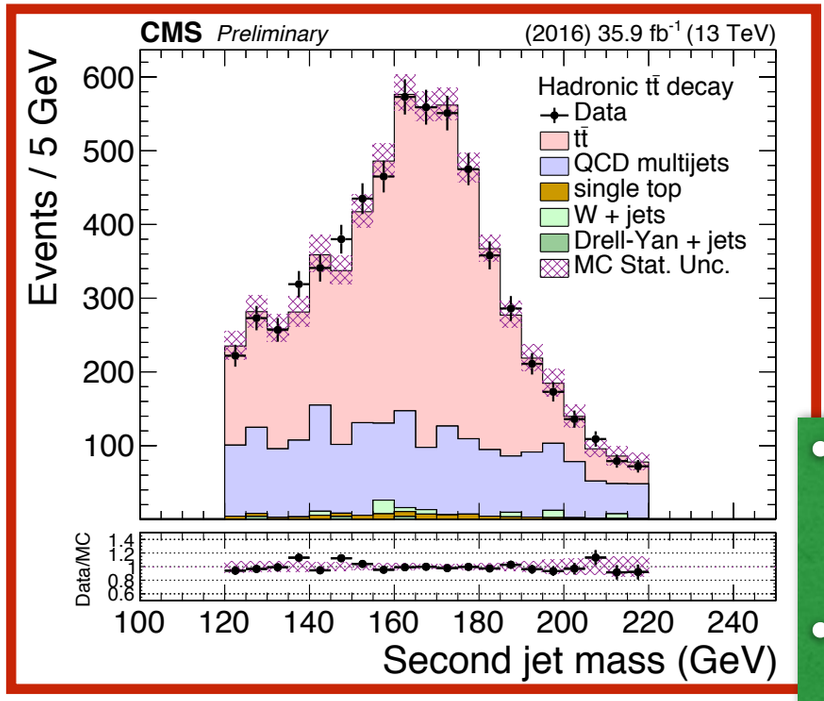
- Two AK8 (anti-kt, R=0.8) jets with p_T > 400 GeV,
- "ttbar event tagging" NN using jet substructure variables as inputs
- Categories based on subjet b tagging
  - 0-btag: control
  - 2-btag: signal
- Backgrounds
  - QCD dominant: taken from data
  - Others (ST, W/Z+jets): negligible



Selection and $t\bar{t}$ reconstruction with NN → ML fit to the mass of the leading jet → Signal Extraction → Unfolding

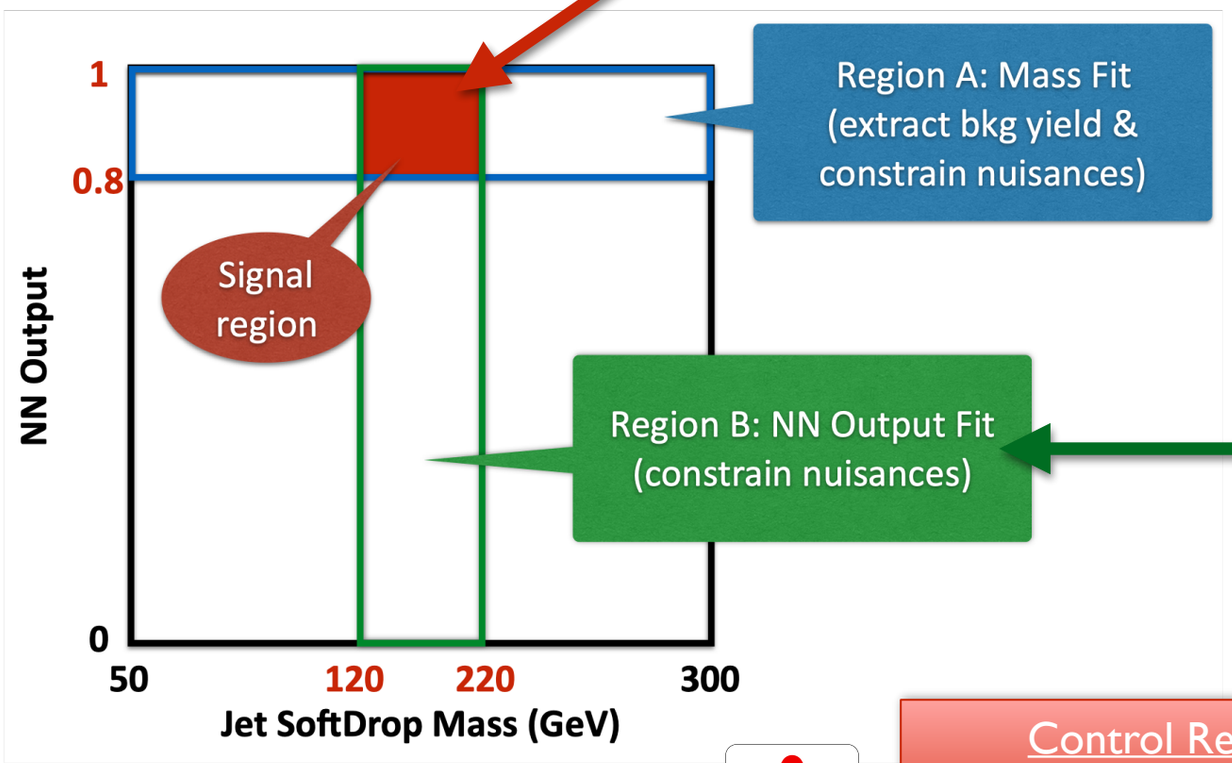
Selection & Analysis Regions (hadronic)

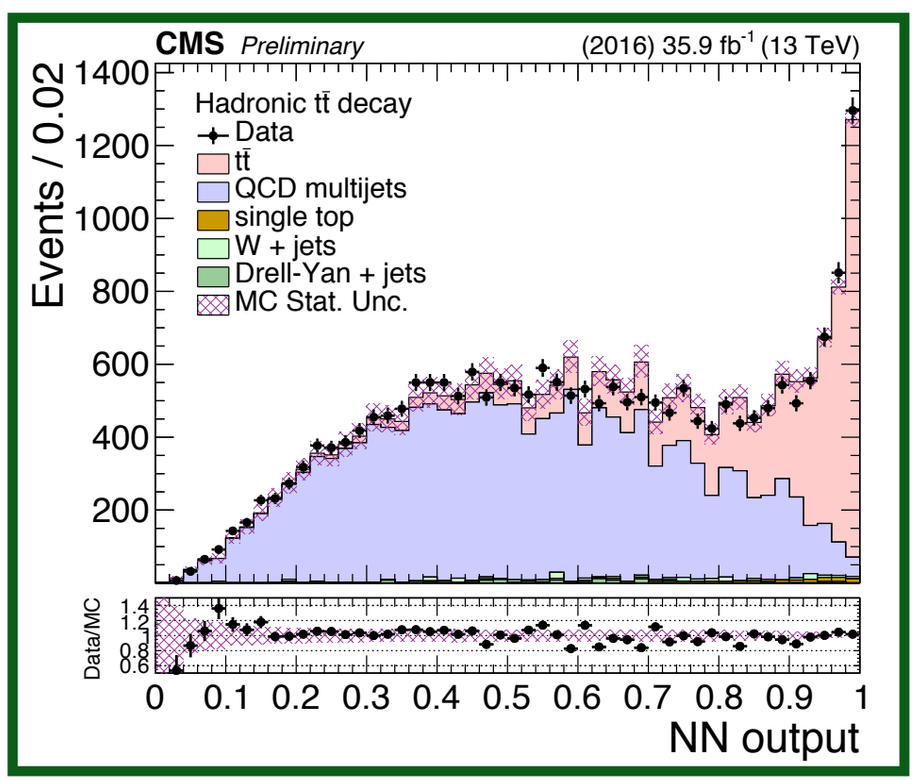




Observable	Requirement
N_{jets}	> 1
N_{leptons}	$= 0$
$p_{\text{T}}^{1,2}$	$> 400\,\text{GeV}$
$m_{SD}^{1,2}$	50–300 GeV

- NN CUT (0.8)compromise between enough tt events and signal purity
- b tagging Medium WP (2 subjets should be b tagged)

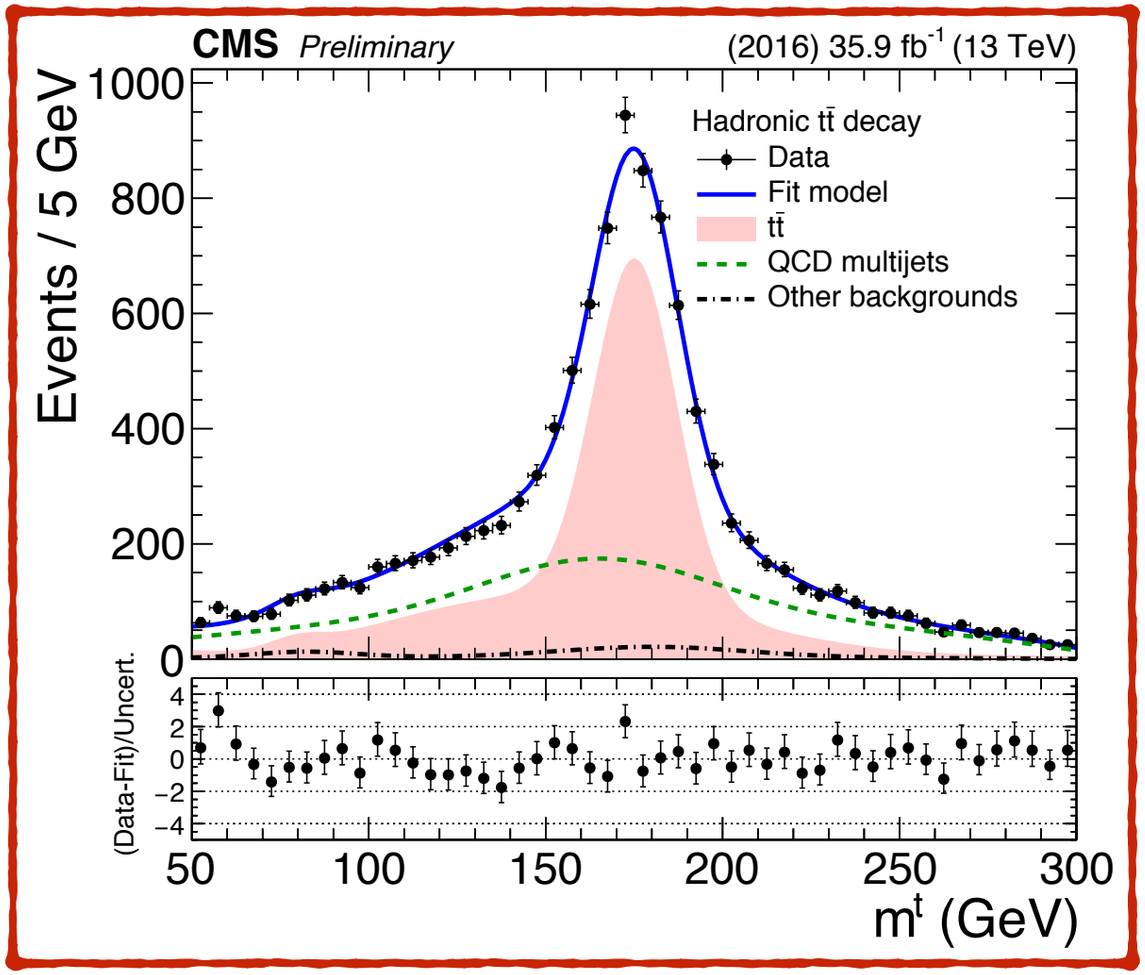




Control Region:
SR kinematics
invert the b tag requirement

Signal Extraction (hadronic)

$$D(m^t) = N_{t\bar{t}} T(m^t; k_{scale}, k_{res}) + N_{qcd}(1 + k_{slope} m^t) Q(m^t) + N_{bkg} B(m^t)$$



Parameter	Value
k_{res}	0.960 ± 0.026
k_{scale}	1.002 ± 0.002
k_{slope}	$(5.7 \pm 1.4) \times 10^{-3}$
N_{bkg}	400 ± 255
N_{QCD}	4539 ± 247
$N_{t\bar{t}}$	6238 ± 181

$t\bar{t}$ Signal Strength
r_{tt} = 0.64

fiducial yield

Data

Transfer factor from SR_A to SR (taken from CR_A/CR)

Bkg shape (taken from CR in data)

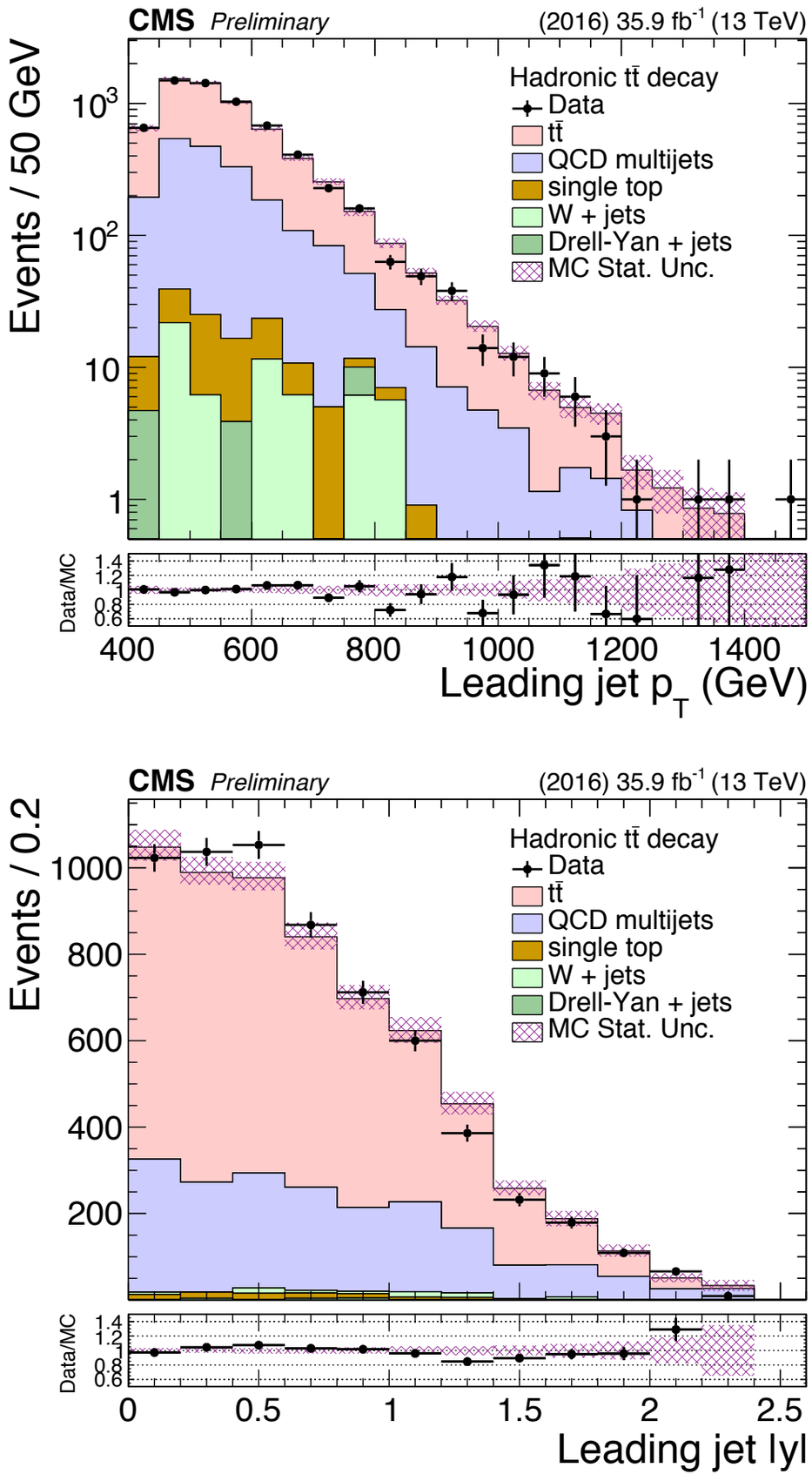
$$S_{fid}(x_{reco}) = D(x_{reco}) - C^{yield}_{bkg} \cdot N^{fit}_{bkg} \cdot C^{shape}_{bkg}(x_{reco}) \cdot B(x_{reco})$$

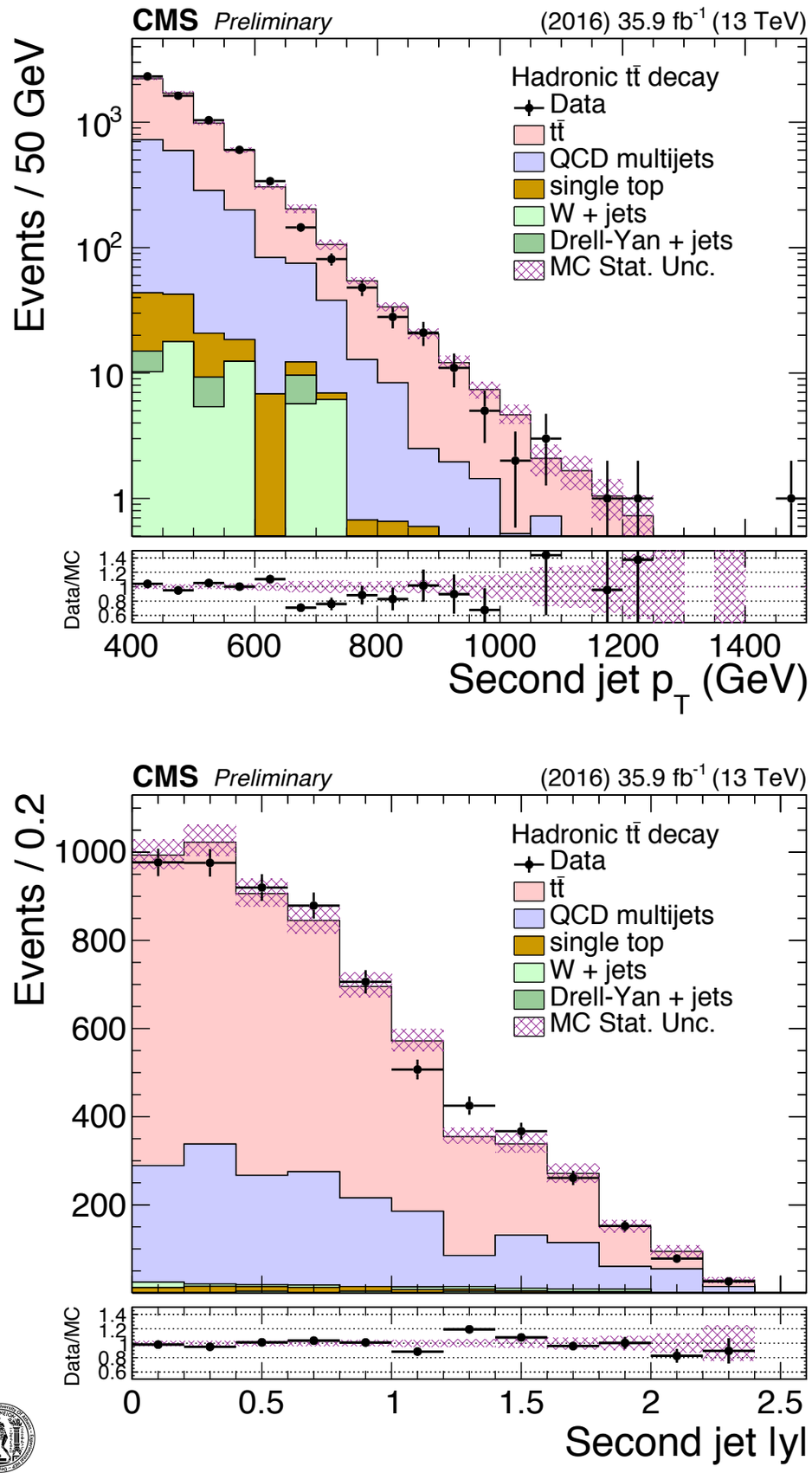
Bkg yield (fit in region SR_A)

Bkg shape correction (taken from MC)

Process	Number of events
tt̄	4244 ± 127
QCD multijets	1876 ± 102
Single t	83 ± 41
W+jets	58 ± 29
Z+jets	12 ± 6
Total	6273 ± 171
Data	6274

Top kinematic distributions (hadronic)

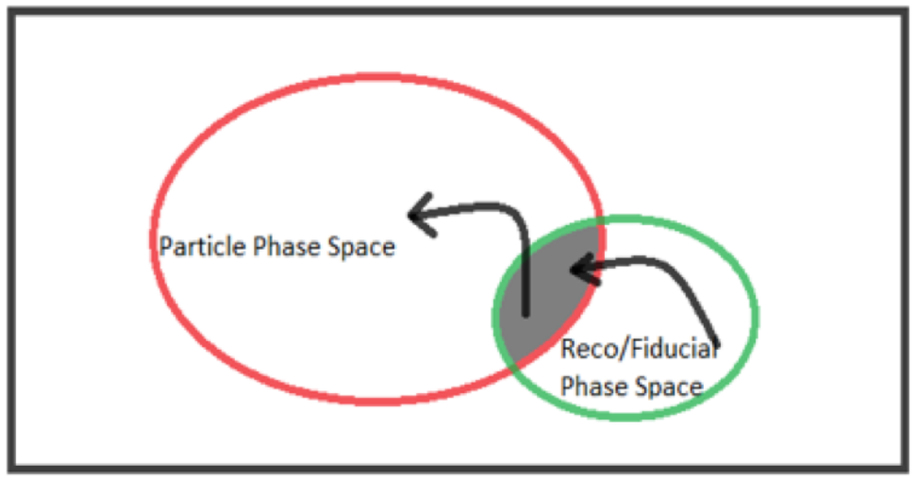




Parton & Particle levels (hadronic)

Parton

Observable	Requirement
$p_T^{t,\,tbar}$	> 400 GeV
$\lvert\eta^{t,tbar}\rvert$	<2.4
m_{ttbar}	> 800 GeV



Particle

top candidates: AK8 genjets

Observable	Requirement
N_{jets}	> 1
$p_T^{jet1,\,2}$	> 400 GeV
$\lvert\eta^{jet1,\,2}\rvert$	<2.4
$m_{SD}^{jet1,2}$	(120,220)GeV
m_{ttbar}	> 800 GeV

$$\frac{d\sigma_i^{\mathrm{unf}}}{dx} = \frac{1}{\mathcal{L}\cdot\Delta x_i}\cdot\frac{1}{f_{2,i}}\cdot\sum_j\left(R_{ij}^{-1}\cdot f_{1,j}\cdot S_j\right)$$

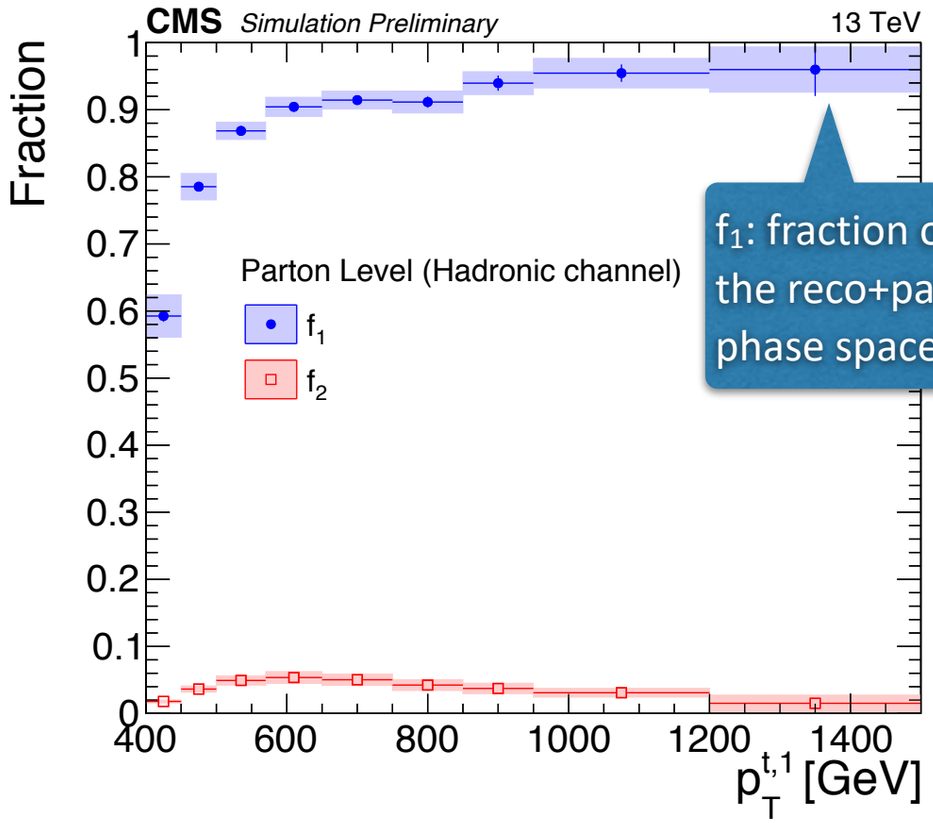
efficiency of the reco+true selection

migration matrix

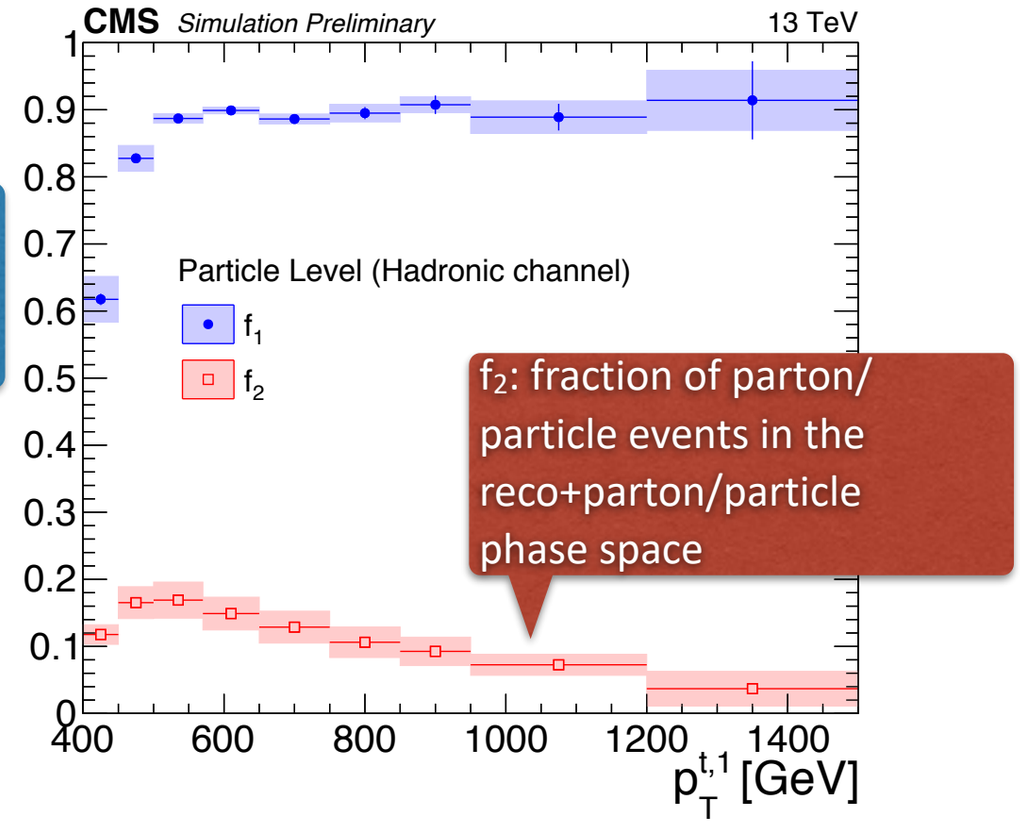
reco efficiency of the reco+true selection

