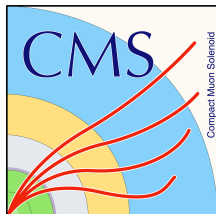


New Top Mass Results and Prospects for Run I Combination

Andrea Knue

on behalf of the ATLAS and CMS Collaborations

TopLHCWG meeting, 2nd November, 2017



New CMS measurement at 13 TeV

► CMS-PAS-TOP-17-007

- first m_{top} measurement using 13 TeV data from 2016: 35.9 fb^{-1}
- select events with exactly one lepton and at least 4 jets
 - ↔ 2 b -tagged jets amongst the leading 4
- compare three different unbinned likelihood fits:



1D Method:

↔ obtain $m_{\text{top}}^{\text{fit}}$



2D Method:

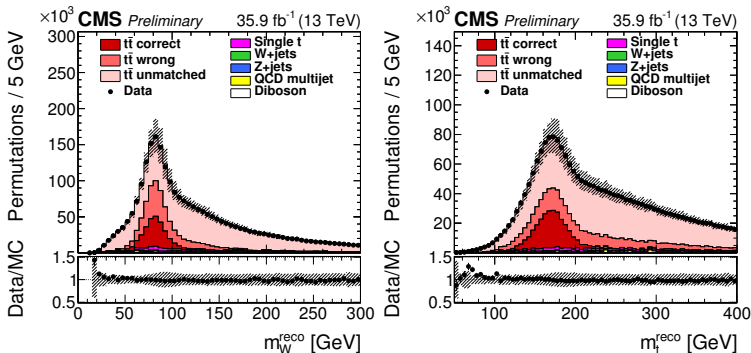
↔ obtain $m_{\text{top}}^{\text{fit}}$ and JSF



Hybrid Method:

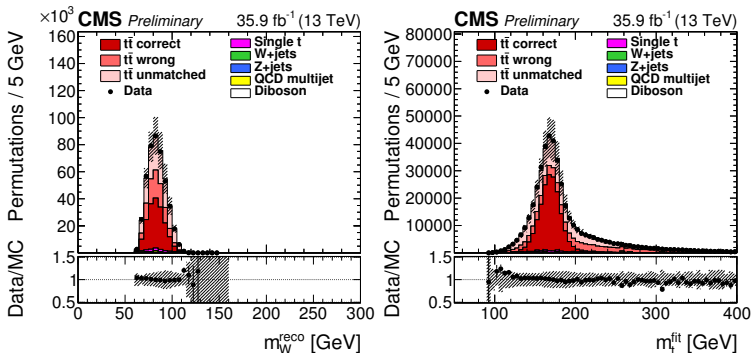
↔ obtain $m_{\text{top}}^{\text{fit}}$ and JSF
↔ constrain JSF with uncertainties from JEC (jet energy corrections)

Event reconstruction



- new nominal sample Powheg (NLO), at 8 TeV: Madgraph (LO)
- use 4 leading jets, charged lepton and E_T^{miss} as input
- per event: two jet-parton assignments, two solutions for p_z^{ν}
- calculate per permutation a goodness-of-fit probability: $P_{\text{gof}} = \exp(-0.5 \cdot \chi^2)$

Event reconstruction



Additional requirements:

- remove events with $P_{\text{gof}} = \exp(-0.5 \cdot \chi^2) < 0.2$: removes 76% events
- in addition: weight the permutations with P_{gof}
- fraction of correct permutations increases from 15 to 48 %

Template fit

Observables used in measurement:

- $m_{\text{top}}^{\text{fit}}$ as obtained from kinematic reconstruction
- $m_{\text{W}}^{\text{reco}}$ before constrained by kinematic fit

Now:

↔ sum over PDFs for correctly, wrongly and unmatched events

↔ sum over permutations, take into account P_{gof}

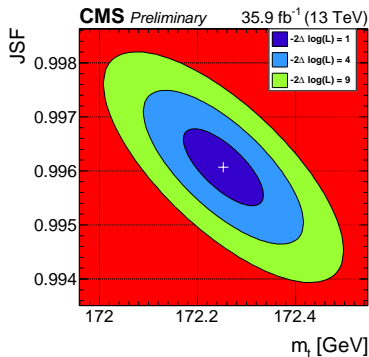
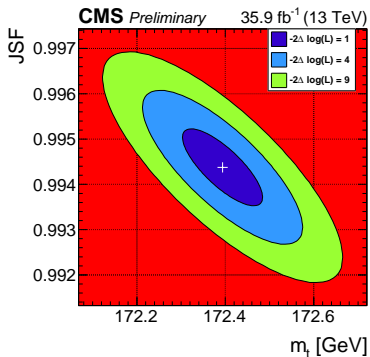
$$\mathcal{L}(\text{sample}|m_t, \text{JSF}) = P(\text{JSF}) \cdot \prod_{\text{events}} \left(\sum_{i=1}^n P_{\text{gof}}(i) \left(\sum_j f_j P_j(m_{t,i}^{\text{fit}}|m_t, \text{JSF}) \times P_j(m_{\text{W},i}^{\text{reco}}|m_t, \text{JSF}) \right) \right)^{w_{\text{evt}}}$$