Unfolding: simple response matrix inversion w/o regularisation

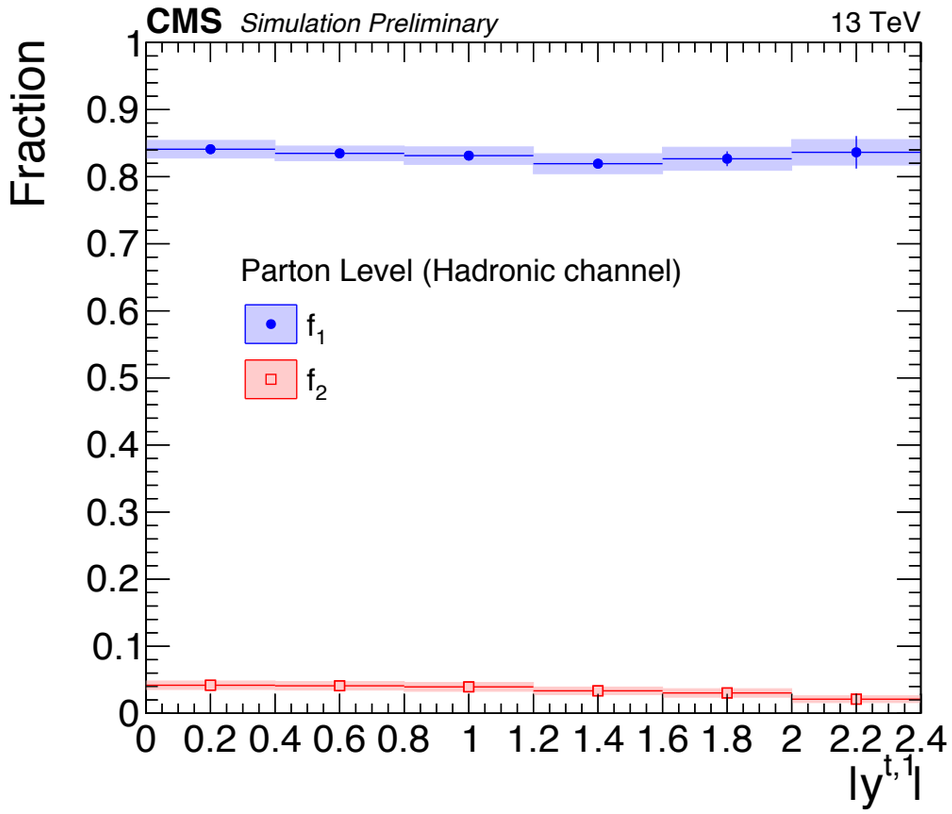
Fractions of parton & particle level definitions

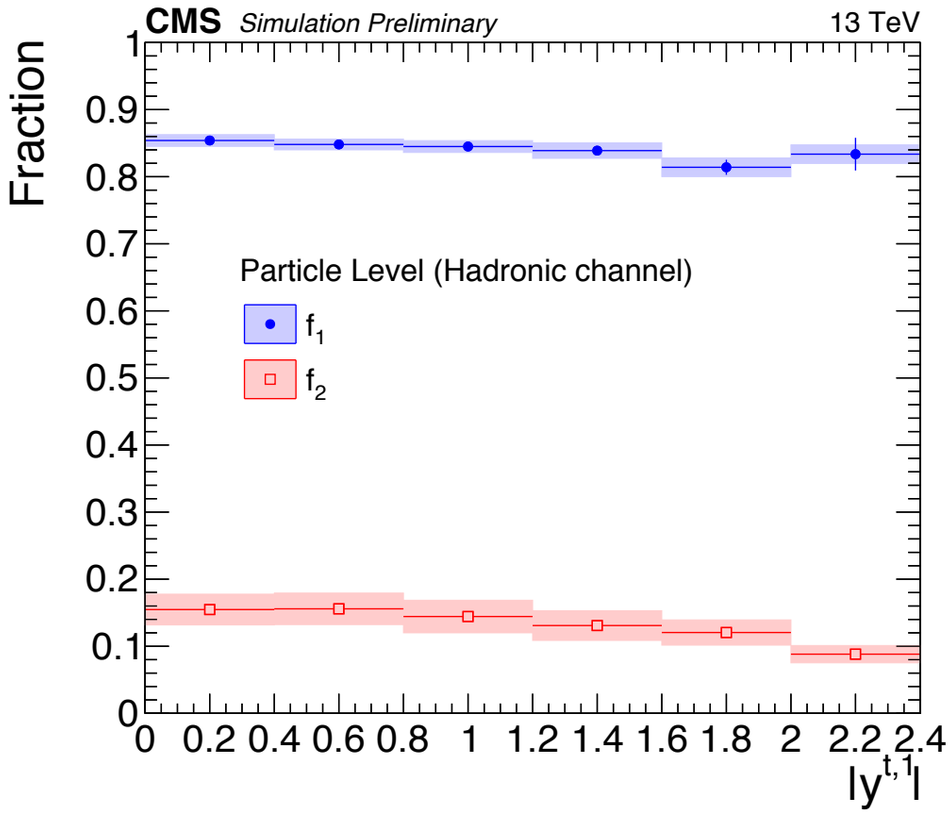


f_1: fraction of reco events in the reco+parton/particle phase space



f_2: fraction of parton/ particle events in the reco+parton/particle phase space



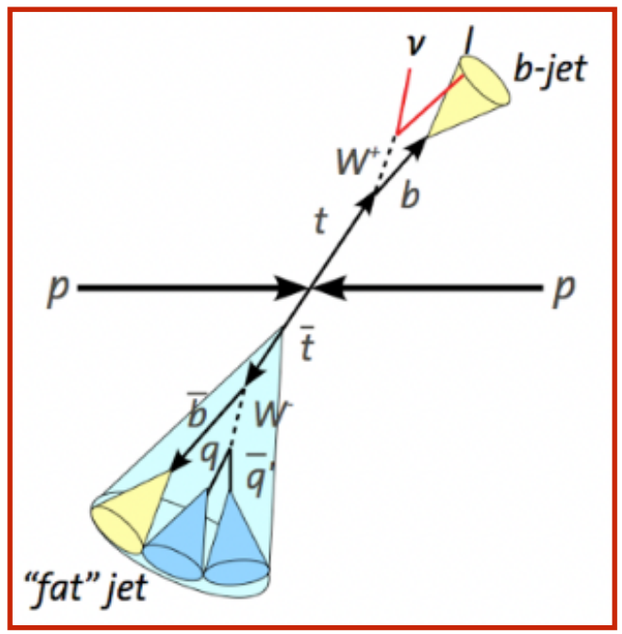


Analysis Overview (l+jets)

–Final State:
- Lepton + b jet + MET + t jet

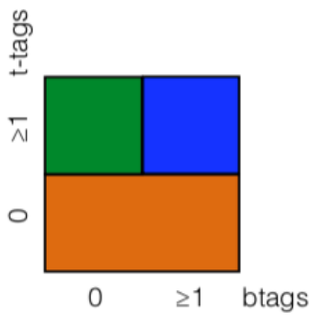
–Trigger:
- Single Lepton: e (μ)
- Two small-R jets

–Selection:
- ==1 e/μ (veto additional leptons)
- ≥1 small-R jet (anti-kt, R=0.4, leptonic top decay)
- ≥1 large-R jet (anti-kt, R=0.8, hadronic top decay)
- E_T^{Miss}
- Characterise events by whether b and t jet candidate pass tagging requirements

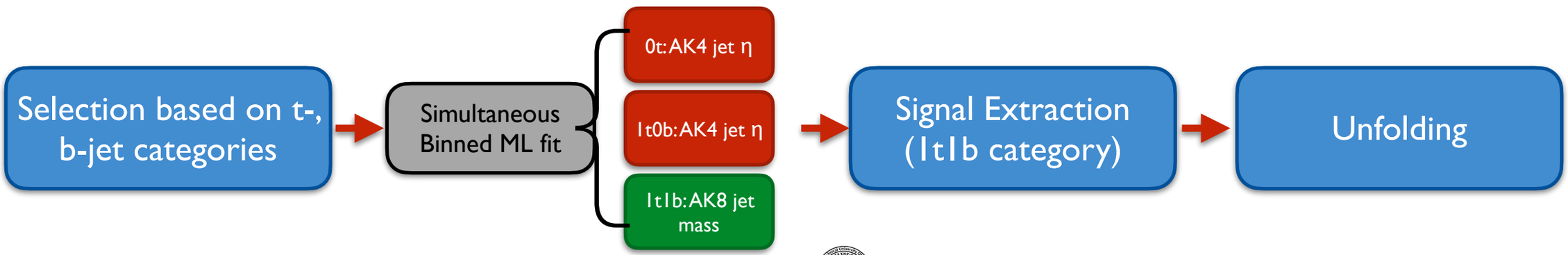


- b tagging:
  - AK4 jet
  - Medium b-tag WP

- t tagging:
  - AK8 jet
  - 105 < m_{top} candidate < 220 GeV
  - N-subjetiness ratio τ_{32} < 0.81 (Loose WP)
  - No b-tagging to improve signal acceptance



- Categories:
  - 0t: t-jet candidate fails to pass the top tagging requirements
  - 1t0b: t-jet candidate passes the top tagging requirements but b-jet candidate fails the b tagging requirements
  - 1t1b: both t-jet and b-jet candidates pass the respective tagging requirements



Background Estimation (l+jets)

- Several sources of background:
  - Non-signal ttbar, single top quark, W+jets, Z+jets, diboson, QCD
  - All except QCD are modelled using MC
  - QCD: data driven technique

- QCD:
  - Data in a sideband dominated by QCD:
    -invert the requirement on the lepton: exactly one with 0.1 < miniIso < 0.2; Medium e ID
  - Subtract expected non-QCD contributions
  - Compare QCD shape in data sideband against QCD MC in data sideband signal region
    Good agreement
    Largest deviations in e p_T and η

Simultaneous Likelihood Fit

- $t\bar{t}$ signal strength, t tagging efficiency SF and bkg normalisations determined using simultaneous Binned Likelihood fit
- Event categories fitted simultaneously:
  - 0t: Background dominated
  - 1t0b
  - 1t1b: Signal Dominated

} → AK4 jet η

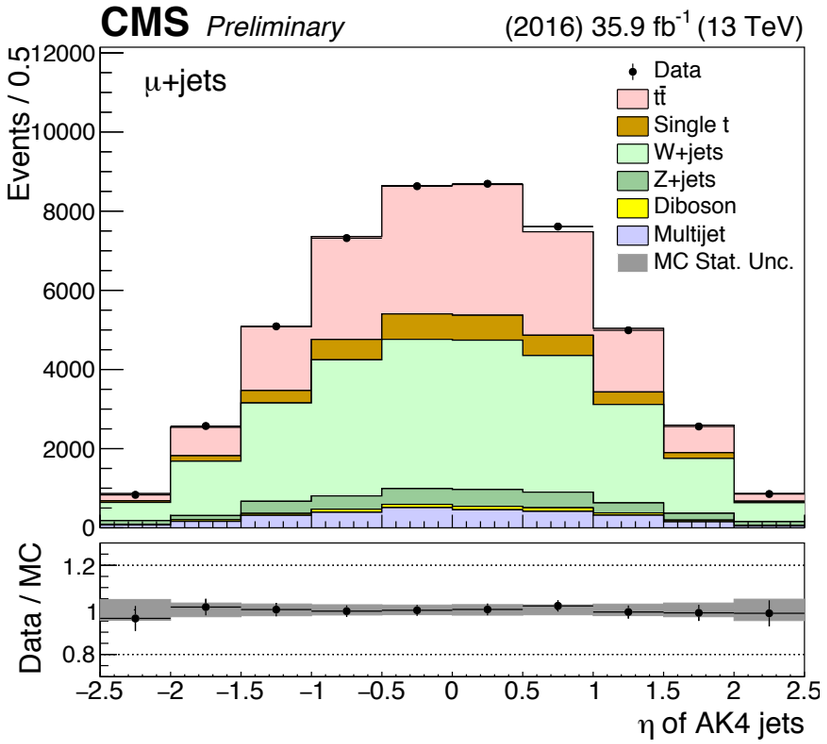
} → AK8 jet SD mass

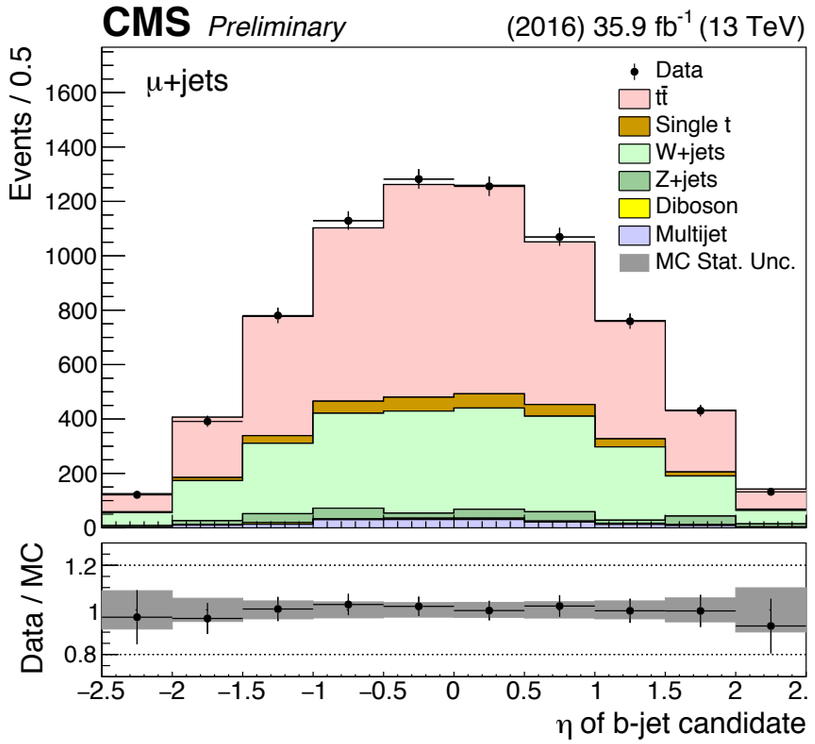
} Best signal/background discrimination, QCD well modelled

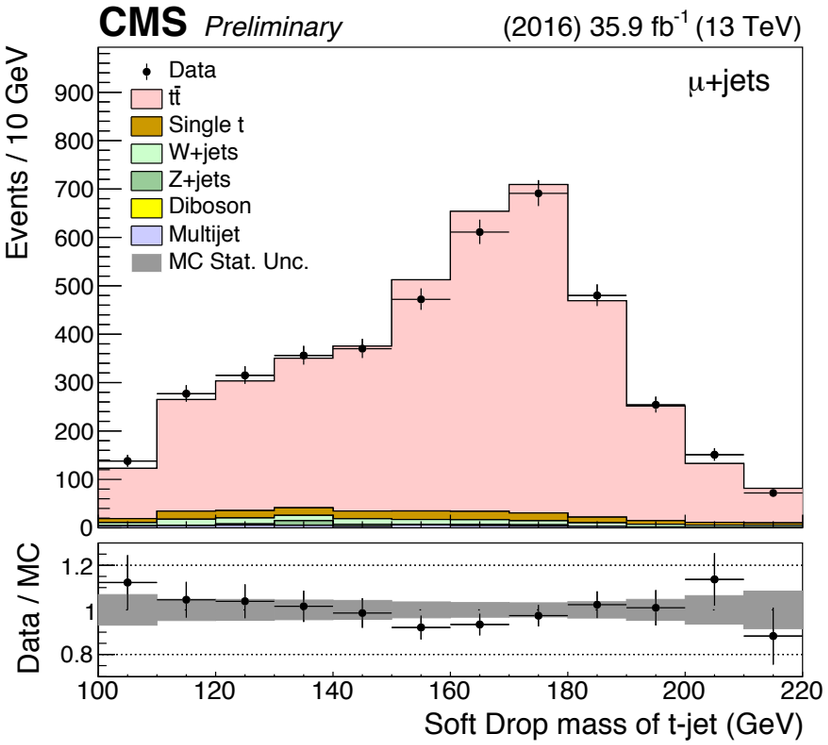
- Systematics, normalisations enter as nuisances
- e/μ + jets channels are fitted simultaneously

$t\bar{t}$ Signal Stength: r = 0.81 ± 0.05

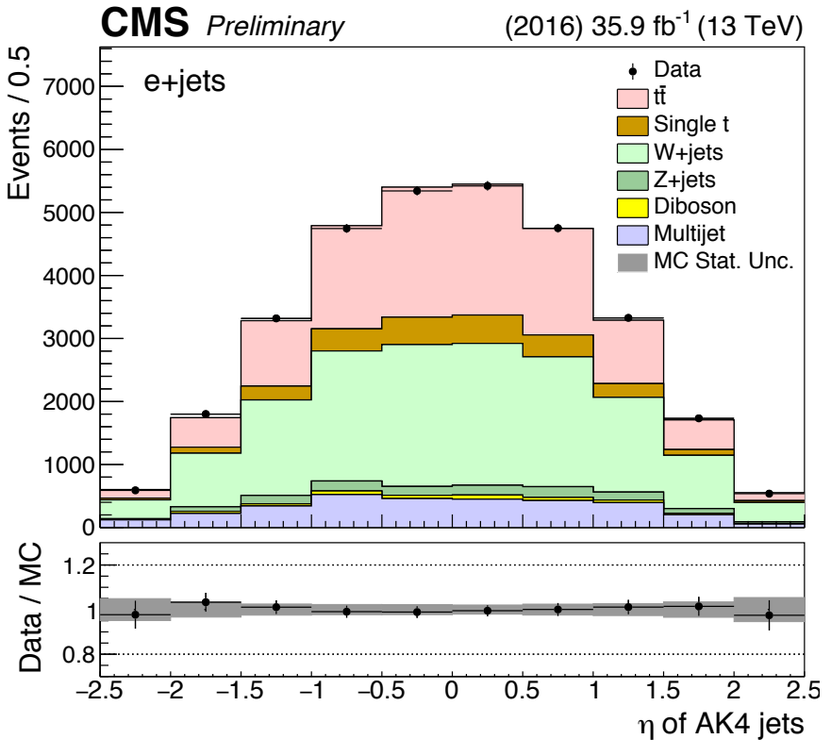
Post fit kinematic distributions (l+jets)



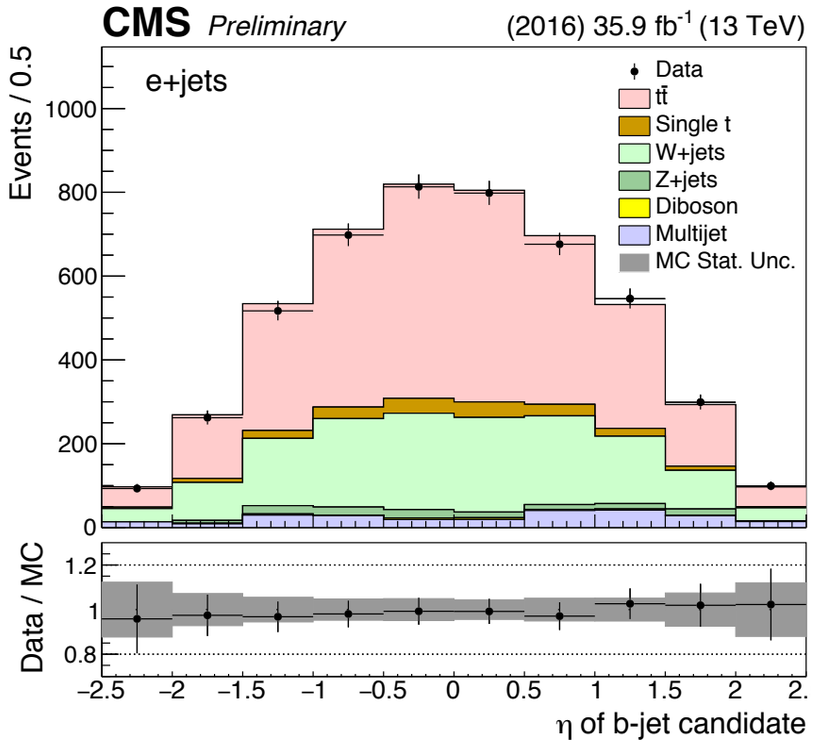




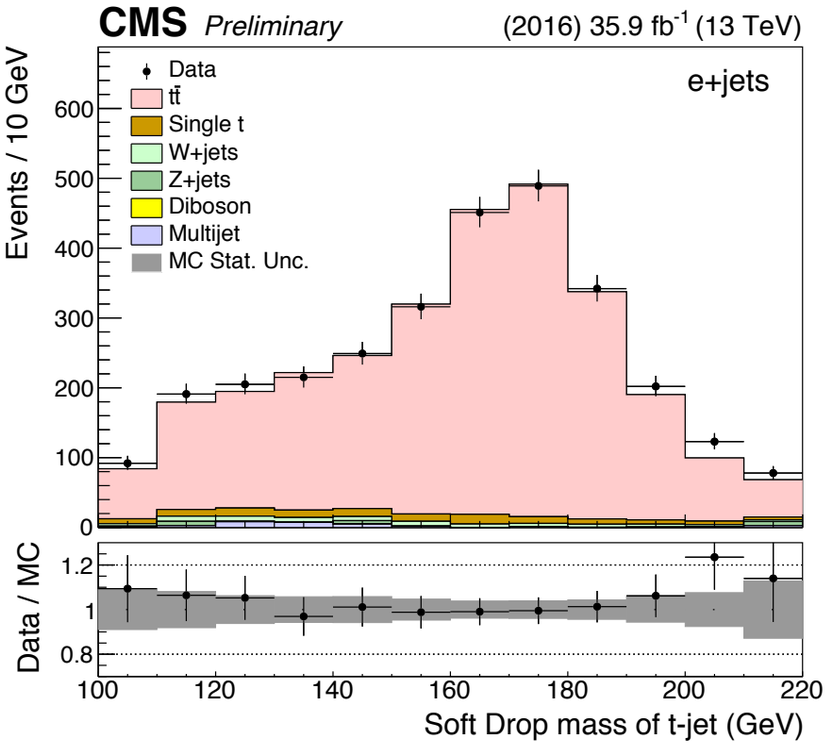
AK4 jet η, 0t



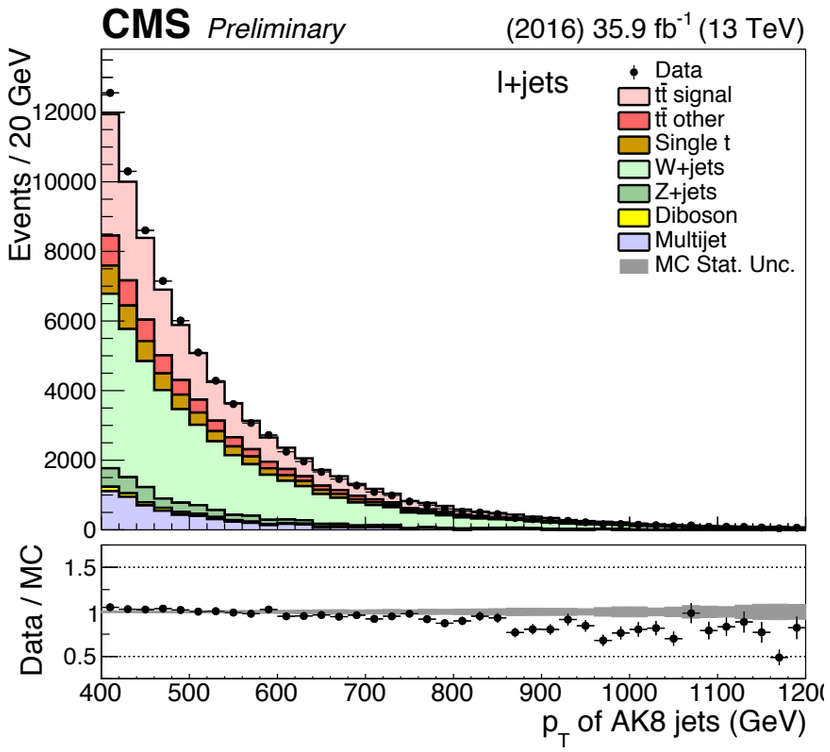
AK4 jet η, 1t0b

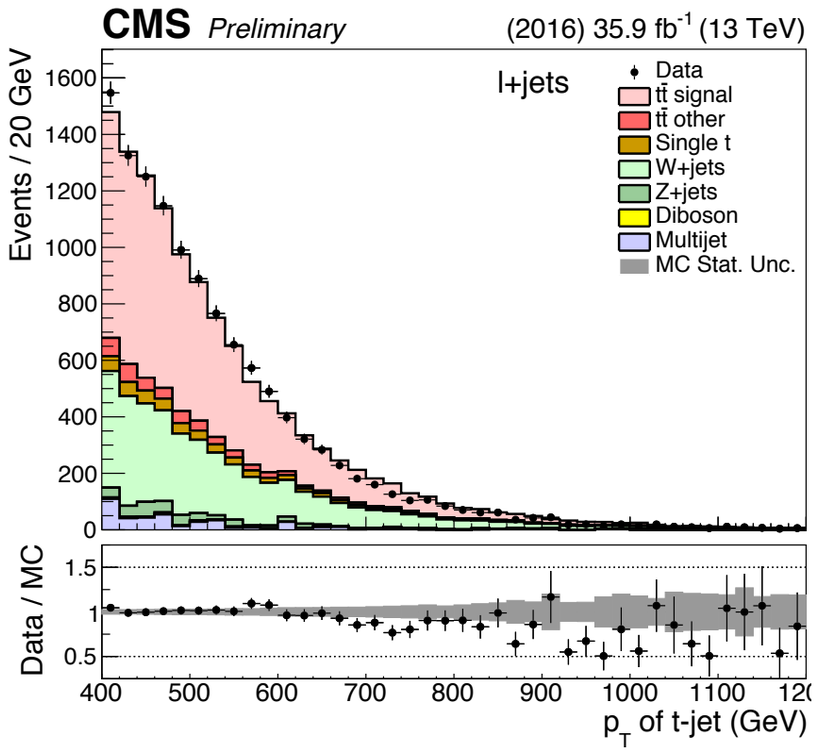


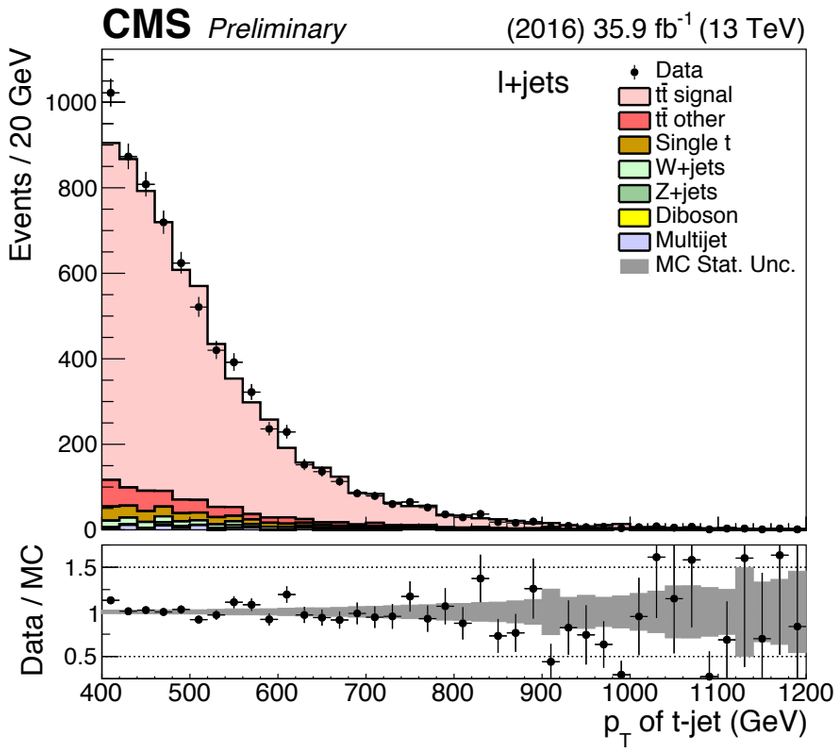
AK8 jet softDrop Mass, 1t1b



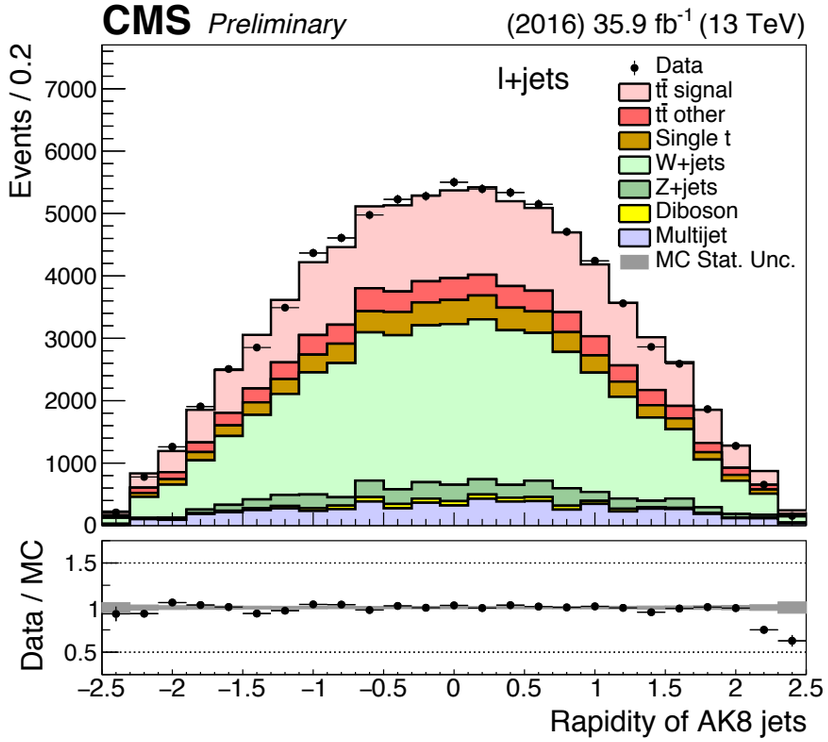
Post fit kinematic distributions (l+jets)



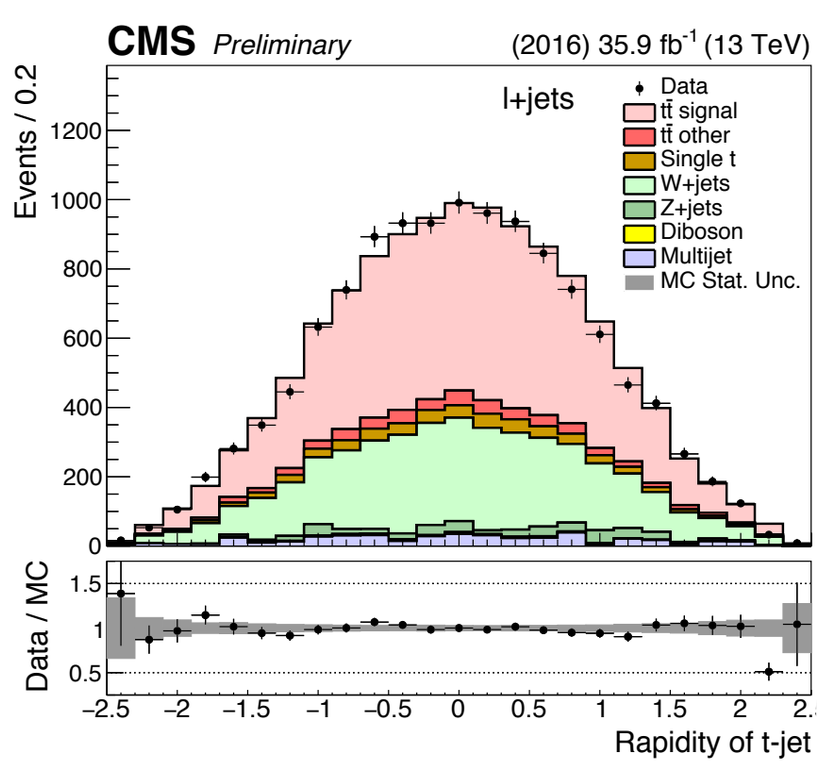




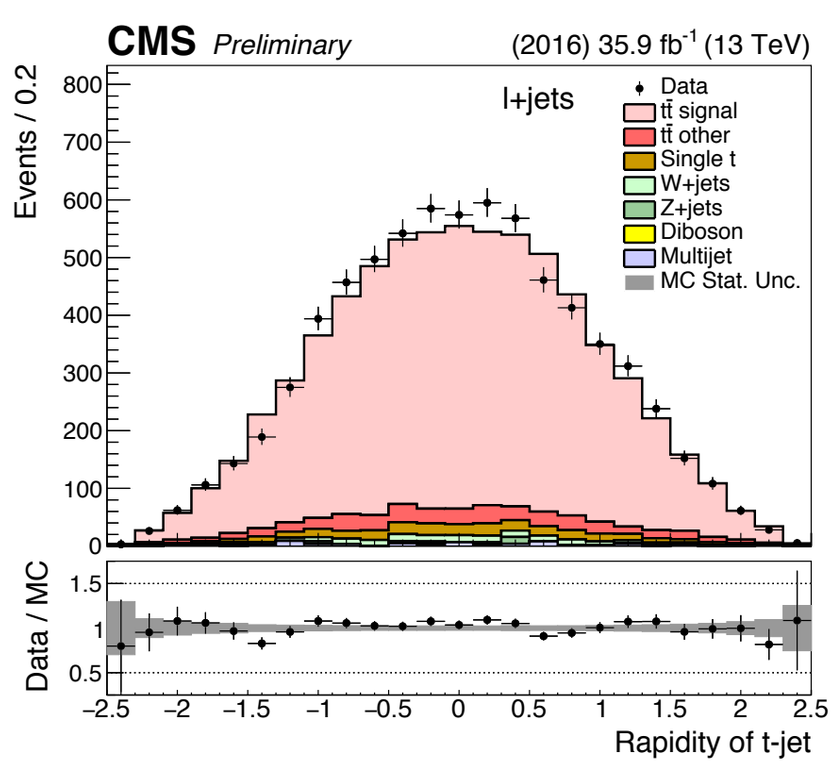
0t



1t0b



1t1b



Parton & Particle levels (l+jets)

Parton

Selection:

- Confined to semi-leptonic ttbar with p_T >400GeV for the hadronically decaying top quark
- ==1 μ/e
- Top Quark: sign opposite of the lepton

Particle

Selection:

- ≥1 particle level AK8 jet:
  - p_T > 400GeV
  - |η| < 2.4
  - 105 < m_{jet} < 220 GeV
- ≥1 particle level AK4 jet:
  - p_T > 50 GeV
  - |η| < 2.4
  - Originating from b quark
- ==1 μ/e:
  - p_T > 50GeV
  - |η|< 2.1
- Top jet: leading particle level AK8 jet

Unfolding

Unfold p_T(top), |y| (top) distributions to give cross section at Particle and Parton level of the hadronically decaying top quark

-Measured distribution: background-subtracted data in 1t1b signal region

Unfolding is performed with TUnfold

- Simple matrix inversion
- Without regularisation

Systematic uncertainties

Two categories:

- Experimental: object performance between data and simulation
- Theoretical: related to the simulation itself(acceptance, efficiency, migration matrix)

Experimental

- Jet Energy Scale
- Jet Energy Resolution
- Pileup
- Luminosity

Hadronic

- QCD multijet Production
- Subdominant Bkgs
- Subjet b-tagging efficiency
- Trigger

L + jets

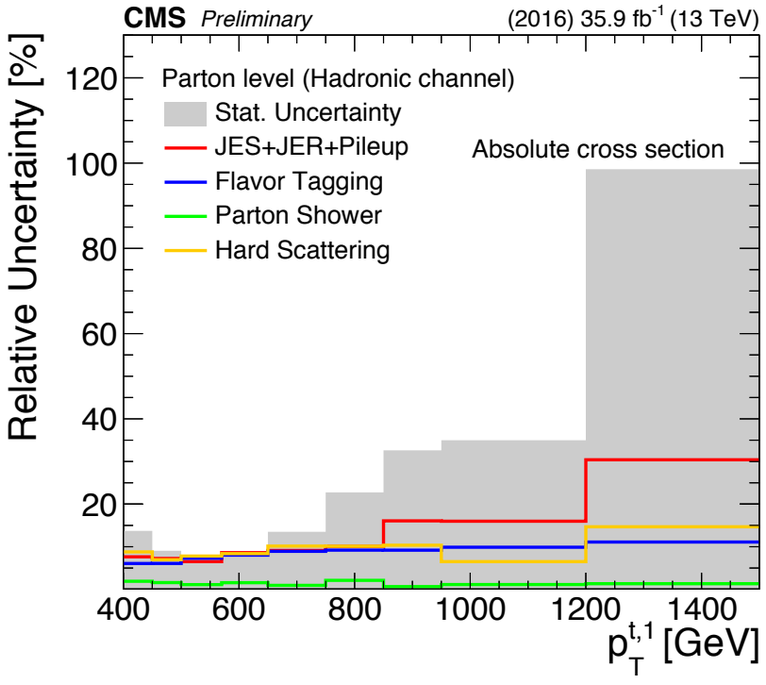
- Background estimate
- t-tagging efficiency
- b-tagging efficiency
- Lepton Identification and Trigger

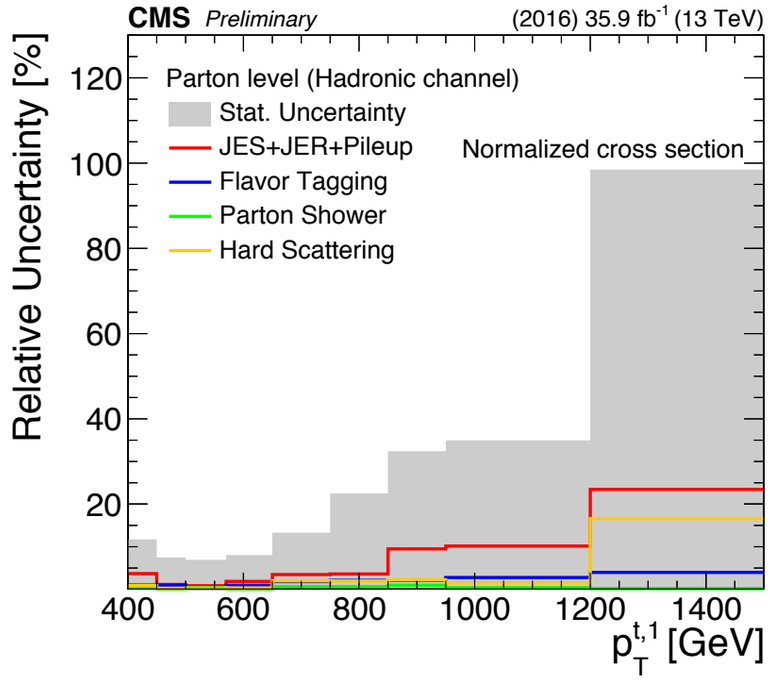
Theoretical: Divided into 2 groups

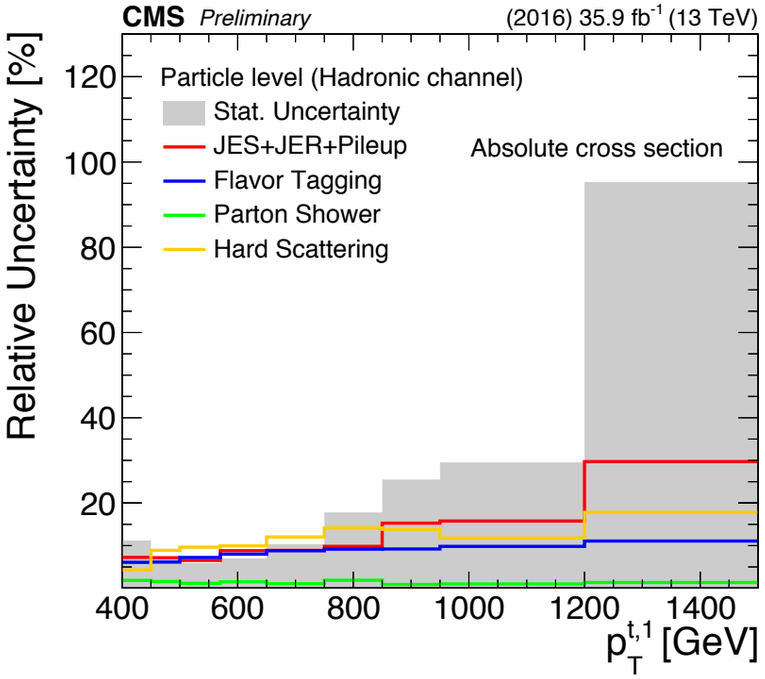
1. Matrix element of the hard process (variations of LHE event weights stored in nominal MC)
2. Modelling of the parton shower and the underlying event (dedicated, alternative MC samples)

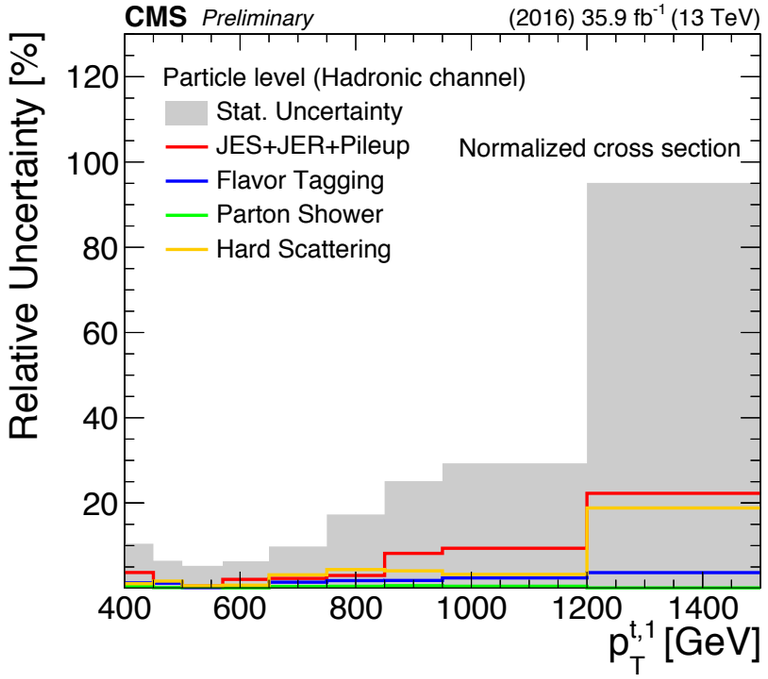
- Parton Distribution Functions
- Renormalisation and factorisation scales:
- Strong coupling constant
- Final State Radiation (FSR): in situ constrained (hadronic)
- Initial State Radiation (ISR)
- Matrix Element-Parton showering matching
- Underlying event tune

Uncertainties vs leading top p_T (hadronic)



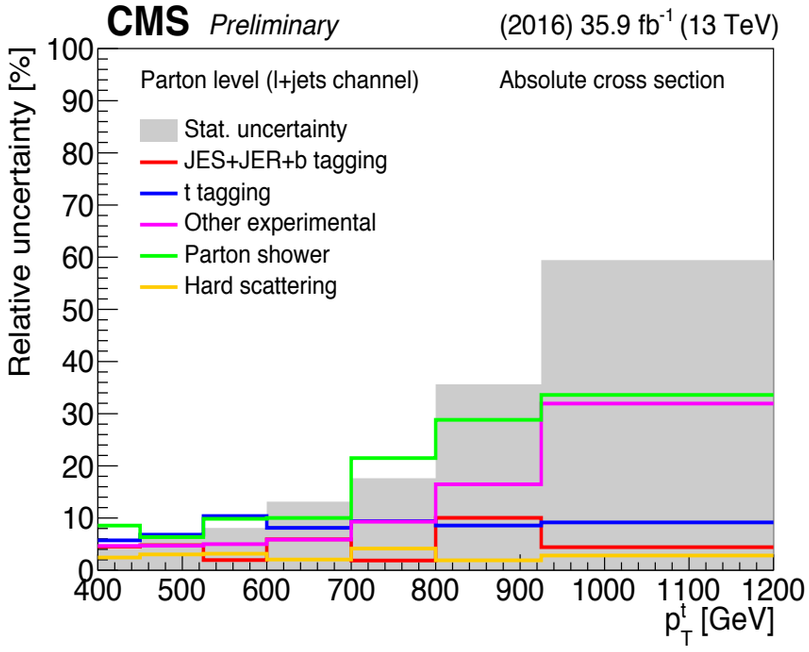


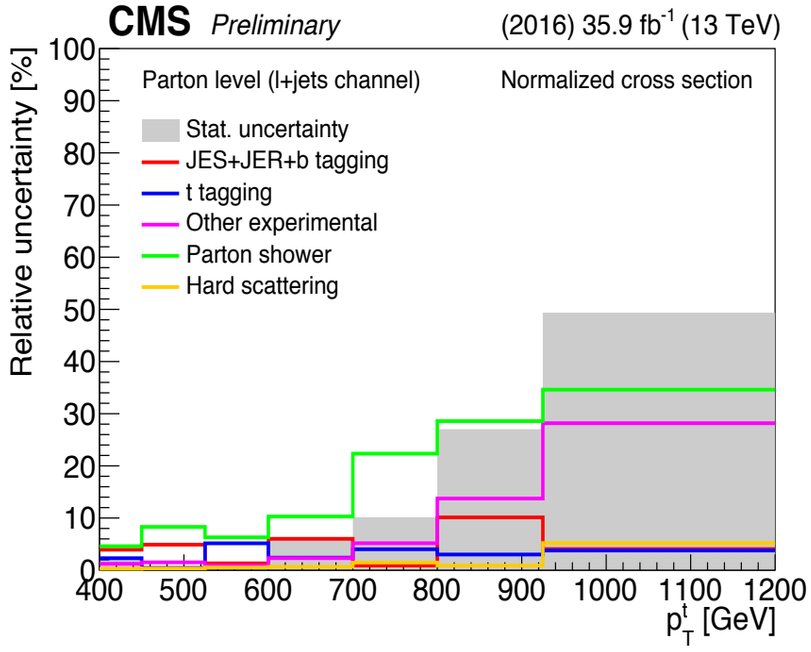


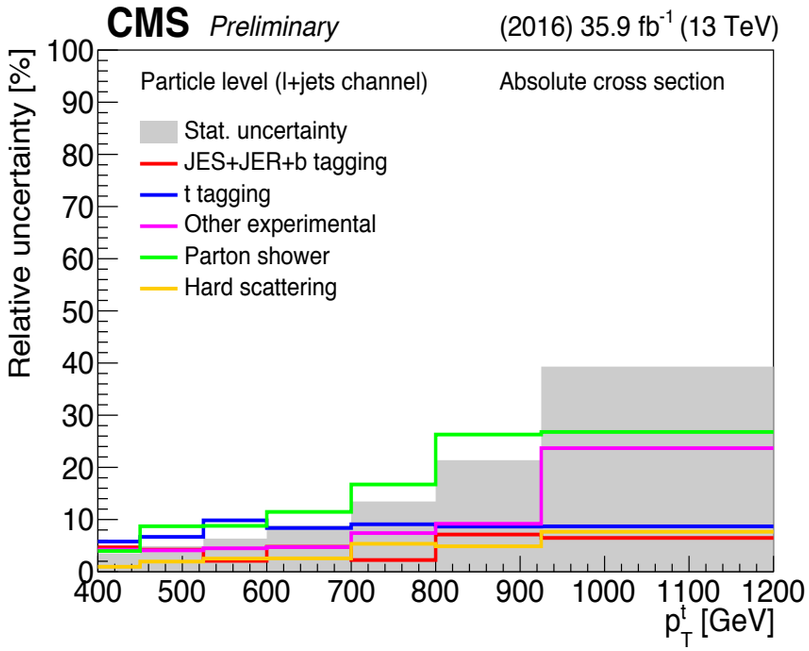


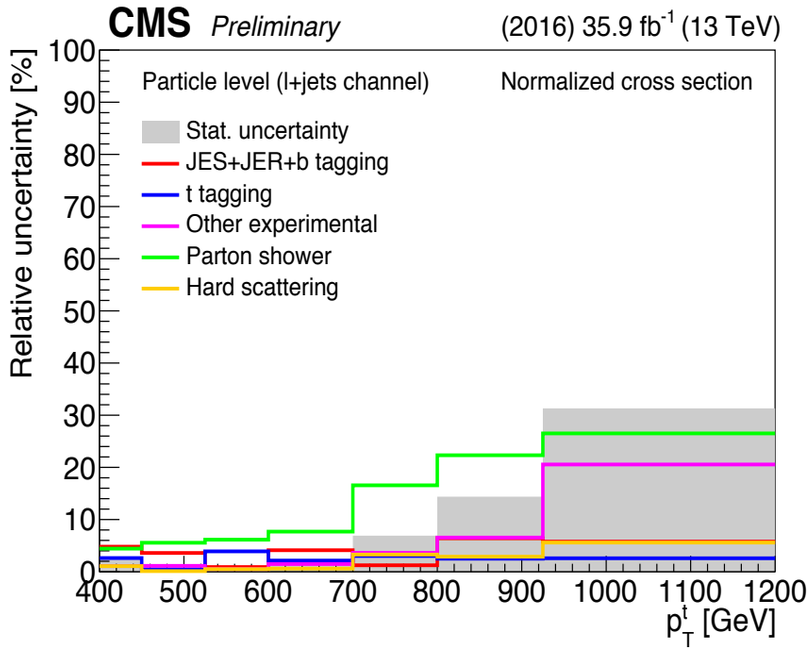
Largest Systematic Uncertainty: JES

Uncertainties vs hadronically decaying top p_T (l+jets)