Systematic uncertainties for three methods

Source	2D approach	1D approach	Hybrid
Jet energy corrections	0.13	0.85	0.19
JEC: Flavor	0.42	0.31	0.39
ME generator	0.19	0.29	0.22
FSR PS scale	0.24	0.22	0.13
Early resonance decays	0.22	0.42	0.03
Color reconnection modeling	0.34	0.23	0.31
Total systematic	0.71	1.09	0.62

- systematic uncertainty reduced by 35% when going from 1D to 2D
- systematic uncertainty reduced by 13% when going from 2D to Hybrid
 - ↪ largest reduction in Jet energy corrections and early resonance decays
- CR uncertainty increased compared to 8 TeV:
 - ↪ include now ERD in Pythia, and test different CR models

Result from fit to data

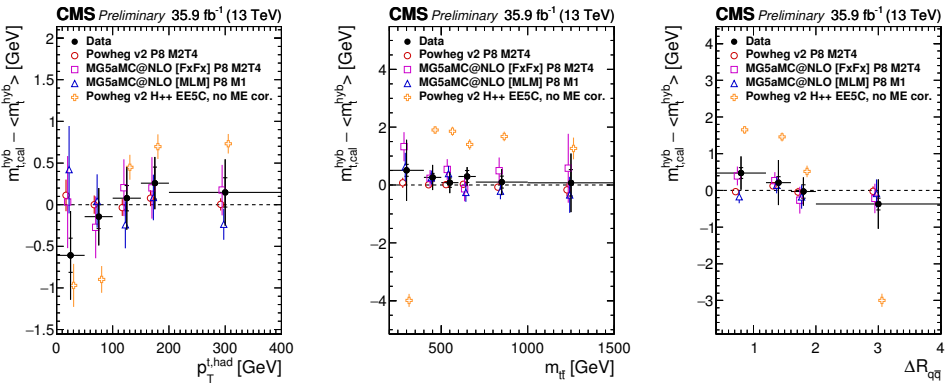


Result for Hybrid Method:

$$m_{\text{top}} = 172.25 \pm 0.08(\text{stat.} + \text{JSF}) \pm 0.62 (\text{syst.}) \text{ GeV}$$

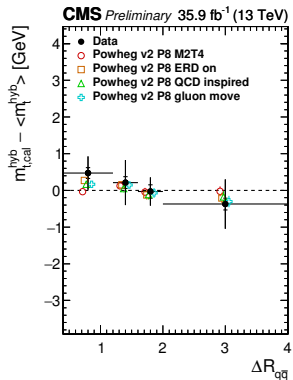
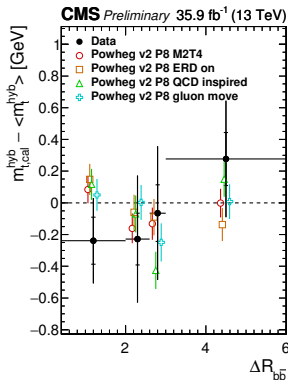
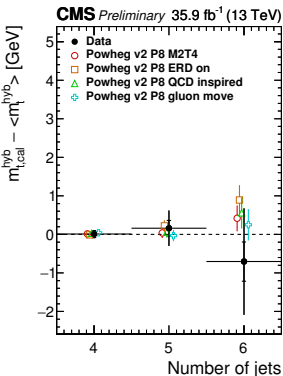
→ relative uncertainty: 0.36 %

Differential measurements: modelling of pQCD



- plot the mass per bin minus the mass obtained from the inclusive sample
- setup using Herwig++ shows large differences with data and other MC setups
 - ↪ has no matrix-element corrections to the top quark decay

Differential measurements: color reconnection models

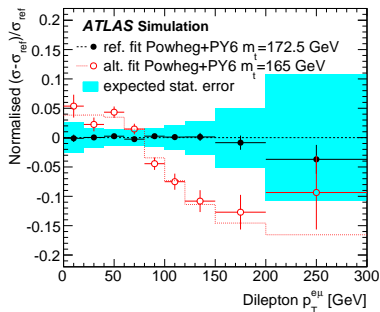
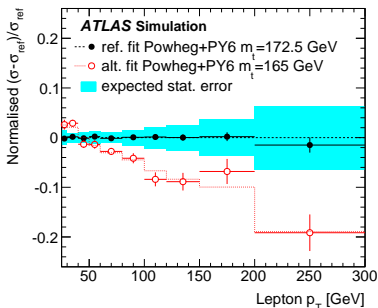


- nominal Powheg+Pythia8 setup models data measurement well
- $\Delta R(q, q)$: setup with early resonance decays agrees best with data
- in general: not enough precision to exclude one of the models