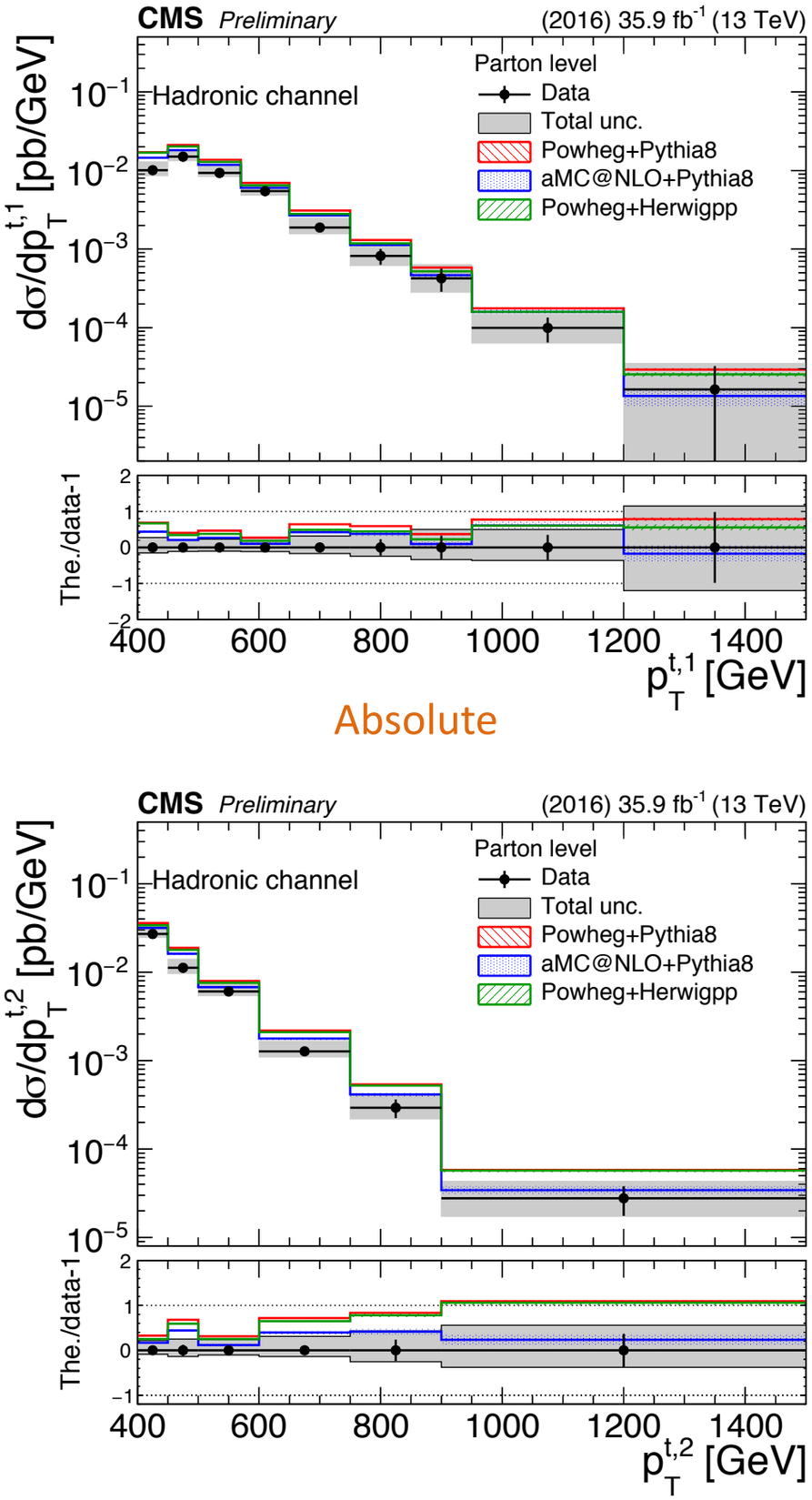




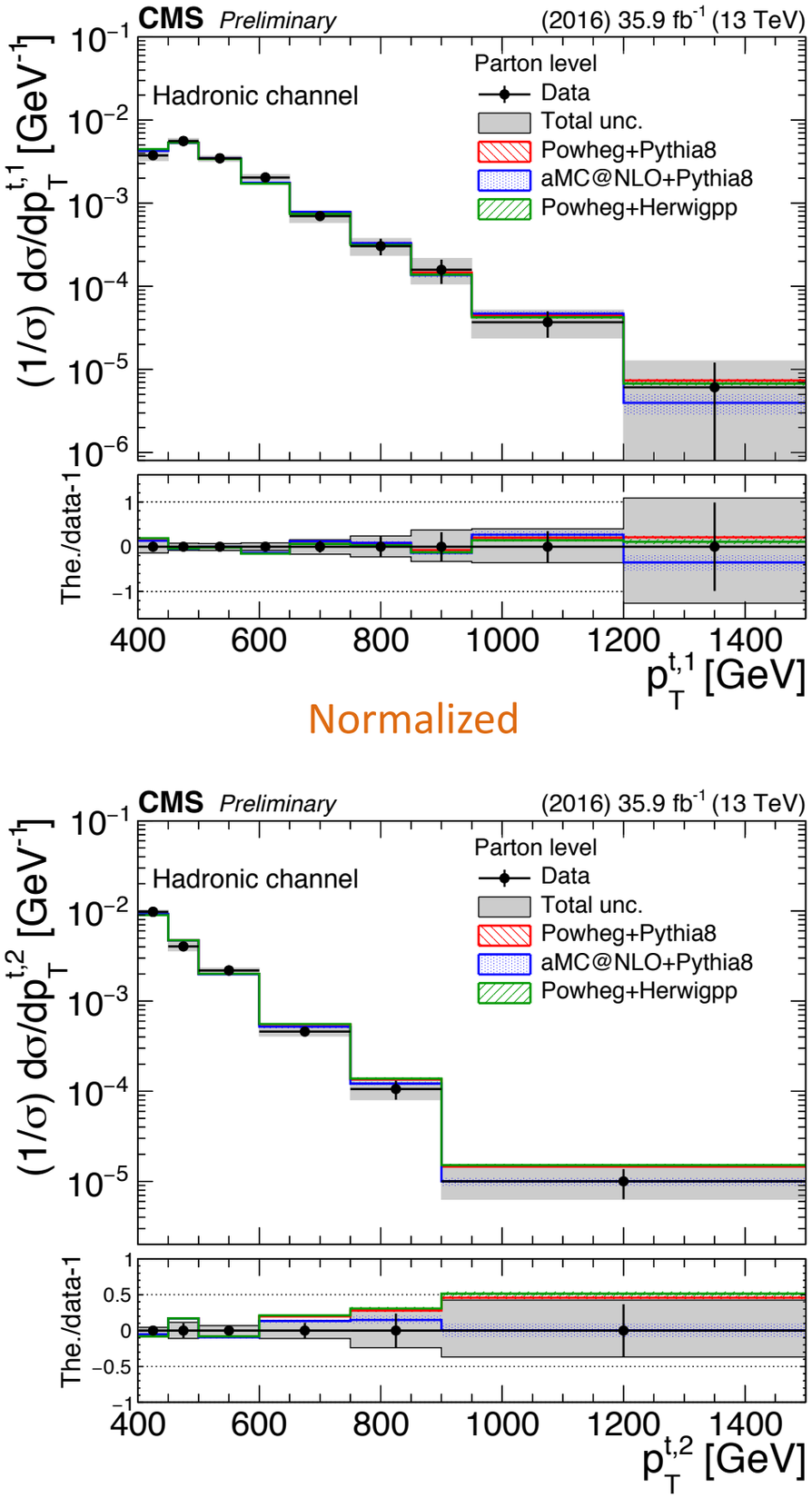


Largest Systematic Uncertainty: Parton Shower

Results (parton, top p_T) (hadronic)



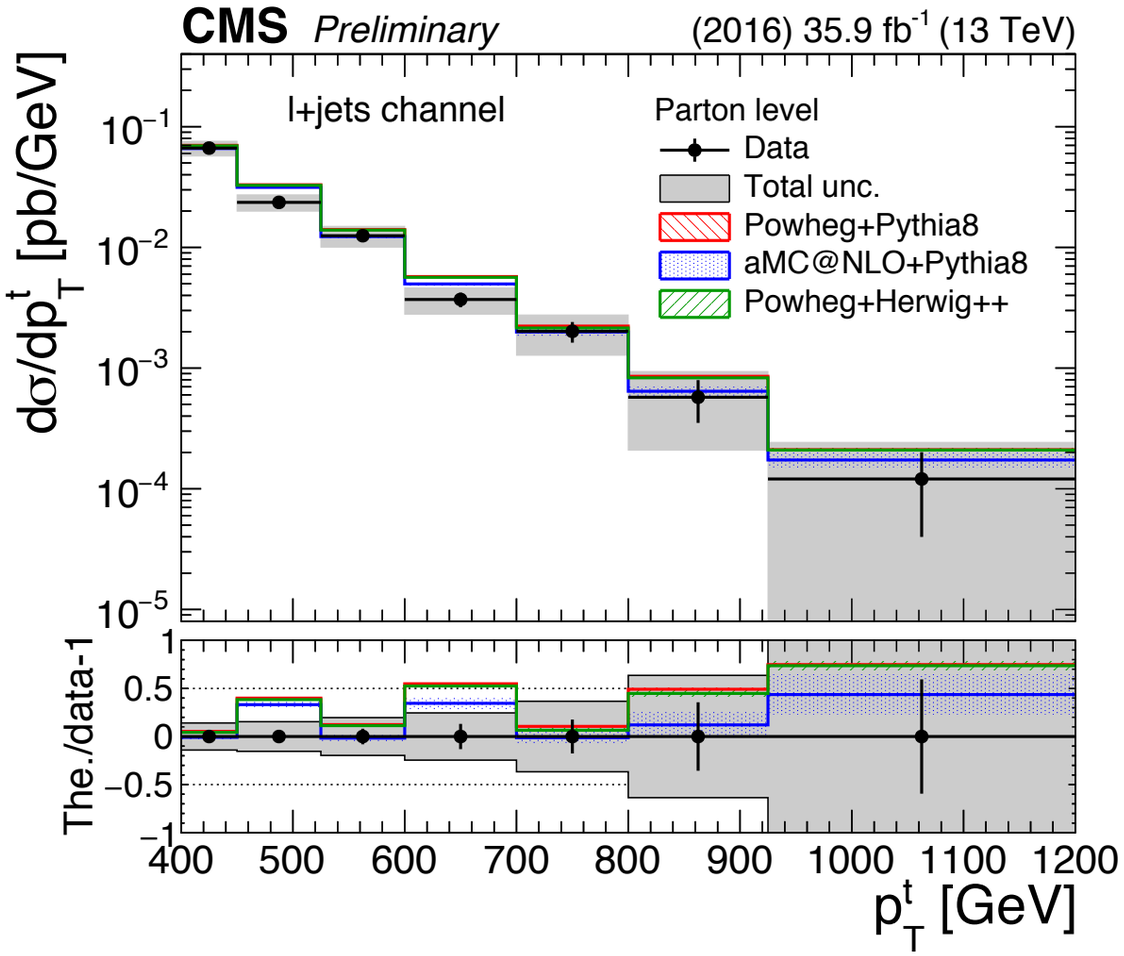
Absolute



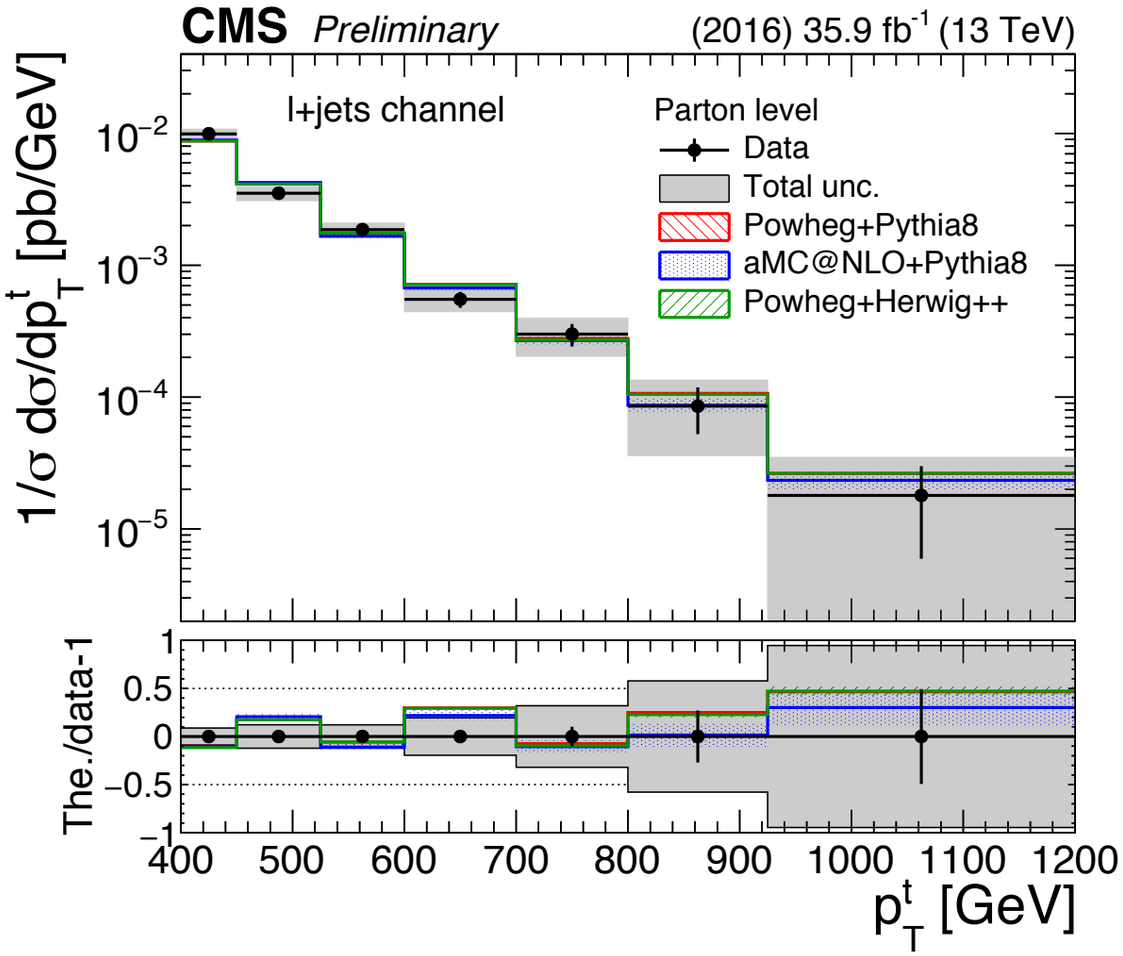
Normalized

Results (parton, top p_T) (l+jets)

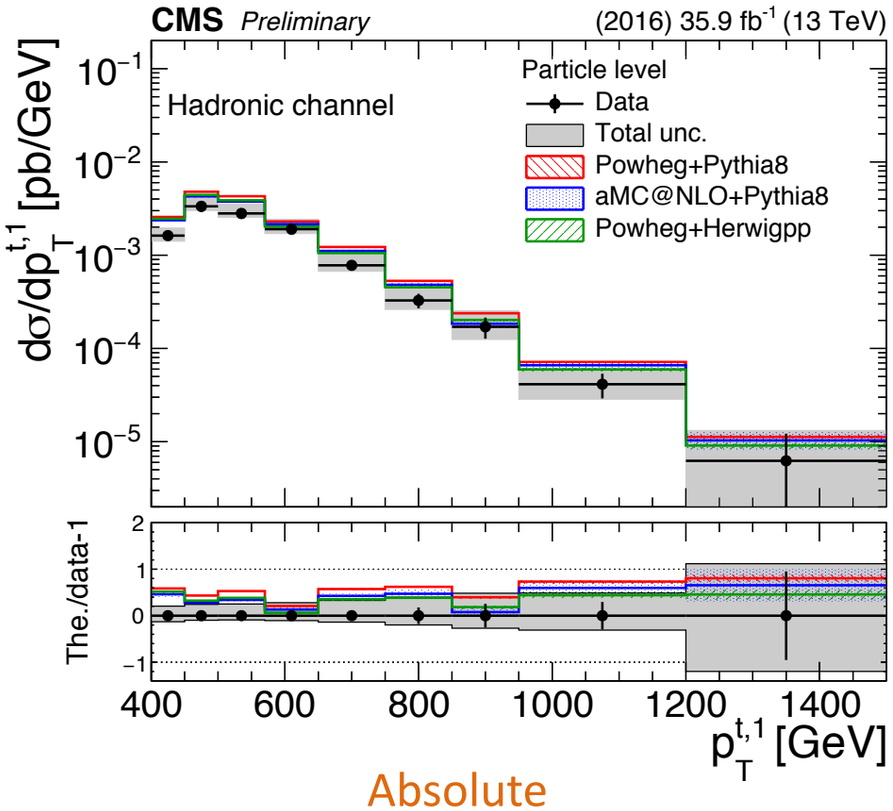
Absolute



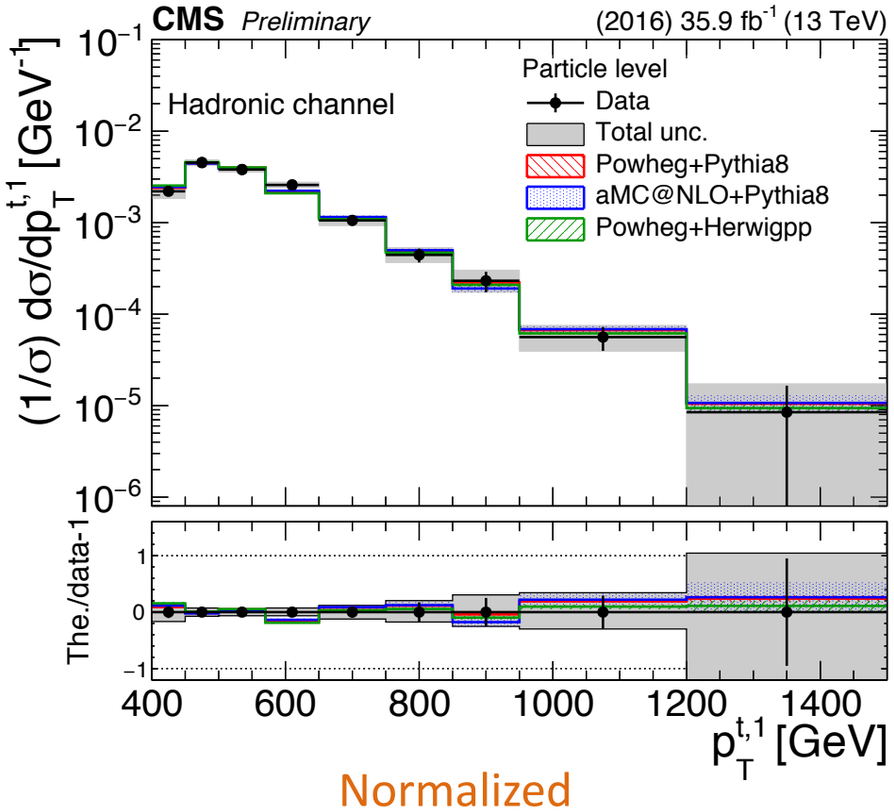
Normalized



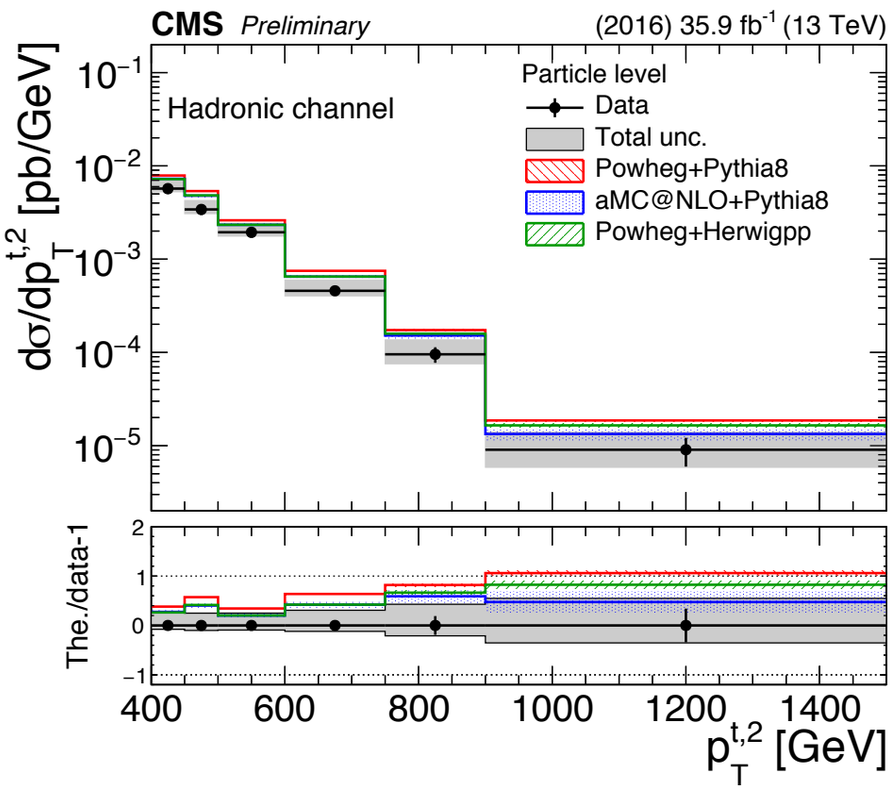
Results (particle, top p_T) (hadronic)