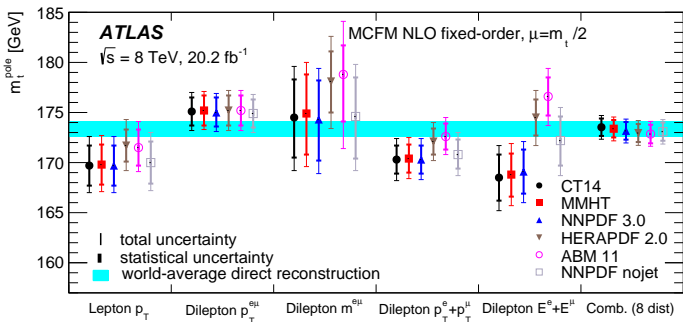
Dilepton differential cross-section at 8 TeV

► [arXiv:1709.09407](https://arxiv.org/abs/1709.09407)

- take events with opposite sign leptons e/μ + 1 or 2 b -tagged jets
- take differential lepton distributions which are sensitive to m_{top}
 \hookrightarrow unfold back to stable particle level
- obtain m_{top} and $m_{\text{top}}^{\text{pole}}$ using fits or from moments



$m_{\text{top}}^{\text{pole}}$ from fixed order predictions



- very good agreement with standard methods and other $m_{\text{top}}^{\text{pole}}$ measurements
- smallest uncertainty obtained from fit to $p_T^{e\mu}$ distribution
- largest uncertainty from choice of functional form for QCD scales
 \hookrightarrow benefit from NNLO predictions with QCD effects in top prod.+decay

Combined fit: $m_{\text{top}}^{\text{pole}} = 173.2 \pm 0.9 \text{ (stat.)} \pm 0.8 \text{ (exp.)} \pm 1.2 \text{ (theor.) GeV}$

New ATLAS result at 8 TeV

▶ ATLAS-CONF-2017-071

- 3D-Template method in the lepton+jets channel as in [Paper @ 7 TeV](#)
- Variable 1: $m_{\text{top}}^{\text{reco}}$ from reconstructed event (KLFitter)
- Variable 2: $m_{\text{W}}^{\text{reco}}$

- Variable 3:
$$R_{bq}^{\text{reco},2b} = \frac{p_{\text{T}}^{b_{\text{had}}} + p_{\text{T}}^{b_{\text{lep}}}}{p_{\text{T}}^{W_{\text{jet}1}} + p_{\text{T}}^{W_{\text{jet}2}}}$$

→ Variable 2 and 3 from chosen jet permutation (but original 4-vectors)

↔ m_{top} has sizable uncertainties from JES and bJES

↔ can be reduced by simultaneous measurement of m_{top} , jet energy scale factor (**JSF**) and relative b-to-light-jet energy scale factor (**bJSF**)

Event yields after preselection

Selection	Preselection		BDT selection	
Data	96105		38054	
$t\bar{t}$ signal	85000 \pm	10000	36100 \pm	5500
Single-top-quark signal	4220 \pm	360	883 \pm	85
NP/fake leptons (data-driven)	700 \pm	700	9.2 \pm	9.2
W +jets (data-driven)	2800 \pm	700	300 \pm	100
Z +jets	430 \pm	230	58 \pm	33
$WW/WZ/ZZ$	63 \pm	32	7.0 \pm	5.2
Signal+background	93000 \pm	10000	37300 \pm	5500
Expected background fraction	0.043 \pm	0.012	0.010 \pm	0.003
Data / (Signal+background)	1.03 \pm	0.12	1.02 \pm	0.15

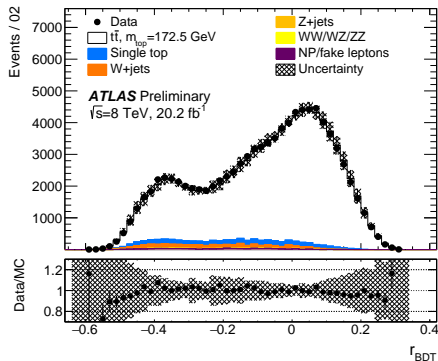
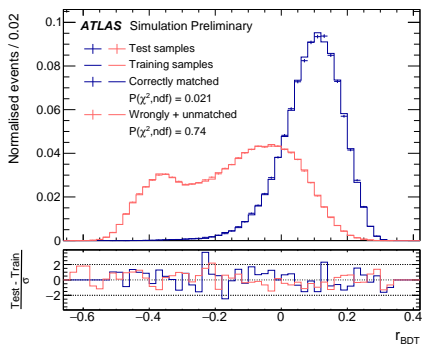
- exactly one lepton, at least 4 jets and exactly two b -tagged jets
- cut on exactly 2 b -tagged jets: significantly reduces background
- good data/MC agreement for both selections shown
- single-top production is m_{top} dependent and hence included in signal

Optimisation studies

- **assumption:** wrongly/unmatched events will have larger systematic uncertainties
- **idea:** find discriminant to distinguish correctly matched from wrongly/unmatched events
- **method:**
 - ↔ take variables (including some KL Fitter variables) and train a BDT algorithm
- cut on classifier output r_{BDT} to increase fraction of well reconstructed events