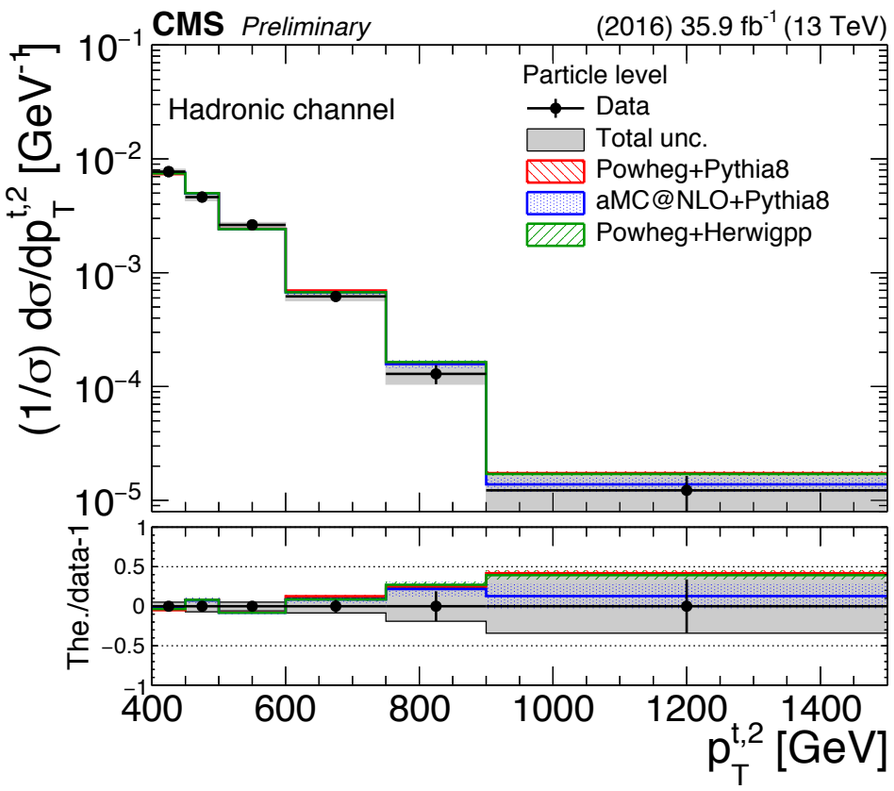


Absolute



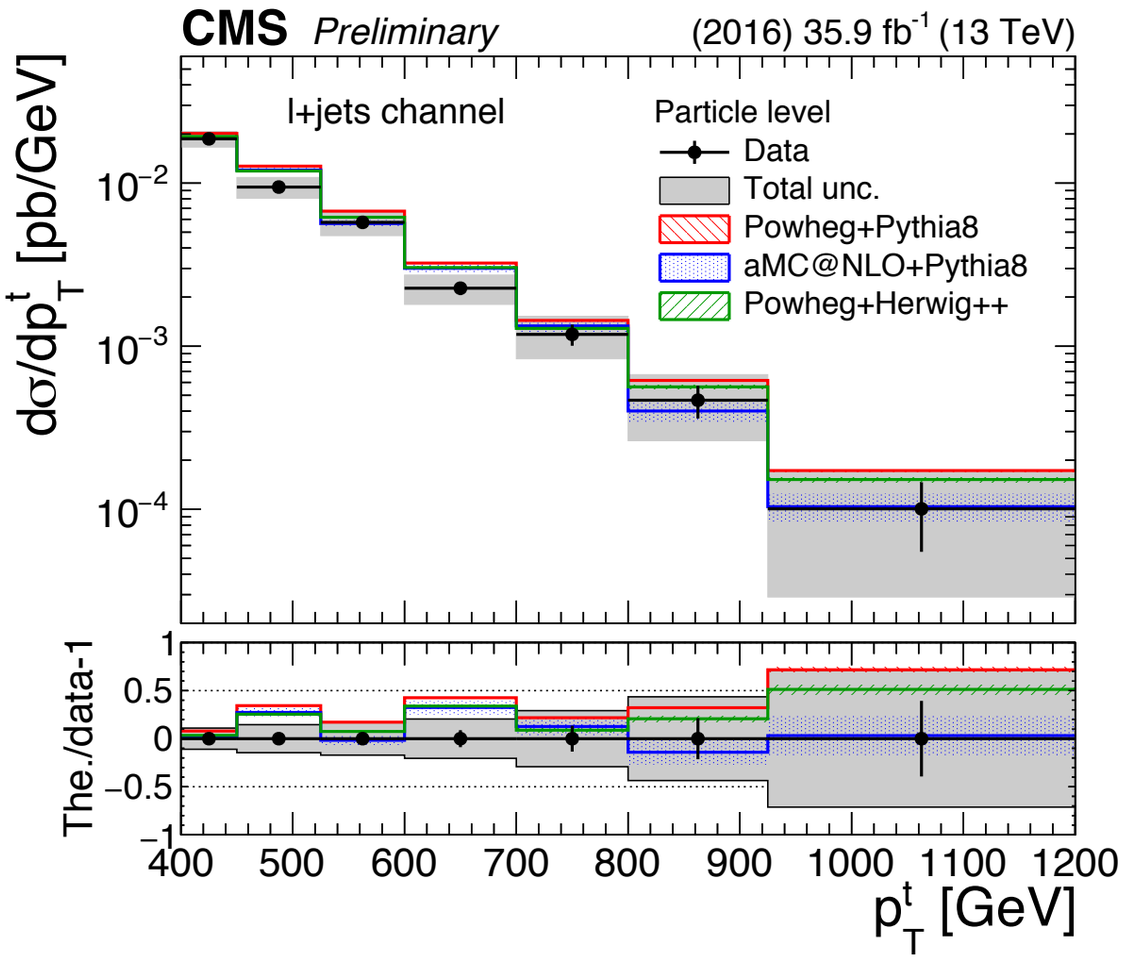
Normalized



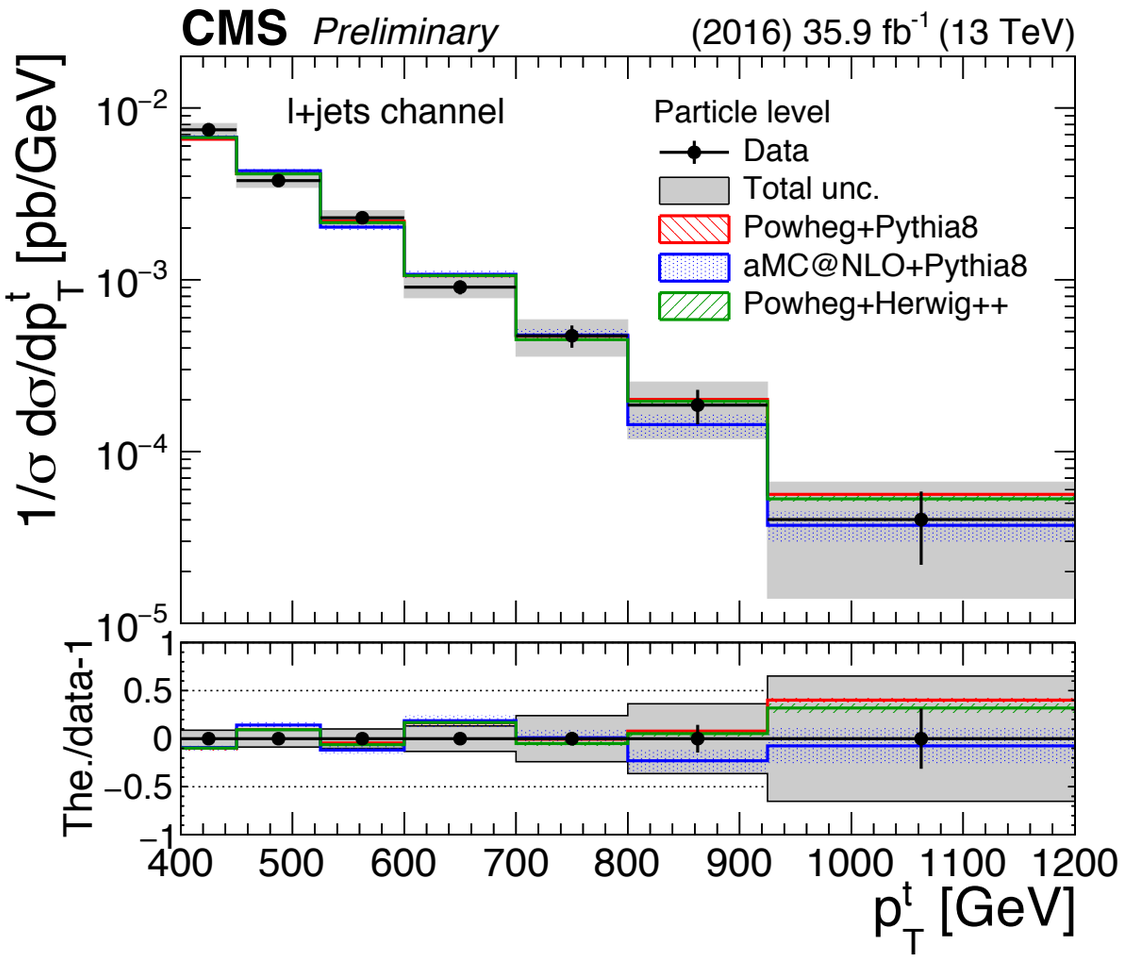


Results (particle, top p_T) (l+jets)

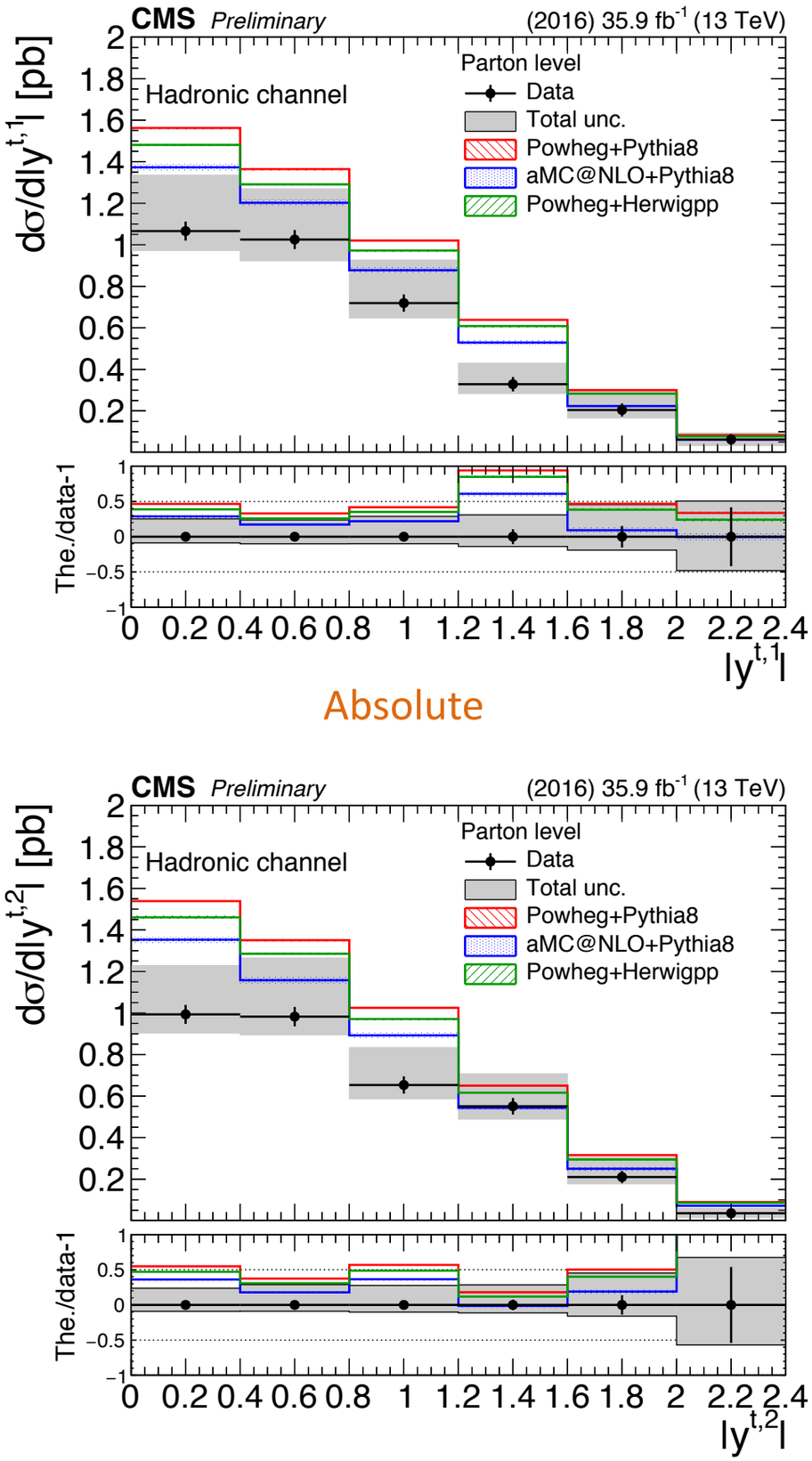
Absolute



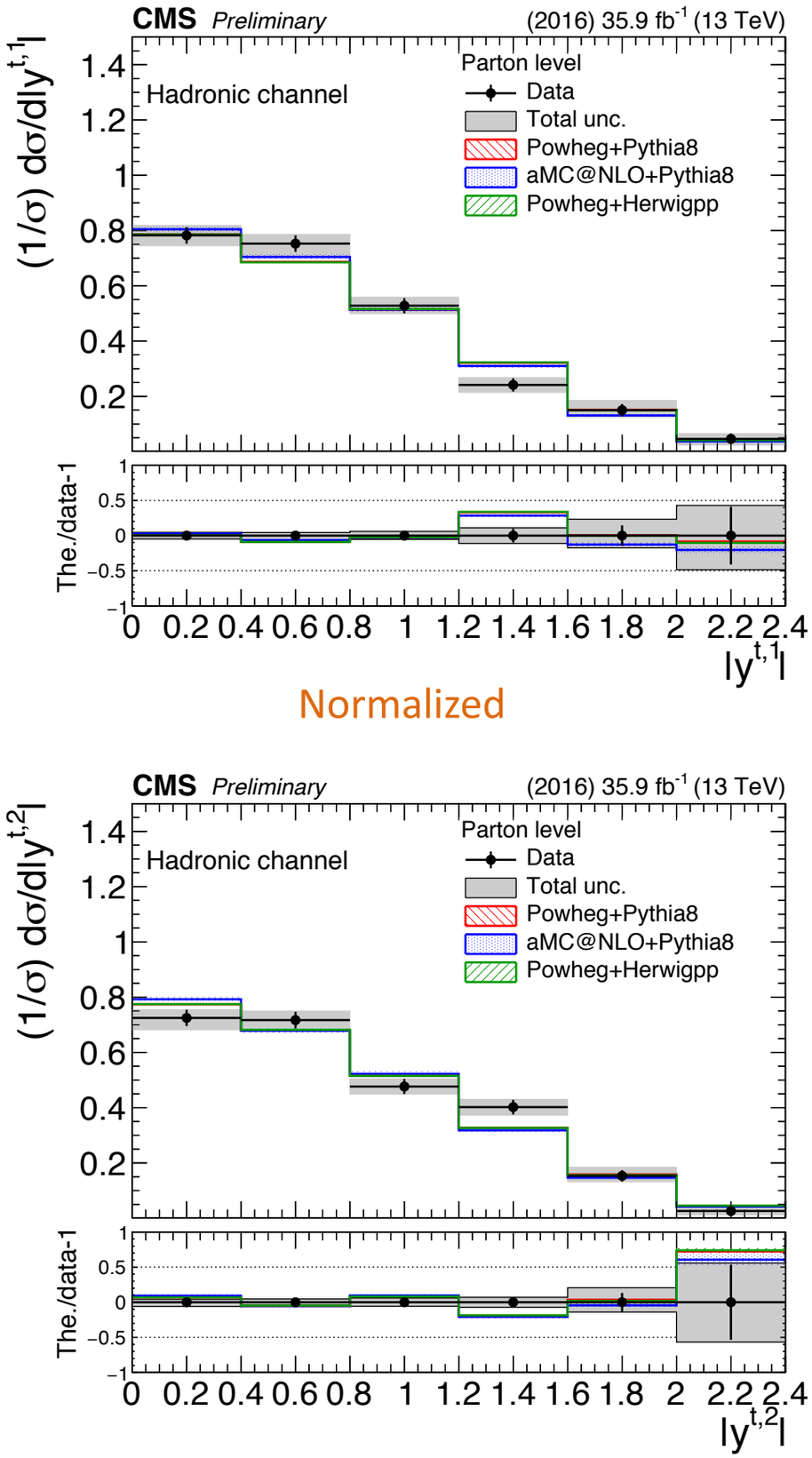
Normalized



Results (parton, top |y|) (hadronic)



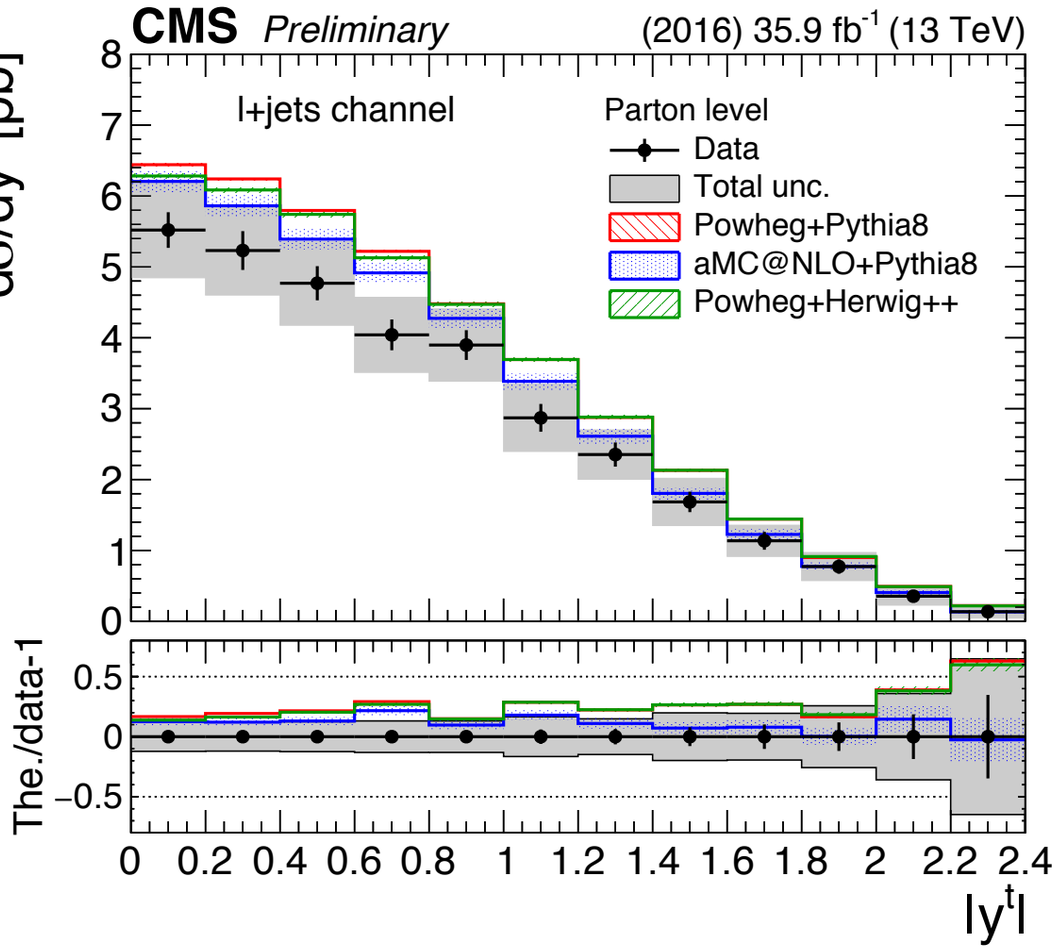
Absolute



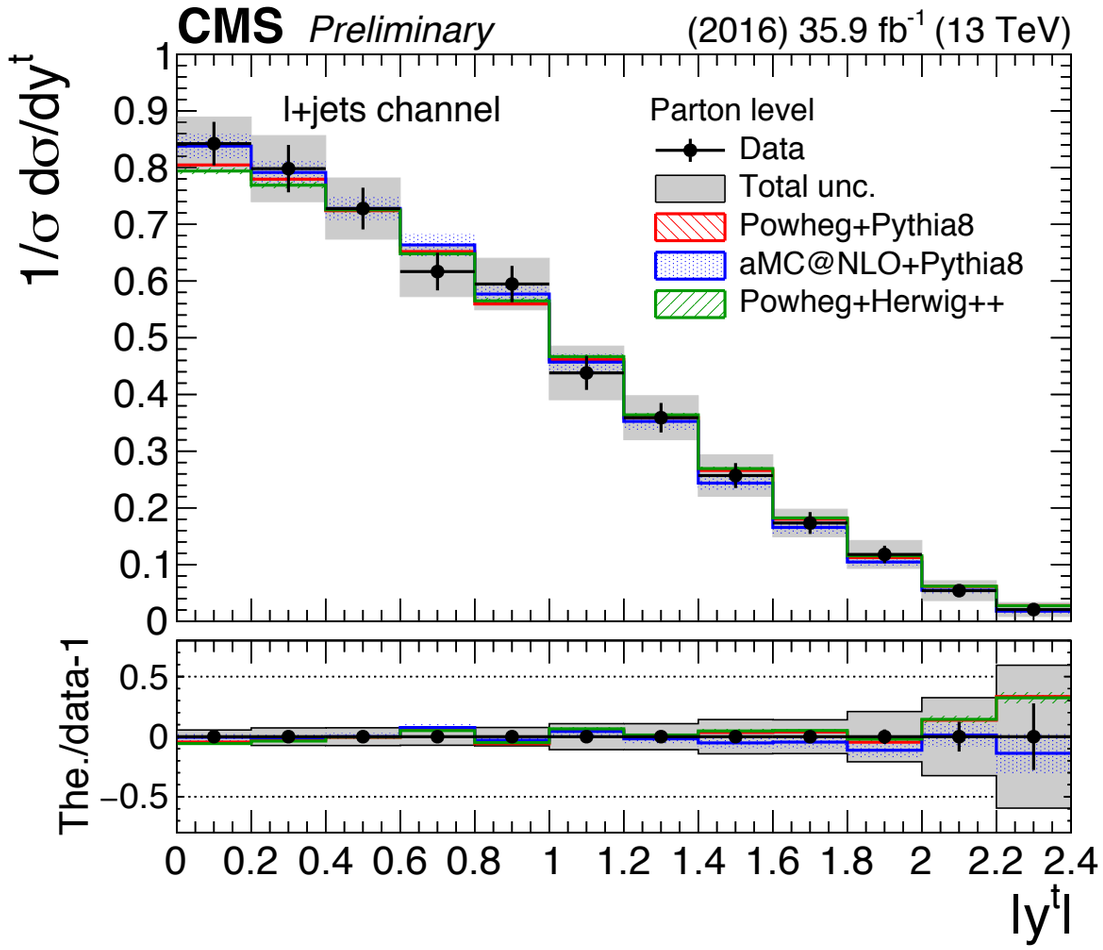
Normalized

Results (parton, top |y|) (l+jets)

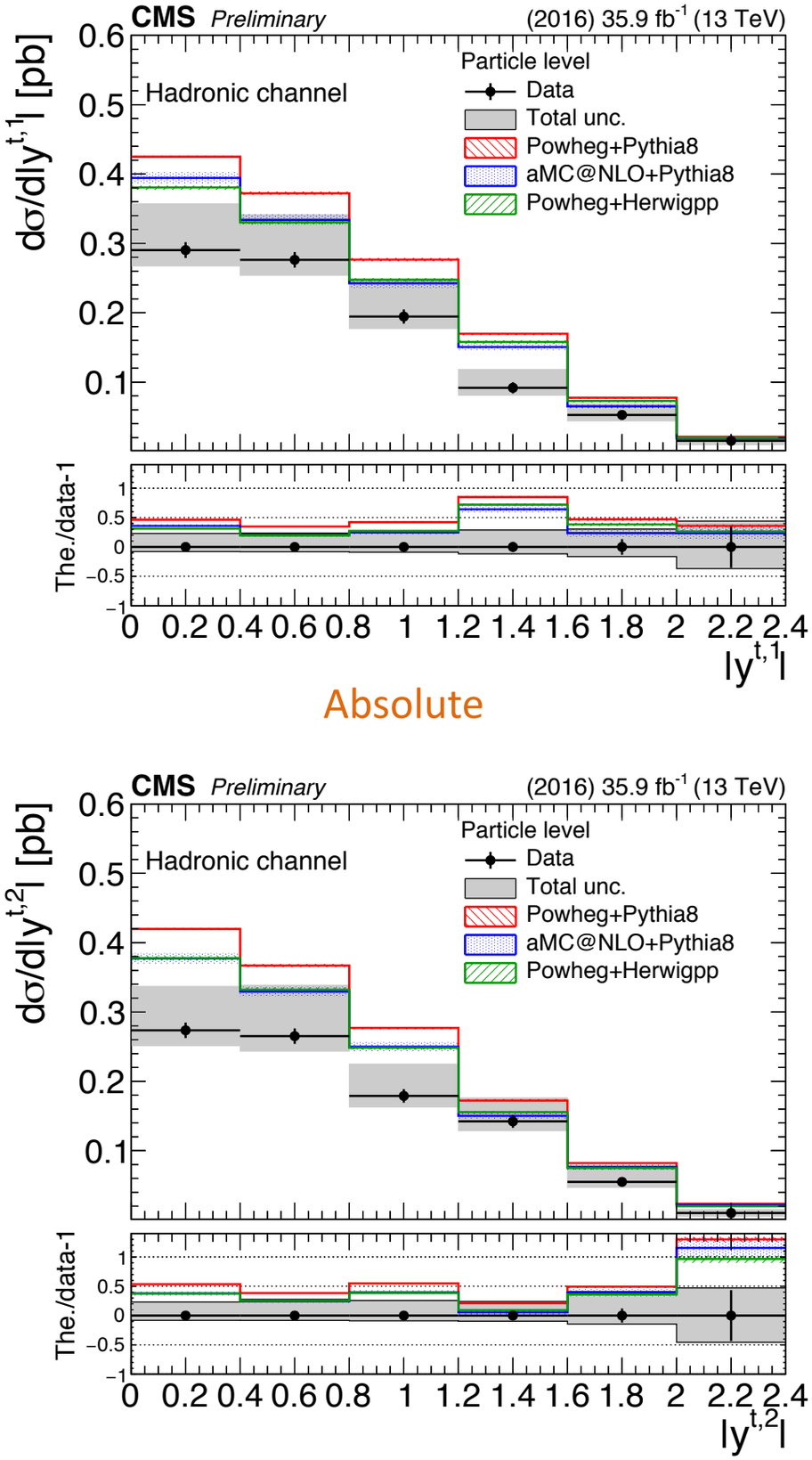
Absolute



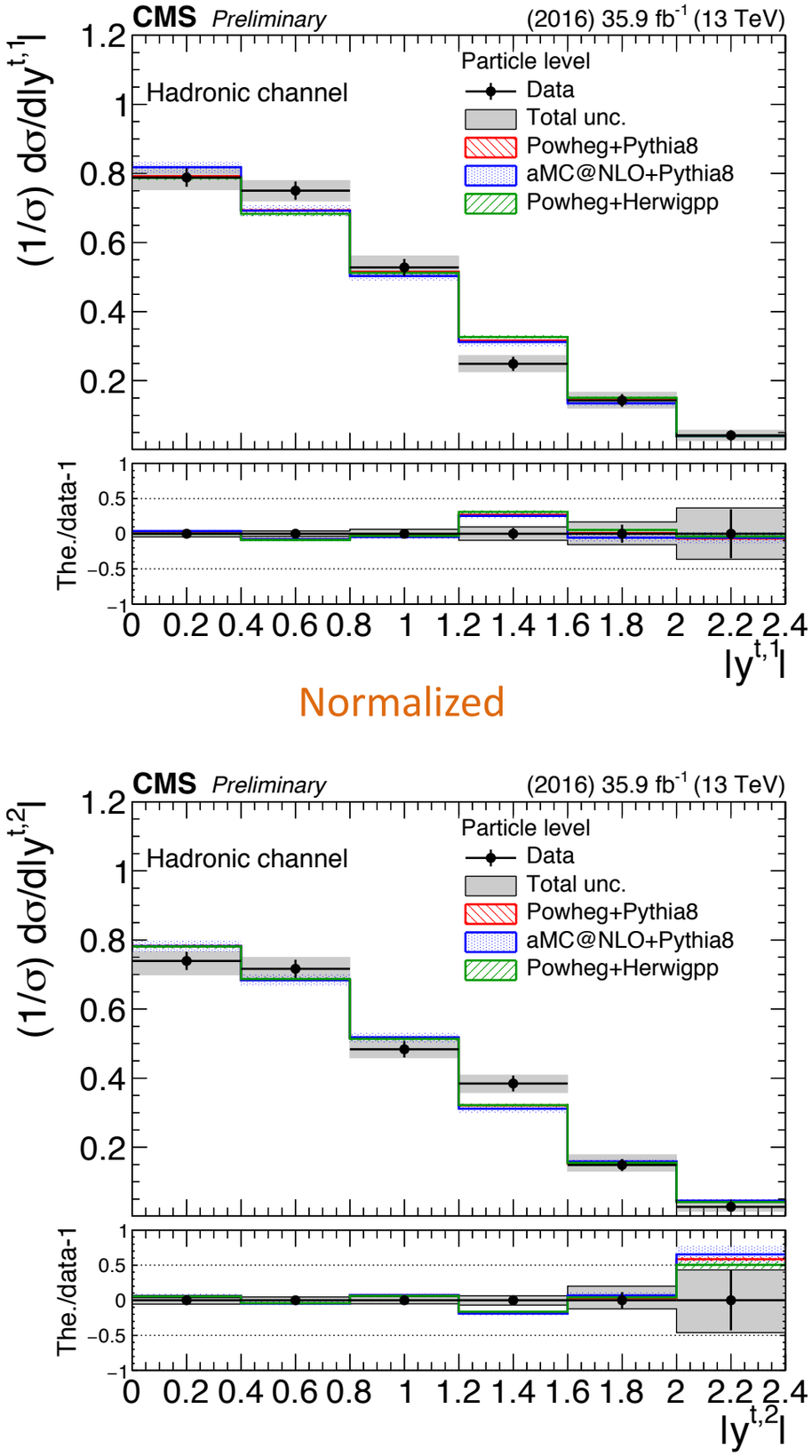
Normalized



Results (particle, top |y|) (hadronic)