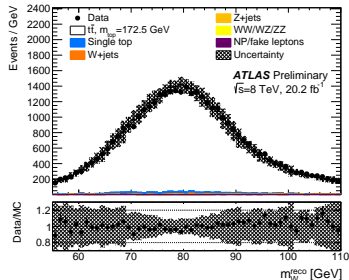
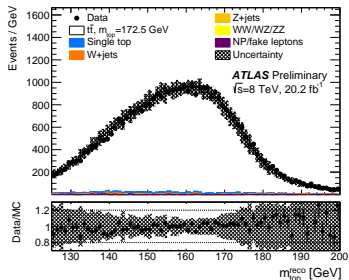
→ perform full analysis and compare total uncertainties

Discrimination power of BDT classifier

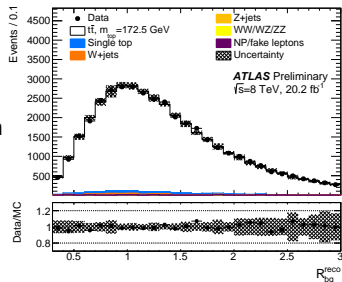


- Good agreement between training and test samples
 - Good agreement between data and prediction
- ⇒ the good performance of the algorithm allows to apply it to data

Data/MC agreement of the three observables $r_{\text{BDT}} > -0.05$



- $m_{\text{top}}^{\text{reco}}$ taken from kinematic fit
- $m_{\text{W}}^{\text{reco}}$ and $R_{\text{bq}}^{\text{reco}}$: chosen jet-permutation but from original 4-vectors
- only shape uncertainties shown here



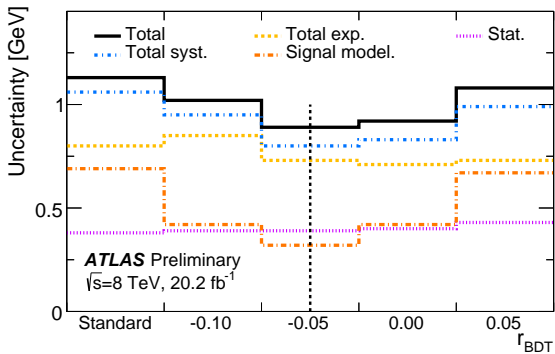
Finding the optimal r_{BDT} cut

- use 3D template method, perform unbinned likelihood fit
- perform full analysis chain for 4 BDT working points
 - ↔ evaluate total uncertainty using pseudo-experiments
- compare with result for “Standard” analysis (no BDT cut)

Likelihood function for unbinned fit:

$$L_{\text{shape}}^{\ell+\text{jets}}(m_{\text{top}}, \text{JSF}, \text{bJSF}, f_{\text{bkg}}) = \prod_{i=1}^N P_{\text{top}}(m_{\text{top}}^{\text{reco},i} | m_{\text{top}}, \text{JSF}, \text{bJSF}, f_{\text{bkg}}) \\ \times P_{\text{W}}(m_{\text{W}}^{\text{reco},i} | \text{JSF}, f_{\text{bkg}}) \\ \times P_{\mathcal{R}_{\text{bq}}}(R_{\text{bq}}^{\text{reco},i} | m_{\text{top}}, \text{JSF}, \text{bJSF}, f_{\text{bkg}})$$

Vary BDT requirement to obtain smallest total uncertainty



- working point with smallest total uncertainty is $r_{\text{BDT}} > -0.05$
- gain mostly caused by strongly reduced modelling uncertainties
 - ↪ trade here statistical for reduced systematic uncertainties
 - ↪ leads to a strongly reduced correlation to other estimates!

What do we gain from the BDT optimisation?

	m_{top} [GeV]		
	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV	
Event selection	Standard	Standard	BDT
Result	172.33	171.90	172.08
Statistics	0.75	0.38	0.39
– Stat. comp. (m_{top})	0.23	0.12	0.11
– Stat. comp. (JSF)	0.25	0.11	0.11
– Stat. comp. (bJSF)	0.67	0.34	0.35
Method	0.11 ± 0.10	0.04 ± 0.11	0.13 ± 0.11
Signal Monte Carlo generator	0.22 ± 0.21	0.50 ± 0.17	0.16 ± 0.17
Hadronisation	0.18 ± 0.12	0.05 ± 0.10	0.15 ± 0.10
Initial- and final-state QCD radiation	0.32 ± 0.06	0.28 ± 0.11	0.08 ± 0.11
Underlying event	0.15 ± 0.07	0.08 ± 0.15	0.08 ± 0.15
Colour reconnection	0.11 ± 0.07	0.37 ± 0.15	0.19 ± 0.15
Parton distribution function	0.25 ± 0.00	0.08 ± 0.00	0.09 ± 0.01
Background normalisation	0.10 ± 0.00	0.04 ± 0.00	0.08 ± 0.00
W +jets shape	0.29 ± 0.00	0.05 ± 0.00	0.11 ± 0.00
Fake leptons shape	0.05 ± 0.00	0	0
Jet energy scale	0.58 ± 0.11	0.63 ± 0.02	0.54 ± 0.02
Relative b -to-light-jet energy scale	0.06 ± 0.03	0.05 ± 0.01	0.03 ± 0.01
Jet energy resolution	0.22 ± 0.11	0.23 ± 0.03	0.20 ± 0.04
Jet reconstruction efficiency	0.12 ± 0.00	0.04 ± 0.01	0.02 ± 0.01
Jet vertex fraction	0.01 ± 0.00	0.13 ± 0.01	0.09 ± 0.01
b -tagging	0.50 ± 0.00	0.37 ± 0.00	0.38 ± 0.00
Leptons	0.04 ± 0.00	0.16 ± 0.01	0.16 ± 0.01
$E_{\text{T}}^{\text{miss}}$	0.15 ± 0.04	0.08 ± 0.01	0.05 ± 0.01
Pile-up	0.02 ± 0.01	0.14 ± 0.01	0.15 ± 0.01
Total systematic uncertainty	1.03 ± 0.08	1.07 ± 0.10	0.82 ± 0.06
Total	1.27 ± 0.08	1.13 ± 0.10	0.91 ± 0.06

Precision of systematic uncertainty:

$$u \pm s = \sqrt{\sum_k u_k^2} \pm \frac{\sqrt{\sum_k (s_k^2 u_k^2)}}{u}$$

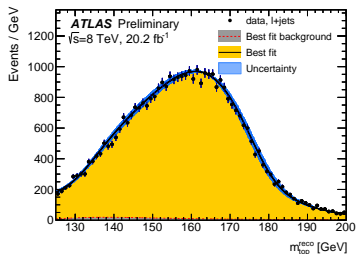
↪ use to test stability of combination

→ total systematic uncertainties almost equal for standard selections at 7 and 8 TeV

→ overall uncertainty is reduced by 19 % when using BDT optimisation

Result in data for $r_{\text{BDT}} > -0.05$

$$m_{\text{top}} = 172.08 \pm 0.39 \text{ (stat) GeV}$$

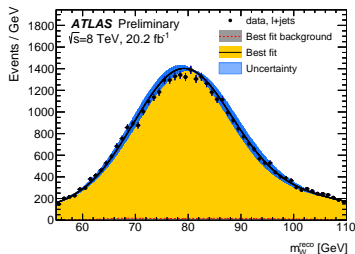


$$\rho_{\text{stat}} = \begin{pmatrix} 1.0 & & & \\ -0.27 & 1.0 & & \\ -0.92 & -0.02 & 1.0 & \\ & & & 1.0 \end{pmatrix}$$

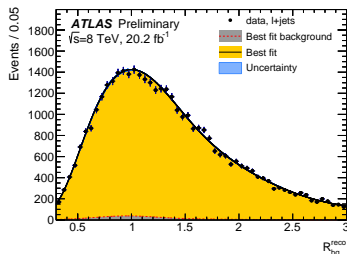
Final result lepton+jets channel:

$$m_{\text{top}} = 172.08 \pm 0.39 \text{ (stat)} \pm 0.82 \text{ (syst) GeV}$$

$$\text{JSF} = 1.005 \pm 0.001 \text{ (stat)}$$



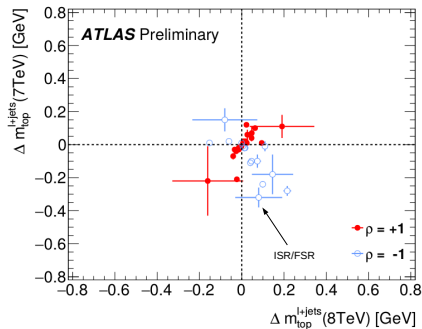
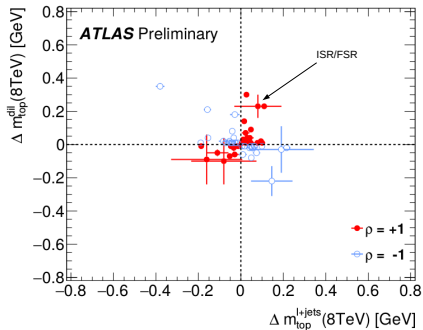
$$\text{bJSF} = 1.008 \pm 0.005 \text{ (stat)}$$



Combination procedure

- use BLUE method (best linear unbiased estimator)
 - ▶ [Link hepforge](#)
 - ▶ [Eur. Phys. J. C \(2014\) 74](#)
- for each source of uncertainty the correlation between any pair of analyses needs to be known
- need to properly map the uncertainty components between 7 and 8 TeV (JES, b -tagging)
- if uncertainty has no equivalent at other centre-of-mass energy:
 - ↪ treat as independent

Pairwise estimator correlation for all systematic uncertainties

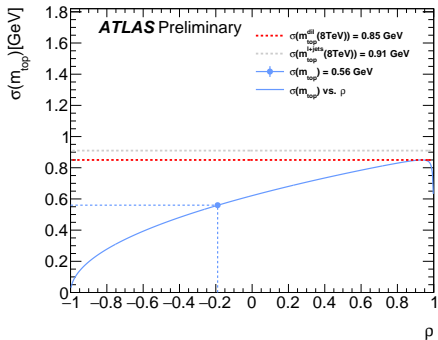
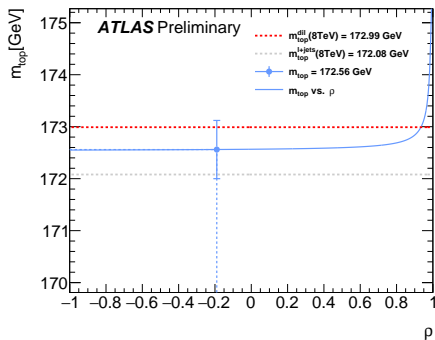