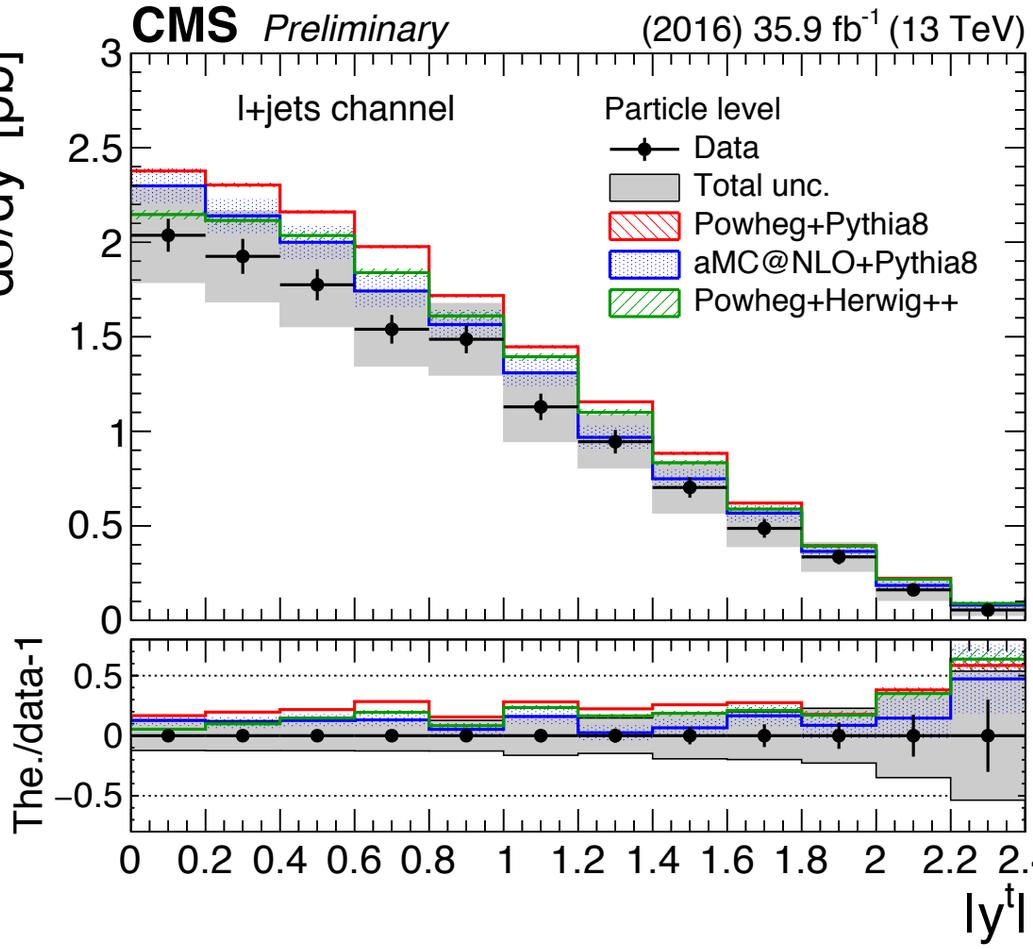
Absolute



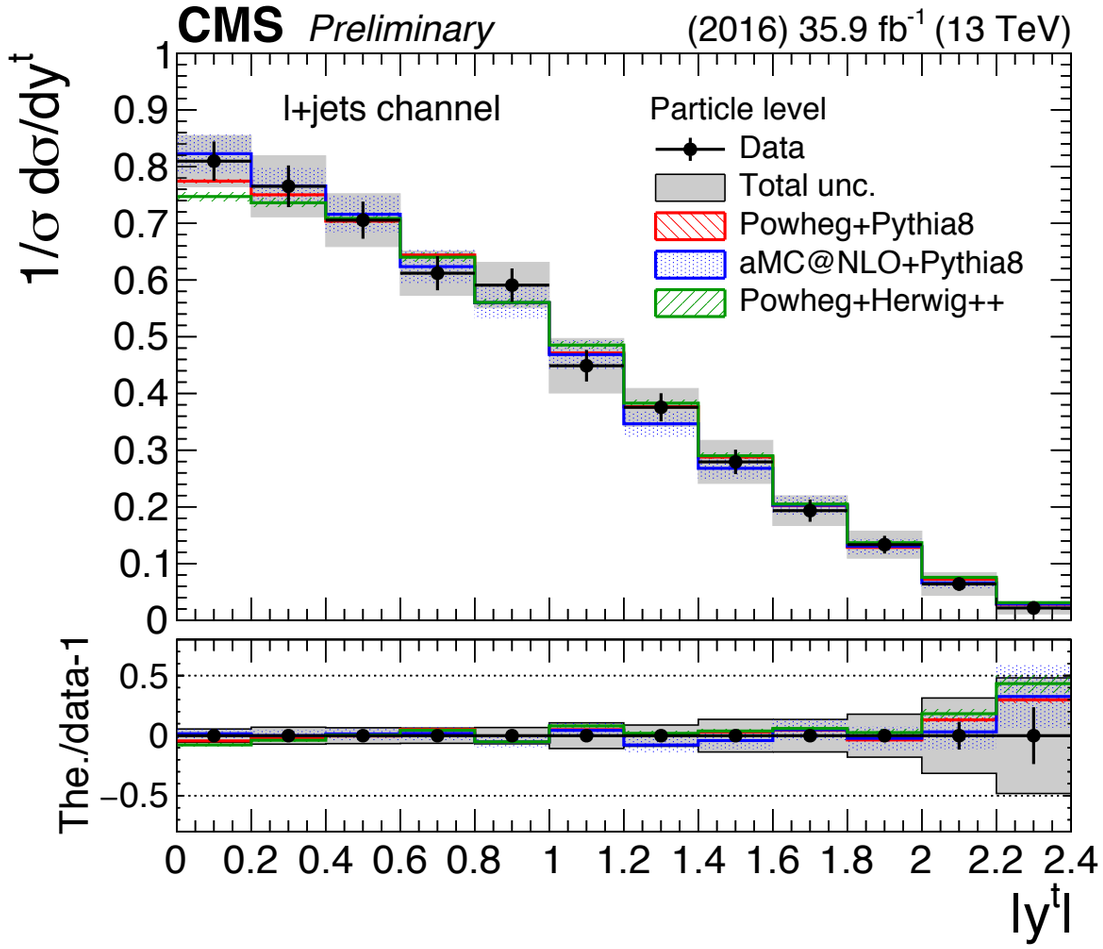
Normalized

Results (particle, top |y|) (l+jets)

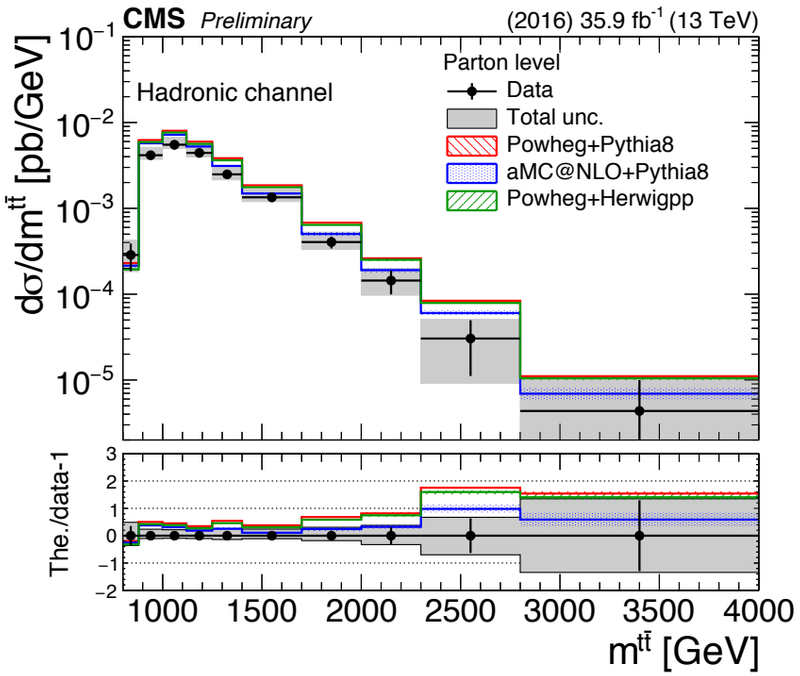
Absolute

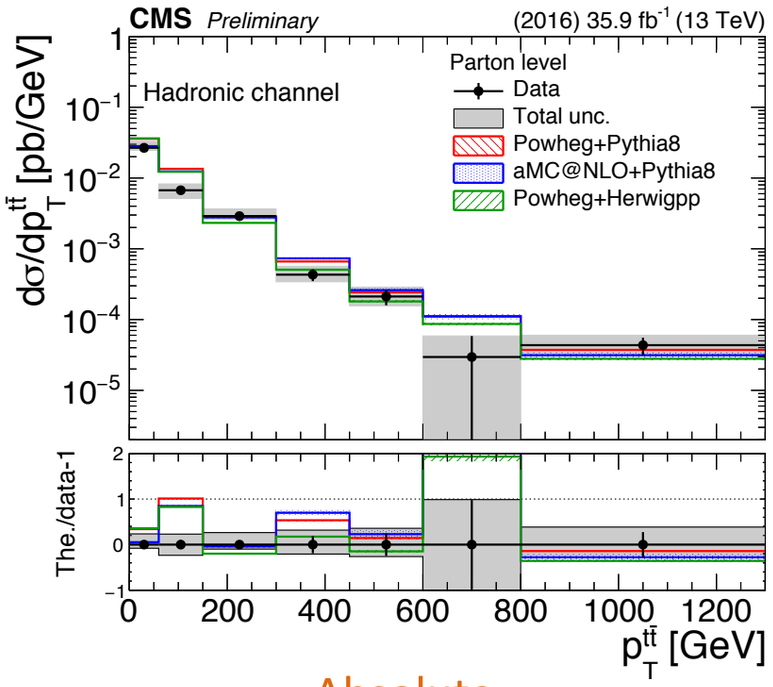


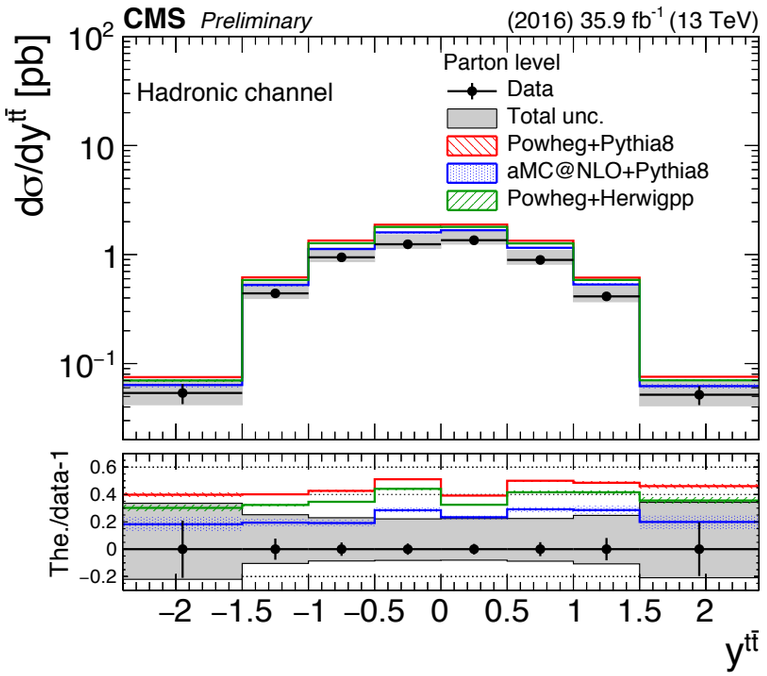
Normalized



Results (parton, ttbar kinematic) (hadronic)

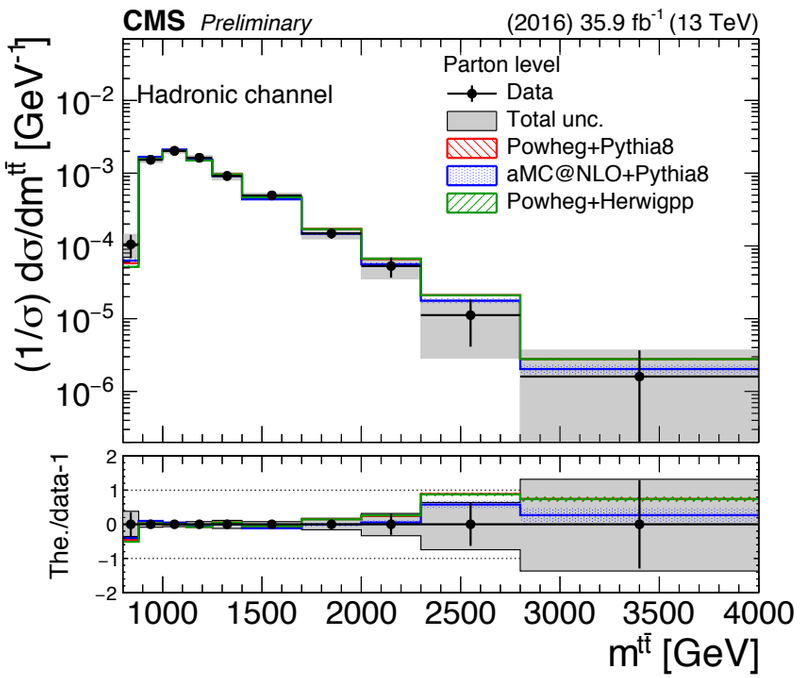


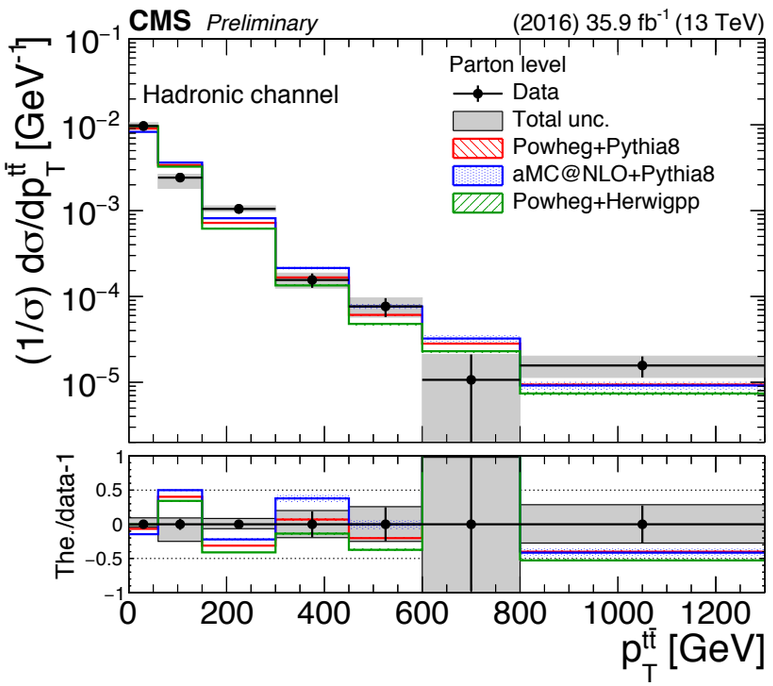


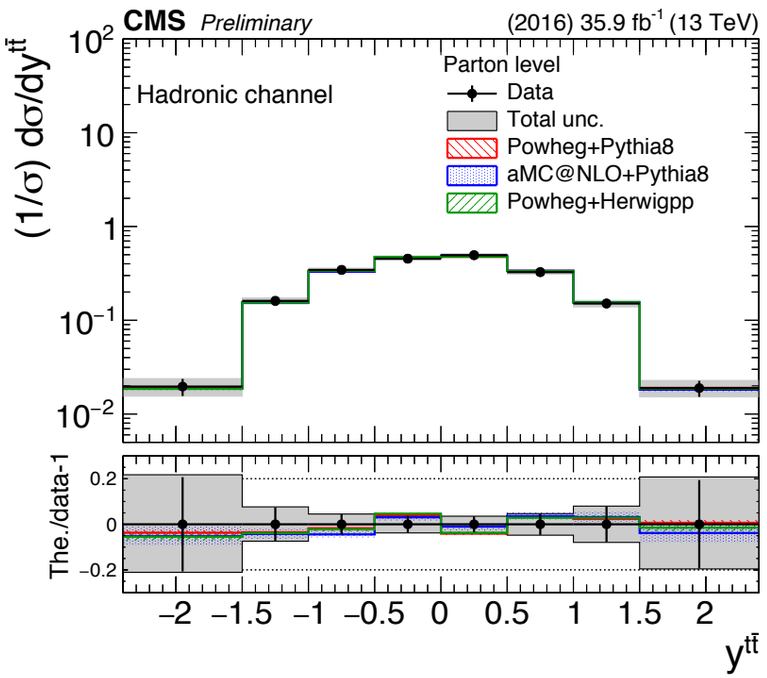


Absolute

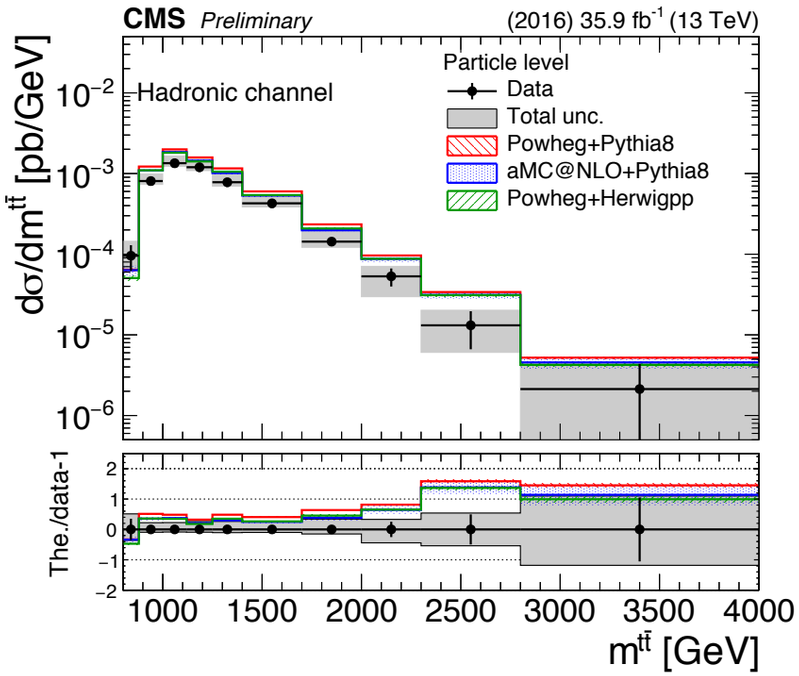
Normalized

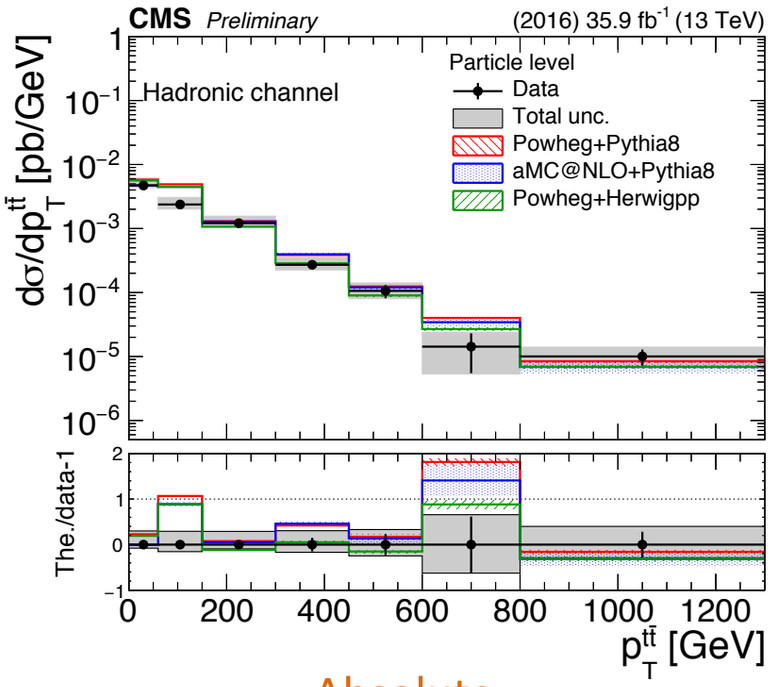


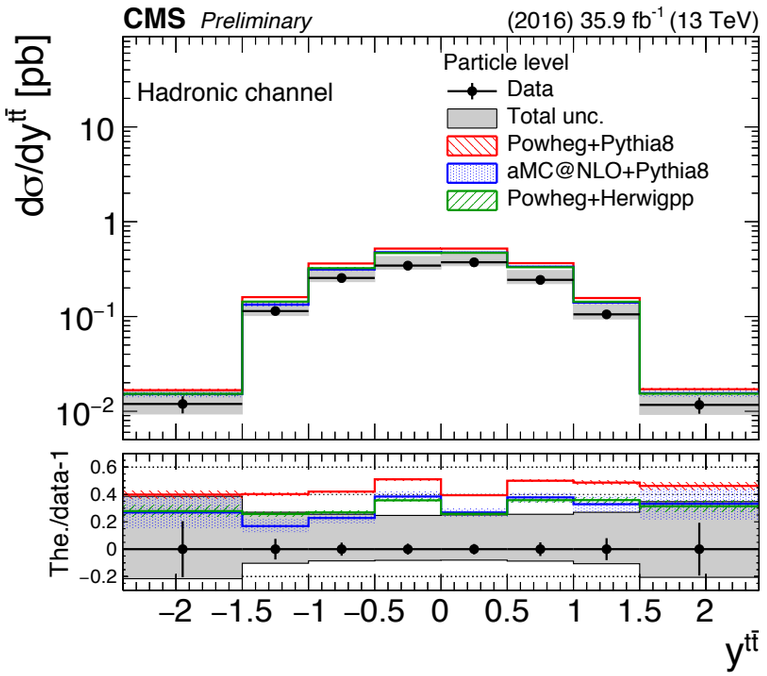




Results (particle, ttbar kinematic) (hadronic)