→ red full points correspond to $\rho = +1$, blue open points to $\rho = -1$

→ ISR/FSR (left): dilepton (8 TeV): 0.23 ± 0.07 , l+jets (8 TeV): 0.08 ± 0.11 GeV

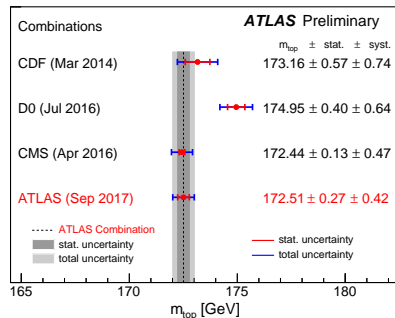
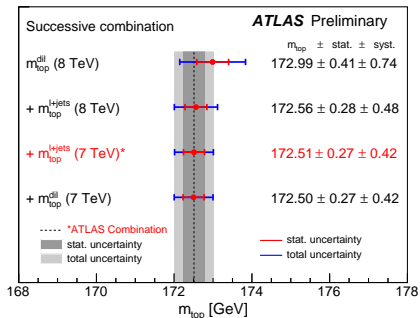
→ ISR/FSR (right): l+jets (7 TeV): -0.32 ± 0.06 , l+jets (8 TeV): 0.08 ± 0.11 GeV

Combination of pair of most important results



- evaluate the combination (value and uncertainty) as function of the correlation
- blue point: actual correlation between lepton+jets and dilepton at 8 TeV
- correlation between two channels is small due to different analysis techniques

Combination with previous ATLAS results



→ left : change of combined result when consecutively adding results

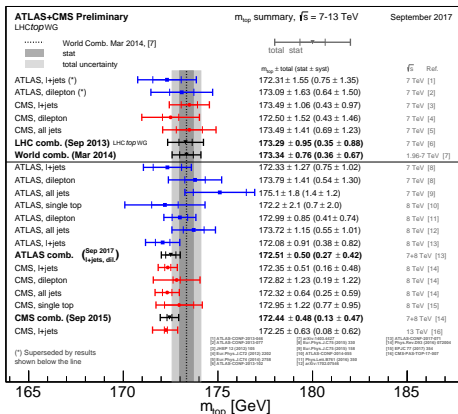
↪ 7 TeV dilepton result does not improve result significantly

↪ has not been included in final combination

→ central values for ATLAS and CMS agree well, both have $< 0.3\%$ uncertainty:

↪ ATLAS ($\sigma_{\text{tot}} = 0.50 \pm 0.04 \text{ GeV}$) and CMS ($\sigma_{\text{tot}} = 0.49 \text{ GeV}$)

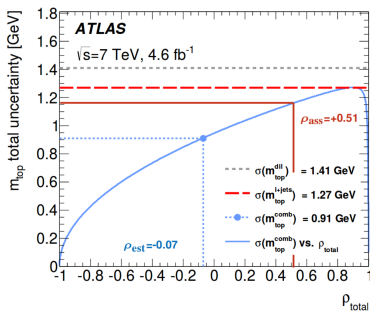
What steps need to be taken for final Run I combination?



- Run 1 results now available in all channels from both experiments
- contribute with very different weights to the combination
- What measurements should be included in the combination?

Important: estimate correlations between measurements

Example: Combination of lepton+jets and dilepton channel @ 7 TeV



→ including the sign of the correlations: $\sigma_{m_{\text{top}}} = 0.91\text{ GeV}$

→ in case the sign of the correlation per sub-component is ignored

↔ the uncertainty is: $\sigma_{m_{\text{top}}} = 1.16\text{ GeV}$

→ Need to evaluate signs in systematic shifts between different centre-of-mass energies and between experiments.

Correlation strategy

▶ Talk Alison

- 1 correlation within each collaboration:
 - ↔ information already exists, has been used for the CMS and ATLAS combinations shown in slide 26
- 2 correlation of estimates from the different collaborations
 - ↔ mapping of uncertainties needs to be found

→ Can we also calculate the statistical precision of the systematics?

Detector uncertainties:

- b -tagging: uncorrelated
- some JES components are correlated between different CME (e.g. η -intercalibration, flavour composition/response, ...)
- studies to compare ATLAS and CMS JES procedures were performed ▶ TOPLHC note → details see next slide