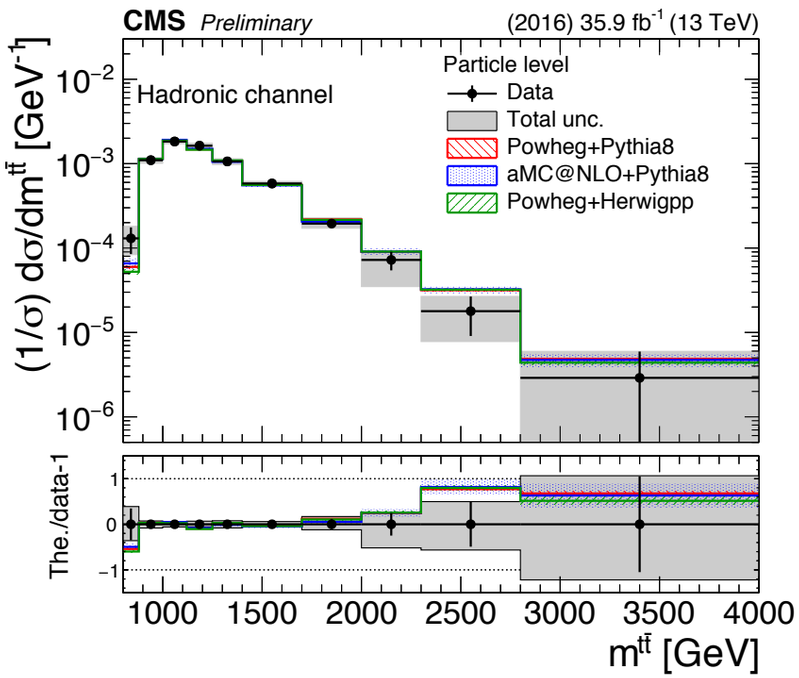


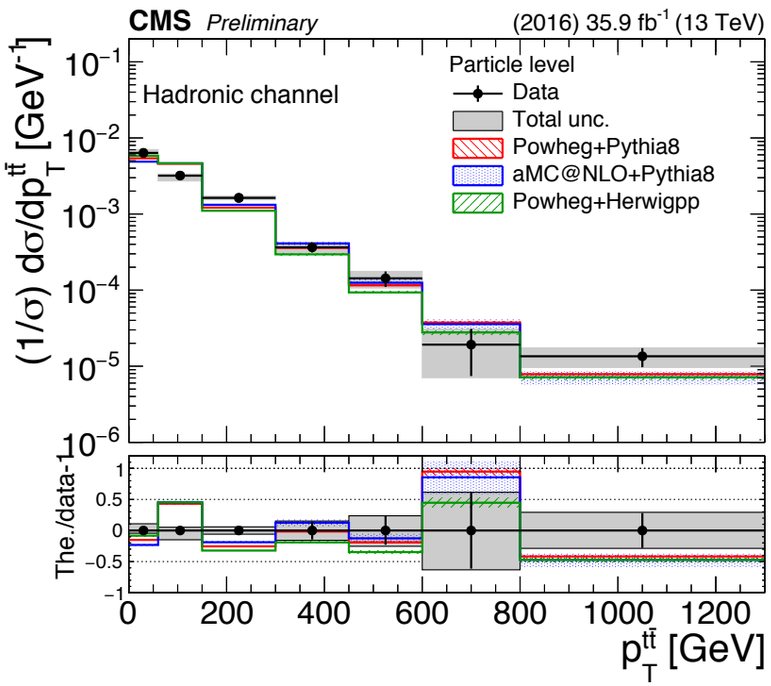


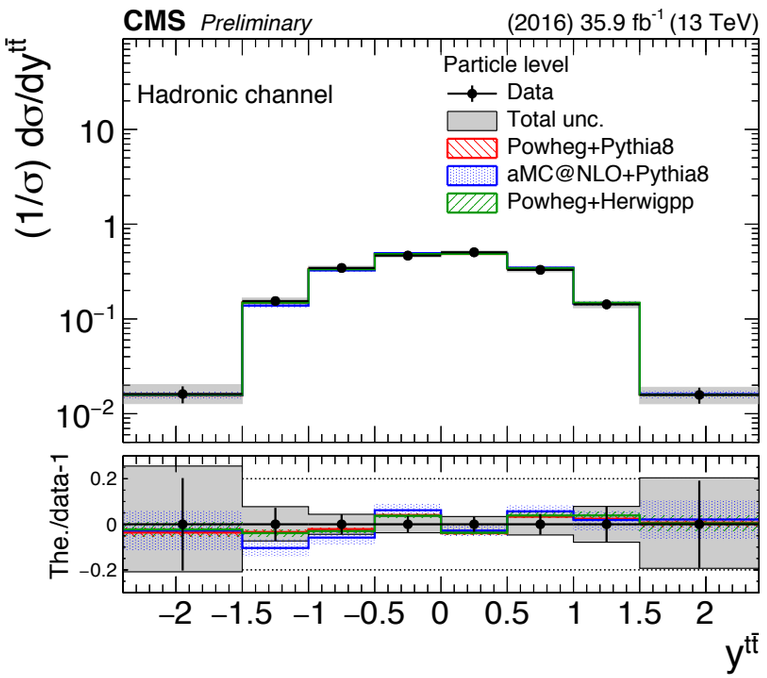


Absolute

Normalized







Summary

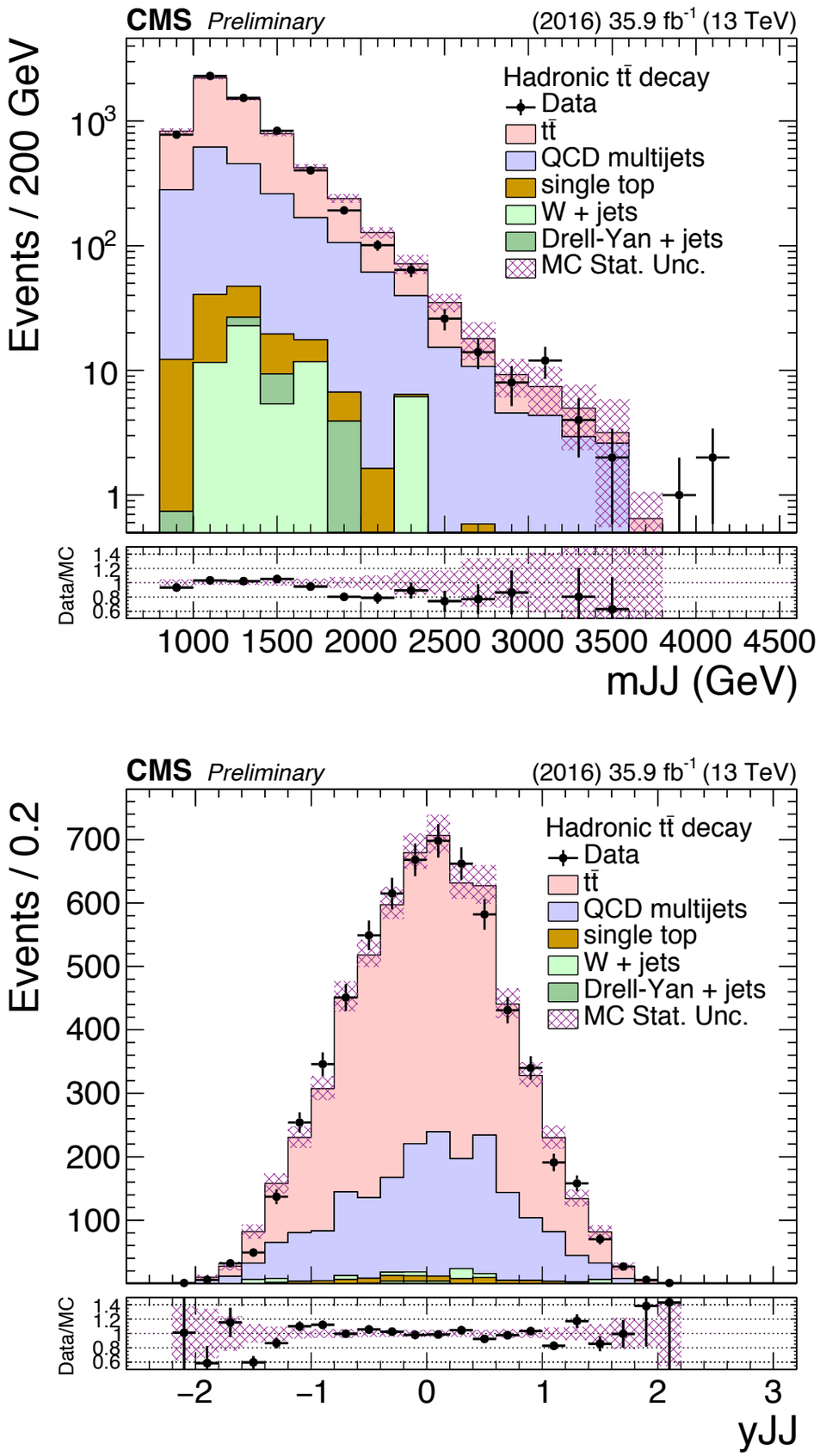
- Differential ttbar cross sections with boosted top quarks are presented in hadronic and l + jets channels using 36fb^{-1} of 13 TeV data

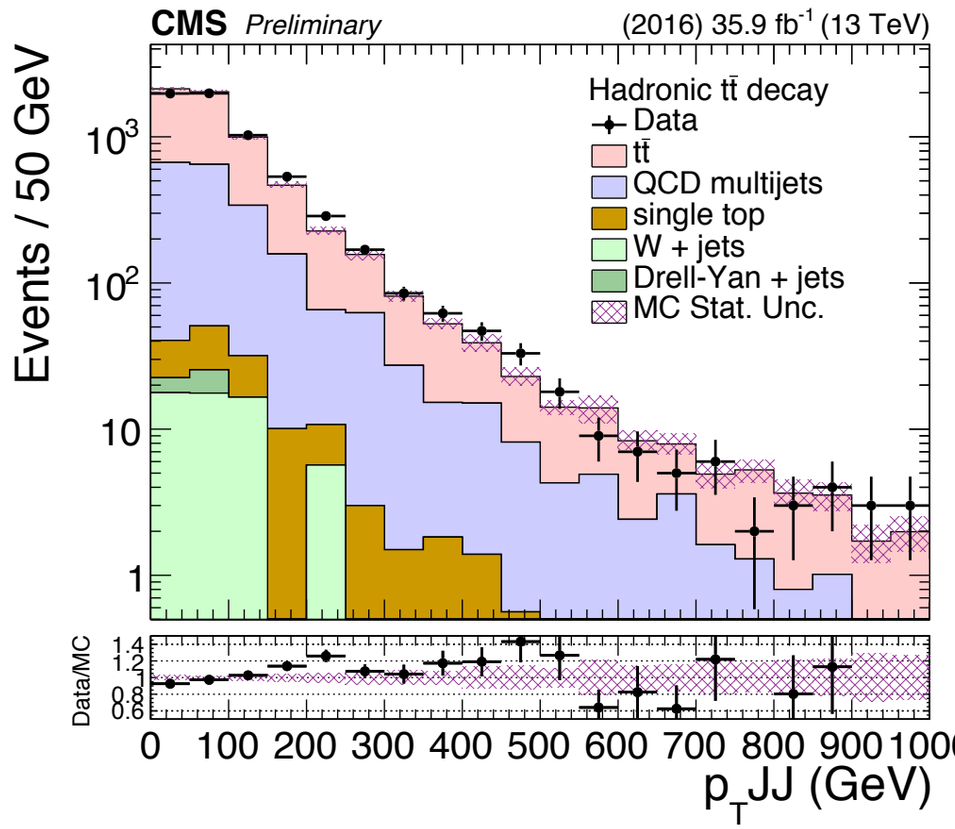
- Observables: leading & subleading top p_T & |y|, $t\bar{t}$ system mass, p_T, y (hadronic channel) and p_T and |y| for the hadronically decaying top quark (l+jets)
  - Parton & particle level
  - Absolute & normalised cross sections

- Results
  - Comparison with MC models: Powheg+Pythia8, Powheg+Herwig++, aMC@NLO+Pythia8
  - Hadronic:
    - Shapes overall compatible with theory: no top p_T slope
    - Overall shift of the order of 35% in the total cross section (Powheg+Pythia8 shows the largest discrepancy)
  - L + jets:
    - Differential distributions generally well described
    - All models overpredict the absolute cross section (~20%)

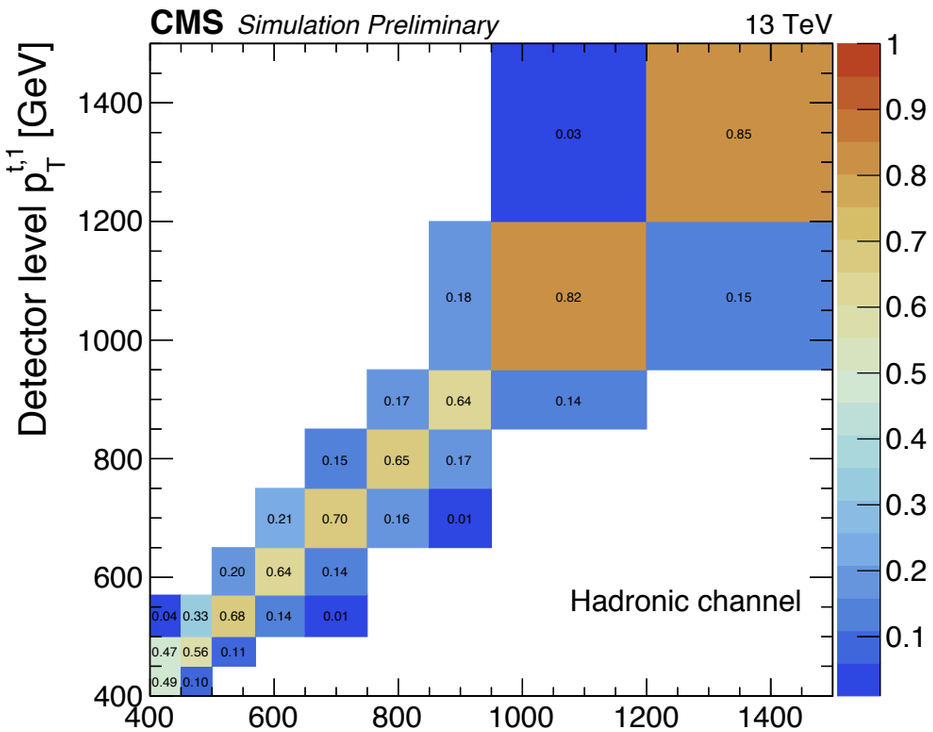
Backup

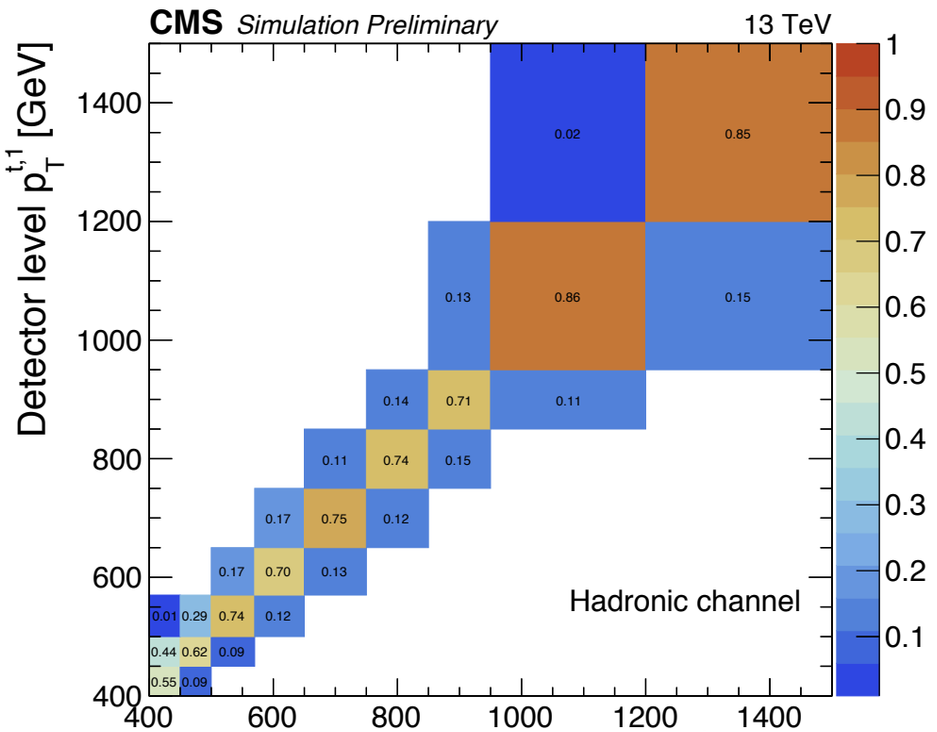
Top kinematic distributions (hadronic)

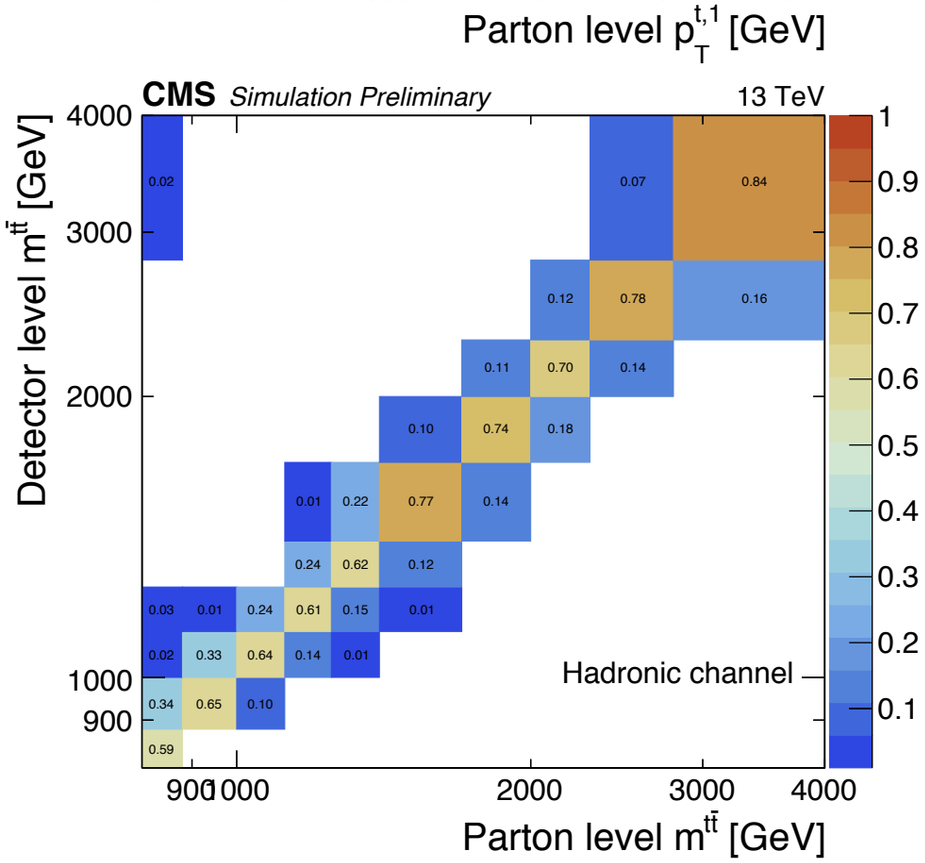


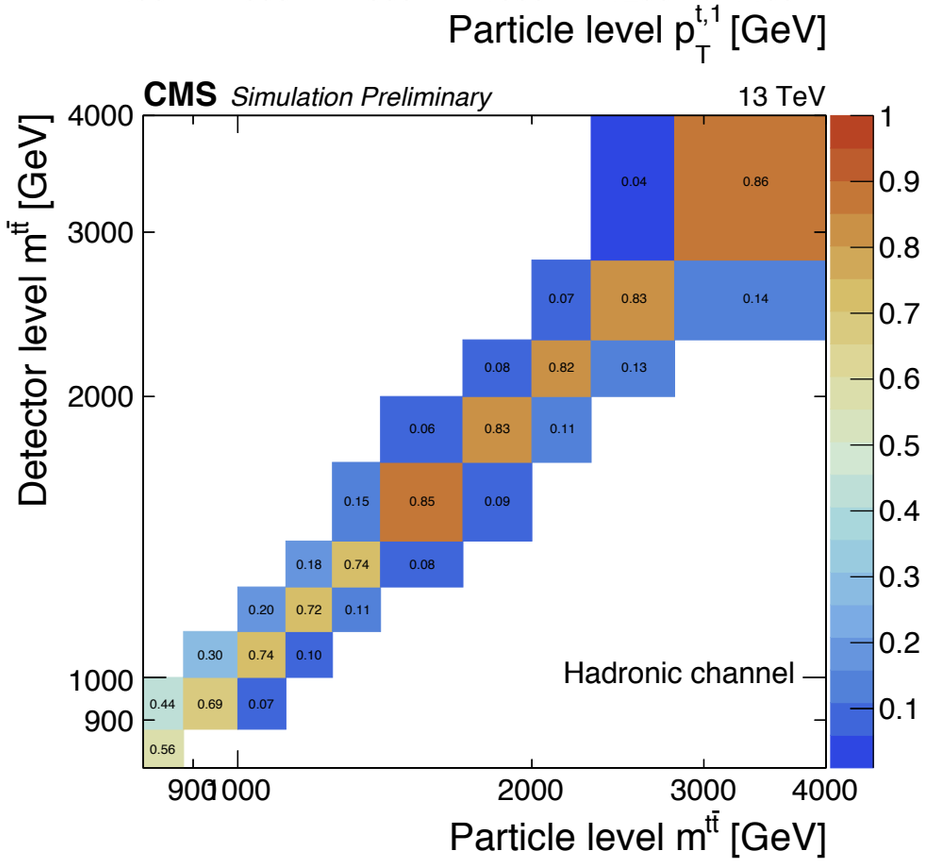


Migration matrices (hadronic)

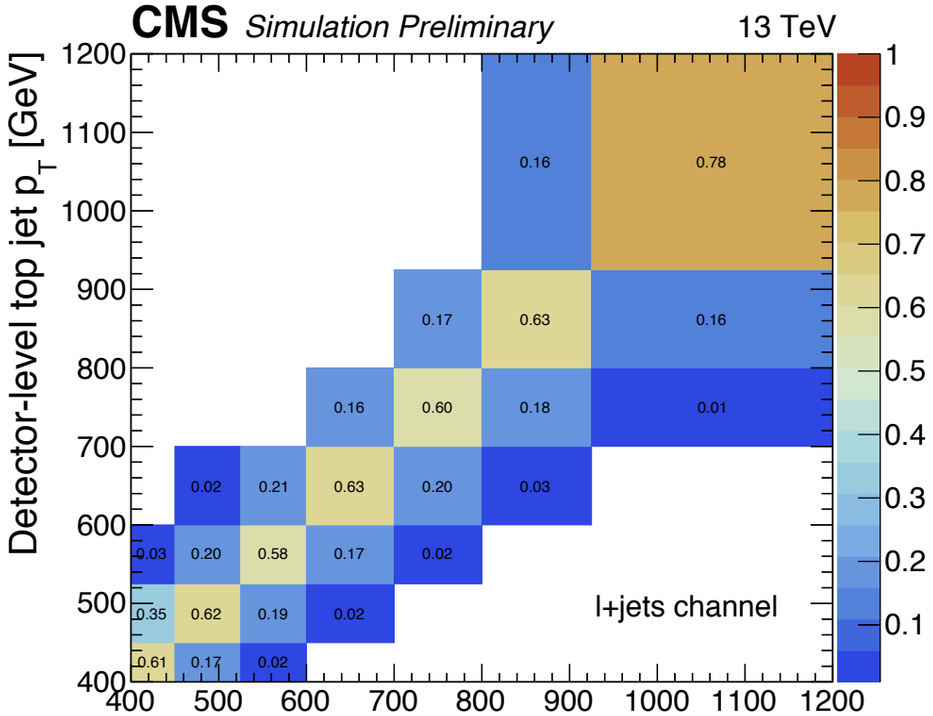


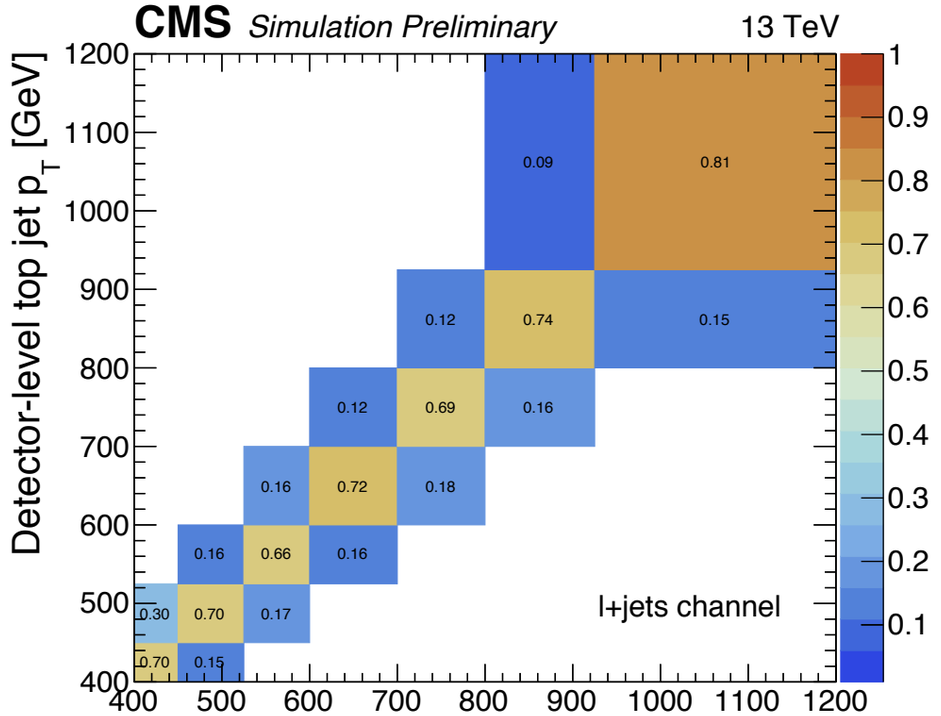


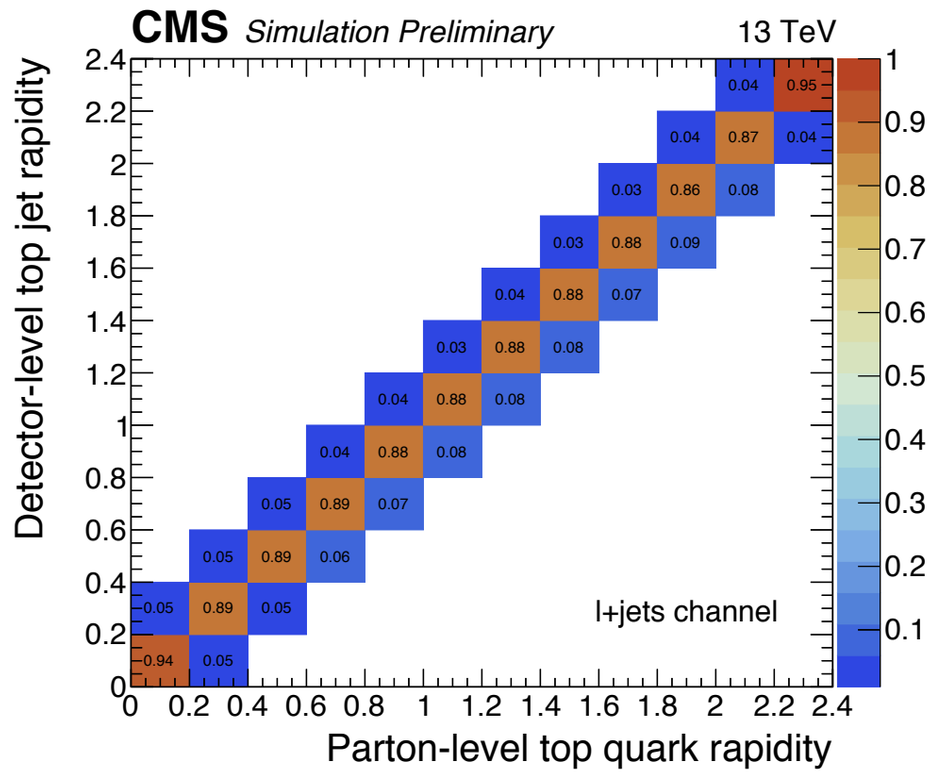


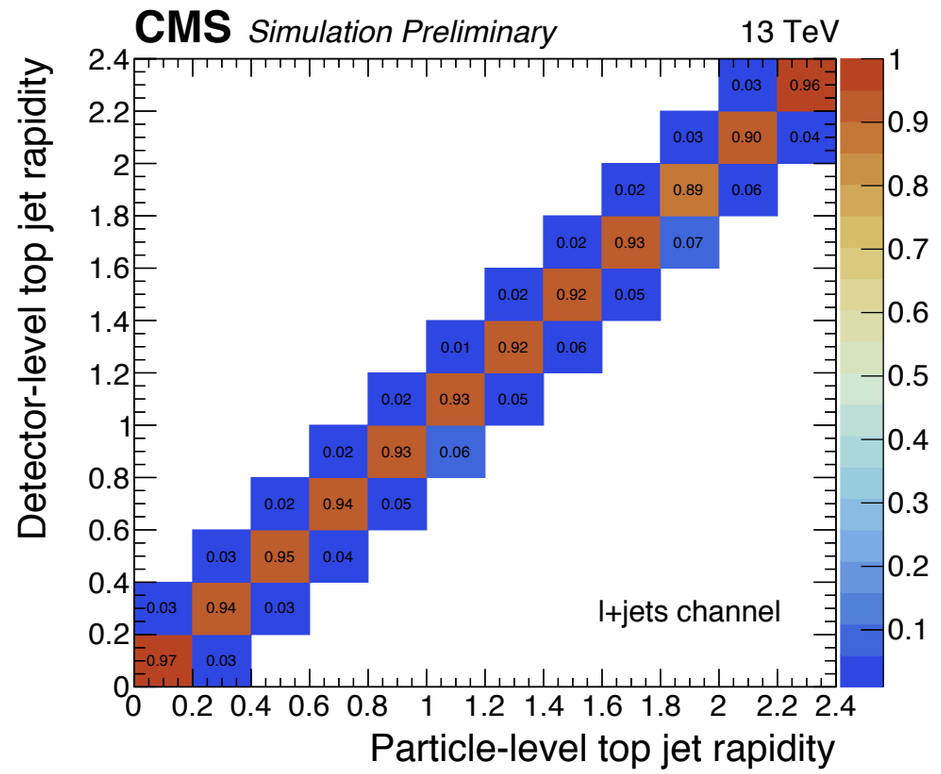


Migration matrices (l + jets)

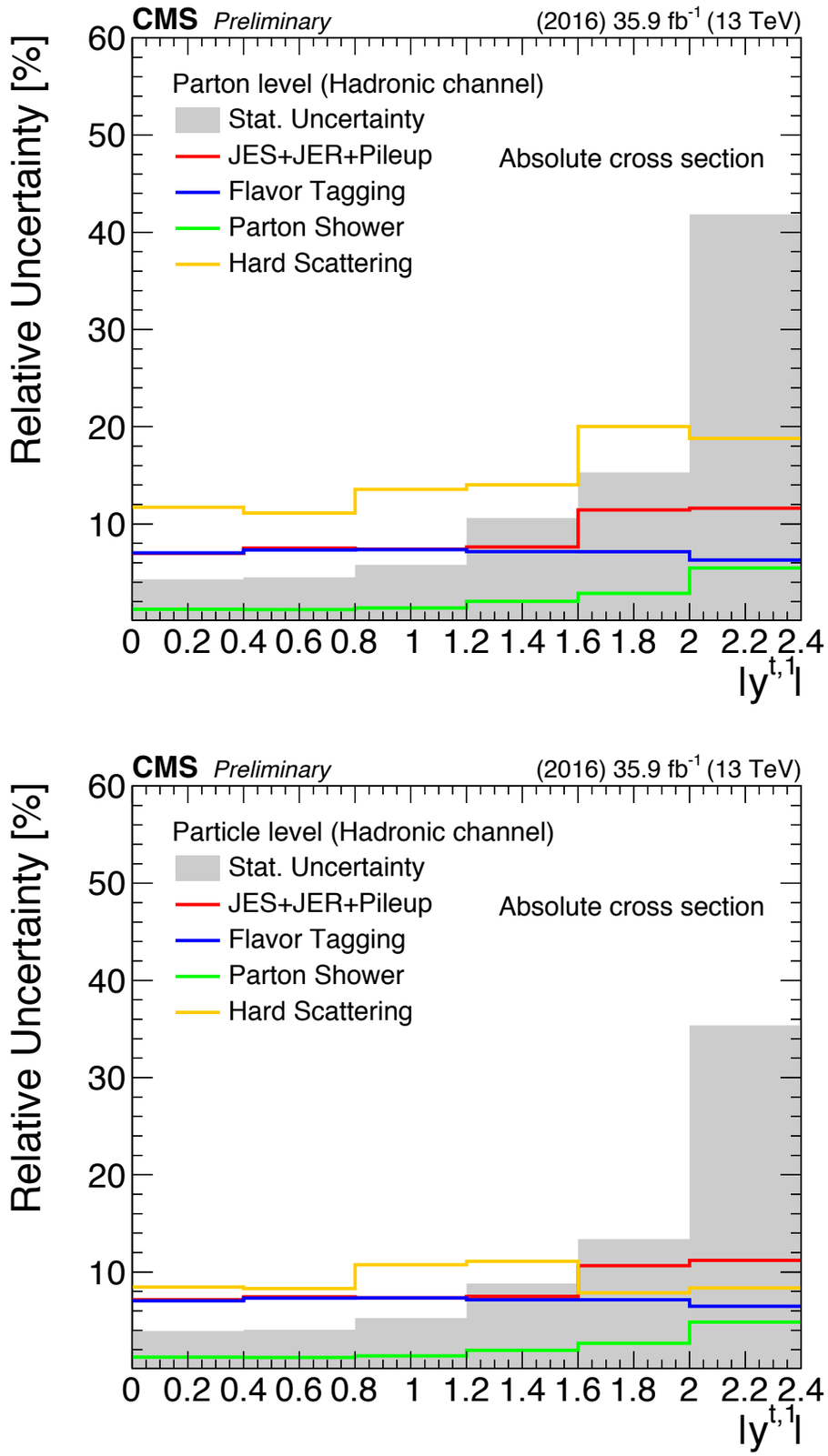


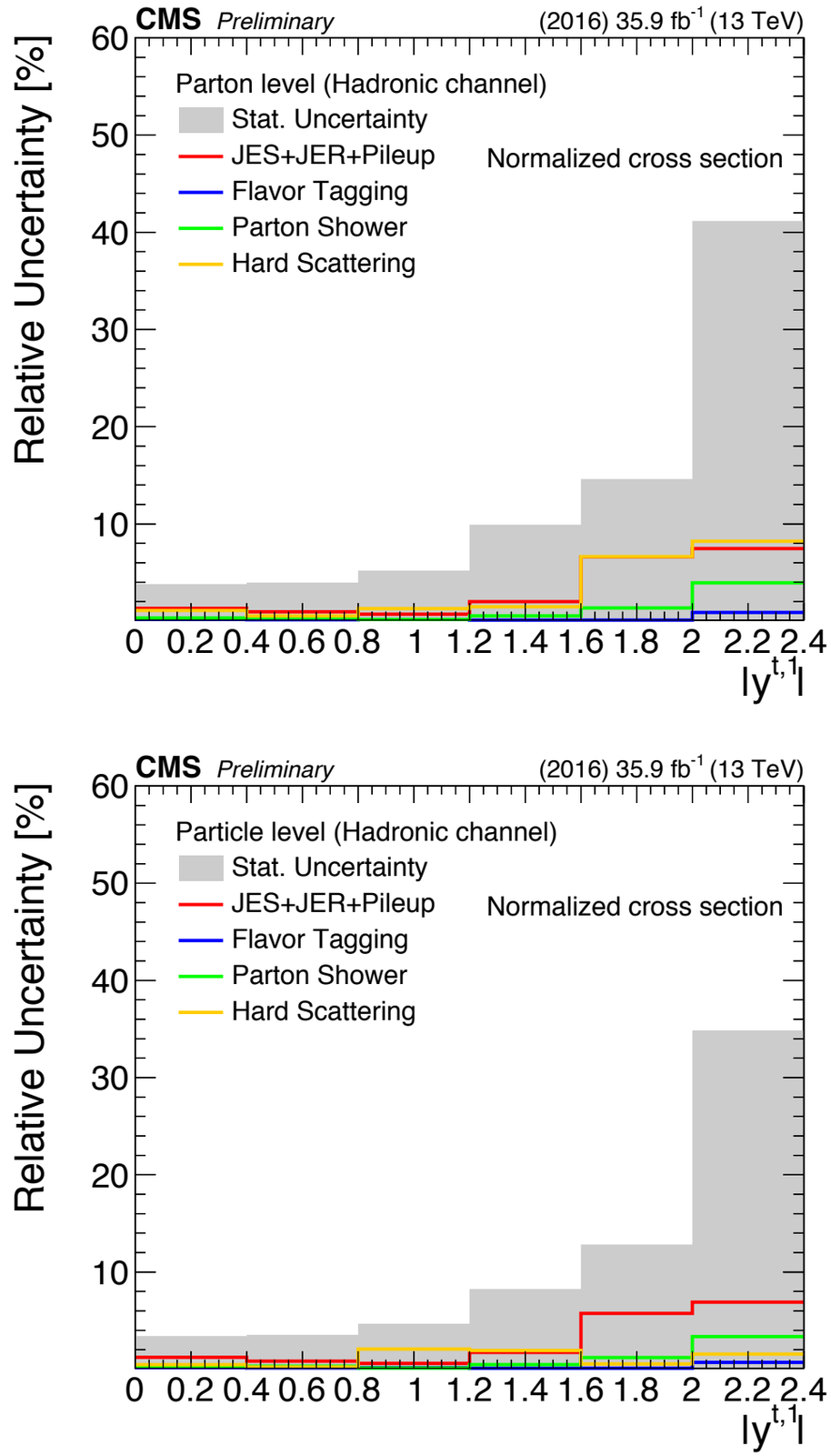




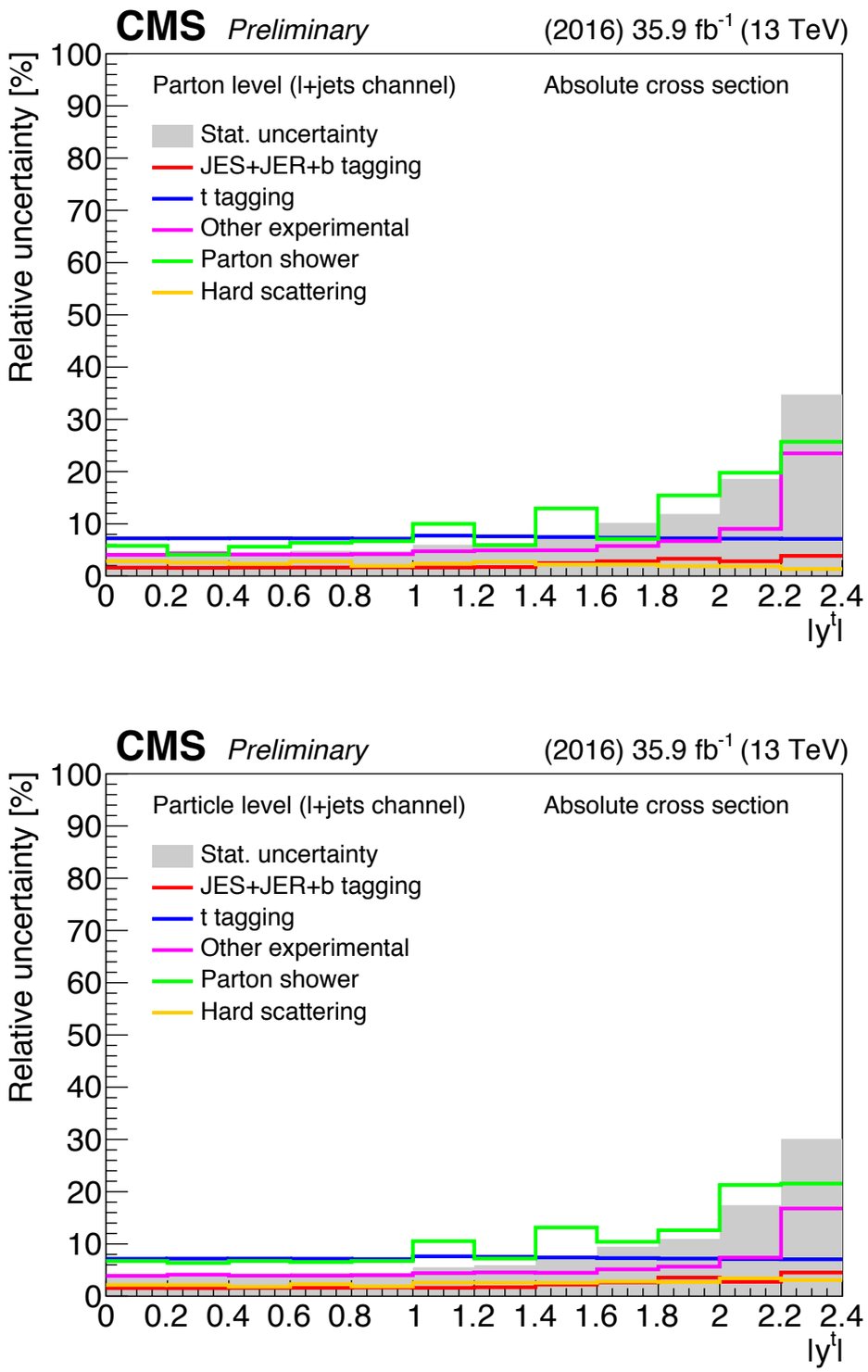


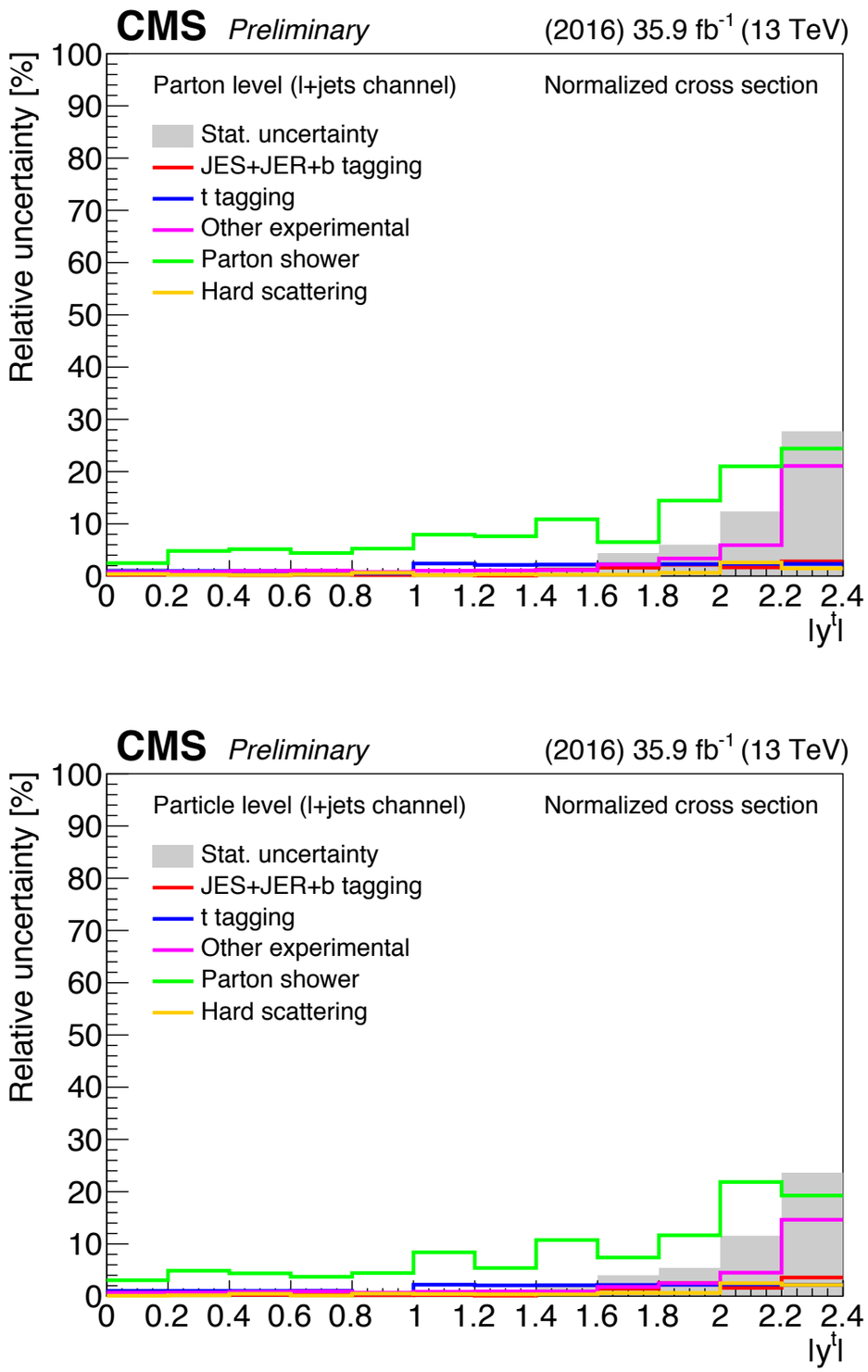
Uncertainties vs leading top |y| (hadronic)





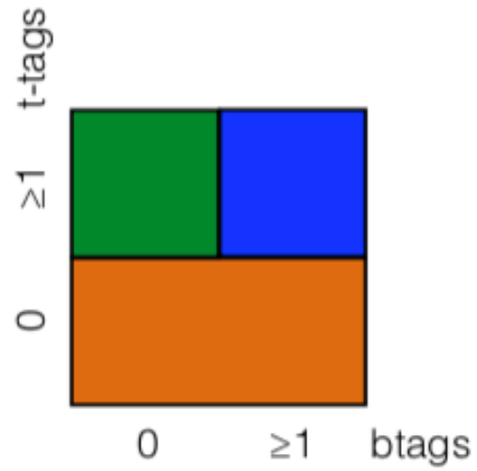
Uncertainties vs hadronically decaying top |y| (l+jets)





Selection & Kinematic Regions (l+jets)

- Lepton selection:
  - 1 μ(e) passing Medium (Tight) ID
  - p_T > 50 GeV, |η|<2.1, miniIso* < 0.1
  - Veto additional leptons
  - In e channel, require $|\Delta\varphi(e/j, E_T^{Miss}) - 1.5| < 1.5 * \frac{E_T^{Miss}}{100GeV}$ for e or leading jet
- ≥ 1b jet candidate:
  - AK4 jet with p_T > 50 GeV and |η| < 2.4
  - 0.3 < ΔR(l, jet) < π/2
- ≥ 1t jet candidate:
  - AK8 jet with p_T > 400 GeV and |η| < 2.4
  - ΔR(l, jet) > π/2
- E_T^{Miss} > 35 (50) GeV for μ (e) channel



Correction Factors (l+jets)

- PU reweighting
- Lepton Efficiencies:
  - Trigger — custom T&P in dileptonic boosted tt sample
  - ID — μ from POG; e from SUS PAG
  - Isolation — μ assessed to be ~ 1; e from SUS PAG
- Jet Energy Corrections
  - Summer16_23Sep2016v3
- Jet Energy Resolution
- b tagging SF
- t tagging SF
  - Correlated with Cross Section, extracted from Simultaneous Fit

Standard POG-approved values used

Fit results and PostFit event counts (l + jets)

Process	Number of events (μ+jets channel) 0t	1t0b	1t1b
$t\bar{t}$	16772 ± 1438	4245 ± 174	3905 ± 80
Single t	3286 ± 587	282 ± 68	153 ± 34
W+jets	23104 ± 2871	2368 ± 318	105 ± 20
Z+jets	2582 ± 680	234 ± 69	19 ± 10
Diboson	557 ± 155	31 ± 10	2 ± 1
QCD multijets	2833 ± 1207	159 ± 76	43 ± 22
Total	49135 ± 3549	7320 ± 383	4228 ± 93
Data	49137	7348	4187

Process	Number of events (e+jets channel) 0t	1t0b	1t1b
$t\bar{t}$	10707 ± 938	2835 ± 116	2670 ± 66
Single t	2267 ± 403	191 ± 47	107 ± 24
W+jets	13945 ± 1742	1445 ± 194	62 ± 12
Z+jets	1068 ± 295	118 ± 37	17 ± 15
Diboson	373 ± 105	22 ± 7	2 ± 1
QCD multijets	3200 ± 735	242 ± 80	31 ± 30
Total	31560 ± 2171	4854 ± 247	2889 ± 79
Data	31559	4801	2953

Posterior t tag	1.04 ± 0.06
Posterior t mistag	0.79 ± 0.06

r = 0.81 ± 0.05

Systematic Uncertainties in Fit (l+jets)

Experimental

- Correction factor uncertainties provided by respective POGs
  - Lepton ID, iso, trigger SFs, JEC, JER, b tag SF
- Uncertainties to be constrained by fit given a priori values
  - t tag SF: 25% uncertainty (separate SF for t tag and t mistag)
  - Background normalisations: 50% for QCD, 30% for other backgrounds
    - Separate normalisations for e and μ channel in QCD
    - Separate normalisations for W+light and W+heavy flavour
- Lumi, pileup uncertainties are not included in the fit

Theoretical Uncertainties

- **PDF**: standard deviation of 100 NNPDF3.0 MC replicas
- **Renormalization and factorization scales (μR/μF):** envelope or separately scaling μR, μF by 0.5, 2.0
- **Initial and Final State Radiation (ISR, FSR):** α_s for ISR (FSR) varied by a factor of $2\sqrt{2}$
- **Matrix element to parton shower (ME-PS) matching:** resumption damping factor h_{damp} varied by ±1σ
- **Color reconnection:** variant sample generated with color reconnection model applied
- **Underlying event:** tune CUETP8M2T4 parameters varied by ±1σ