Correlation strategy: Jet energy scale

► TOPLHC note

Description	Components, CMS	Components, ATLAS	Corr. range
1a. Statistical <i>in situ</i> terms	AbsoluteStat, SinglePionHCAL, RelativeStat[FSR][EC2][HF]	[11] Z -jet balance stat./meth. terms (p_T), [13] γ -jet balance stat./meth. terms (p_T), [10] multi-jet balance stat./meth. terms (p_T), η -intercalibration statistical term (p_T, η)	0%
1b. Detector <i>in situ</i> terms	AbsoluteScale, SinglePionECAL, RelativeJER[EC1][EC2][HF], RelativePt[BB][EC1][EC2][HF]	Z -jet balance det. term, γ -jet balance det. term, [2] correlated Z/γ -jet balance det. terms (p_T)	0%
2. Absolute balance modeling	AbsoluteMPFBias	[7] Z -jet balance model + mixed terms (p_T), [4] γ -jet balance model + mixed terms (p_T), [2] correlated Z/γ -jet balance terms (p_T), [5] multi-jet balance model + mixed terms (p_T)	0-50%
3. Relative balance modeling	RelativeFSR	η -intercalibration modeling (p_T, η)	50-100%
4. g -jet fragmentation	FlavorPureGluon	Flavor response (p_T, η)	100%
5. b -jet fragmentation	FlavorPureBottom	b -jet response (p_T)	50-100%
6. Other fragmentation types	FlavorPureQuark, FlavorPureCharm	Flavor composition (p_T, η)	0%
7. Pileup	PileupDataMC, PileupPt[Ref][BB][EC1][EC2][HF]	N_{PV} offset (p_T, η, N_{PV}), $\langle \mu \rangle$ offset ($p_T, \eta, \langle \mu \rangle$), p_T term ($p_T, \eta, N_{PV}, \langle \mu \rangle$), ρ topology (p_T, η)	0%
8. High- p_T	Fragmentation	High- p_T (p_T)	0%
9. Single-experiment terms	TimeEta, TimePt	Fast simulation closure (p_T, η), punch-through ($p_T, \eta, N_{segments}$)	0%

Components with non-zero correlation:

- absolute/relative balance modelling (category 2 and 3)
- g/b -jet fragmentation (category 4 and 5)

Correlation strategy: modelling

- MC background normalisation uncertainties: correlated
- signal and background modelling uncertainties:
 - ↔ Needs to be seen case by case. Did not use same generators / variations, but some methods are similar

Example ISR/FSR uncertainty:

Check if m_{top} increases/decreases for more/less radiation. Do not distinguish how the change in variation is achieved

↔ example ATLAS 7 TeV: based on AcerMC, 8 TeV: based on Powheg

Example ME Generator uncertainty:

→ CMS: Powheg vs. Madgraph

→ ATLAS: Powheg vs. MC@NLO

Summary

First 13 TeV result using 2016 data by CMS in lepton+jets channel:

- largest impact from JEC and signal modelling uncertainties
- **already relative uncertainty of 0.36 %** [▶ CMS-PAS-TOP-17-007](#)
- new nominal MC generator and higher centre-of-mass energy
↔ result consistent with 8 TeV measurement

8 TeV data:

- ATLAS lepton+jets result was made public for TOP conference
- BDT optimisation allowed to reduce systematic uncertainties
- dominating uncertainties: JES, *b*-tagging, signal modelling

↔ **performed ATLAS combination: relative uncertainty of 0.29 %**

→ Time to combine ATLAS and CMS results to create the Run I legacy!

Object definition

Electrons

- $E_T > 25$ GeV, $|\eta| < 2.47$
- no crack region, Tight++
- e24vhi_medium
OR e60_medium1

Muons

- $p_T > 25$ GeV, $|\eta| < 2.5$
- mu24i_tight OR mu36_tight
- mini-isolation

small- R jets

- antiKt 0.4, LC TopoClusters
- JVF > 0.5 for $p_T < 50$ GeV
and $|\eta| < 2.4$
- b -tagging: MV1, 70% WP
- $p_T > 25$ GeV for $|\eta| \leq 2.5$

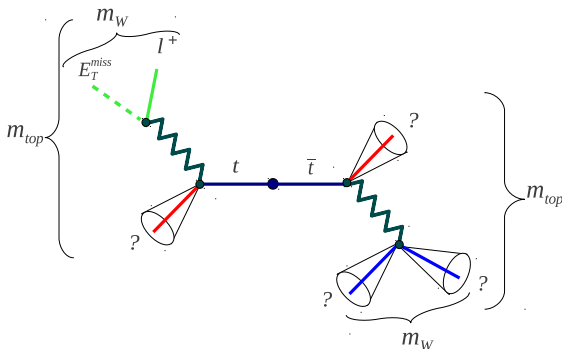
$E_T^{\text{miss}} / M_T^W$

- $E_T^{\text{miss}} > 30$ GeV, $M_T^W > 30$ GeV
(e +jets)
- $E_T^{\text{miss}} > 20$ GeV,
 $E_T^{\text{miss}} + M_T^W > 60$ GeV (μ +jets)

Event reconstruction

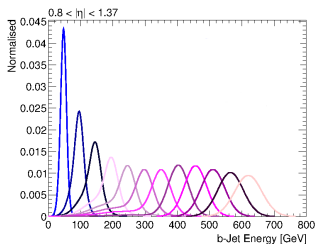
First approach:

- expect 4 jets
 - ⇒ 24 possible jet-parton assignments
- do not distinguish light jets within W
 - ⇒ 12 permutations left
- pick one for calculation of variables
 - ⇒ use kinematic fit to find best permutation



Reconstruction with KLFilter

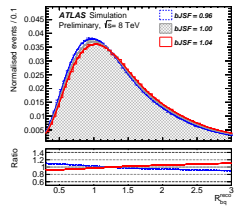
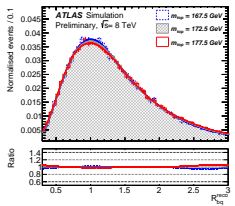
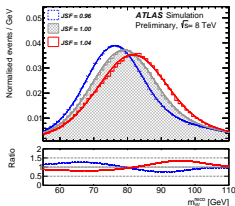
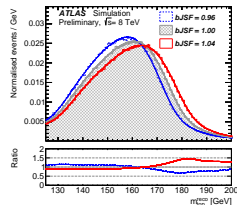
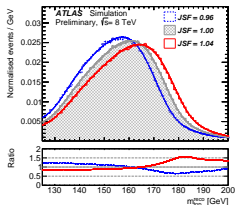
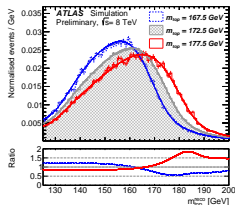
$$\begin{aligned}
 L = & BW(m_{q_1 q_2} | m_W, \Gamma_W) \cdot BW(m_{l\nu} | m_W, \Gamma_W) \\
 & BW(m_{q_1 q_2 b_{had}} | m_{top}, \Gamma_{top}) \cdot BW(m_{l\nu b_{lep}} | m_{top}, \Gamma_{top}) \\
 & W(\tilde{E}_{jet_1} | E_{b_{had}}) W(\tilde{E}_{jet_2} | E_{b_{lep}}) W(\tilde{E}_{jet_3} | q_1) W(\tilde{E}_{jet_4} | q_2) \\
 & W(\tilde{E}_x^{miss} | p_{x,\nu}) W(\tilde{E}_y^{miss} | p_{y,\nu}) \left\{ \begin{array}{l} W(\tilde{E}_l | E_l) \\ W(\tilde{p}_{T,l} | p_{T,l}) \end{array} \right\}
 \end{aligned}$$



- to increase reconstruction efficiency:
 - ↪ use up to 6 jets with highest p_T as input
- detector resolution: encoded in transfer functions
- veto b -tagged jets in position of light jets and vice versa

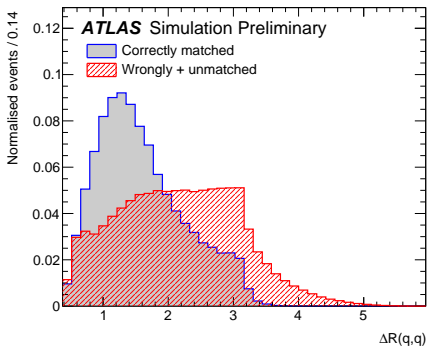
→ Choose jet-parton assignment with maximum Likelihood for analysis

Template parametrisation



→ Template fit functions are used as pdfs in an unbinned likelihood fit to data

BDT: use 13 input variables



Separation	Variable	Description
31%	$\ln L$	Logarithm of the likelihood of best permutation
13%	$\Delta R(q, q)$	ΔR of the two untagged jets from the hadronically decaying W boson
5.0%	$p_T(W_{\text{had}})$	p_T of hadronically decaying W boson
4.3%	$p_{T,\text{had}}$	p_T of hadronically decaying top quark
4.2%	P_{Evt}	Event probability of best permutation
2.0%	$p_T(\bar{t}t)$	p_T of reconstructed $\bar{t}t$ system
1.7%	$p_{T,\text{lep}}$	p_T of leptonically decaying top quark
1.2%	m_T^W	Transverse mass of leptonically decaying W boson
0.3%	$p_T(W_{\text{lep}})$	p_T of leptonically decaying W boson
0.3%	N_{jets}	Number of jets
0.2%	$\Delta R(b, b)$	ΔR of reconstructed b -tagged jets
0.2%	E_T^{miss}	Missing transverse momentum
0.1%	$p_{T,\ell}$	p_T of lepton

Correlation between ATLAS results

Correlations	$m_{\text{top}}^{\text{dil}}$ (7 TeV)	$m_{\text{top}}^{\text{l+jets}}$ (7 TeV)	$m_{\text{top}}^{\text{dil}}$ (8 TeV)	$m_{\text{top}}^{\text{l+jets}}$ (8 TeV)
$m_{\text{top}}^{\text{dil}}$ (7 TeV)	1.00			
$m_{\text{top}}^{\text{l+jets}}$ (7 TeV)	-0.07	1.00		
$m_{\text{top}}^{\text{dil}}$ (8 TeV)	0.52	0.00	1.00	
$m_{\text{top}}^{\text{l+jets}}$ (8 TeV)	0.06	-0.07	-0.19	1.00
BLUE weight (m_{top})	-	0.17	0.43	0.40

→ correlation between two channels is small

↔ due to different analysis techniques

→ measurements not strongly correlated between 7 and 8 TeV

↔ due to many uncorrelated detector uncertainties