ATLAS and CMS b-tagging calibration and uncertainties

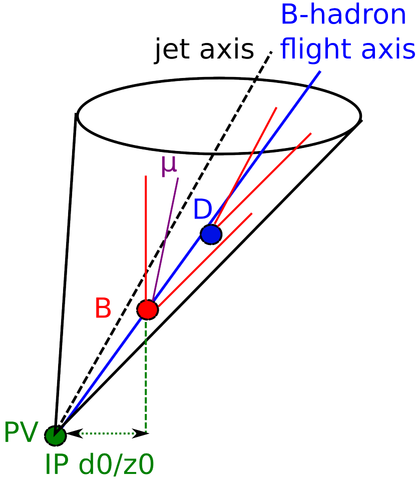
L. Mijović[1], M. E. G. zur Nedden[2], L. Scodellaro[3]

[1] Universität Bonn
[2] Humboldt-Universitaet zu Berlin
[3] Instituto de Fisica de Cantabria - CSIC

Outline

- ATLAS and CMS b-tagging algorithms in Run2
- ATLAS and CMS b-tagging algorithm calibration in Run2
- b-tagging uncertainties correlations proposal in Run1
- b-tagging uncertainties correlations in Run2:
  proposal for first combinations and next steps

Introduction to b-tagging



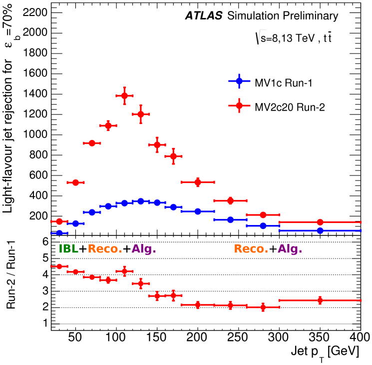
b-tagging & discriminating variables used in tagging algorithms:

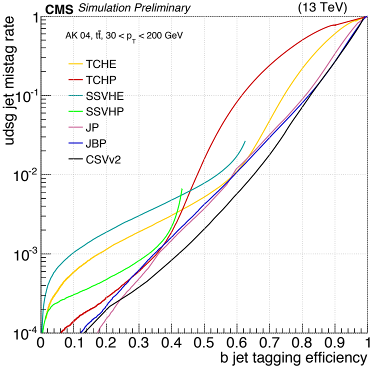
- track Impact Parameter
  signed transverse and longitudinal impact parameters ($d0$,$z0$) and their significance
- presence of secondary vertices
  SV mass, energy fraction, Ntracks, decay length ...
- B decay chain reconstruction
- B hadron decays to leptons
  useful for tagging and calibrations

Discriminating variables are combined into tagging algorithms.

Expected performance

ATL-PHYS-PUB-2015-022 and CMS DP-2015/038
Both ATLAS and CMS have improved their b-tagging performance for Run2: algorithm improvements, ATLAS: IBL.



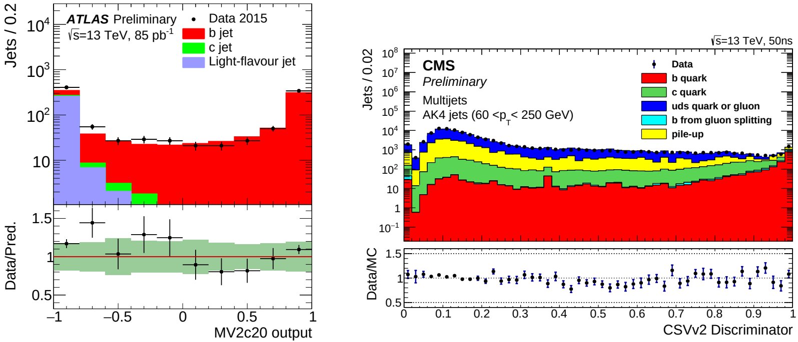


b-tagging efficiency of about 70%: light jet rejection of about 100 and c-jet rejection of about 10 (AntiKt0.4 jets in $t\bar{t}$ events).

Run2 commissioning

universitätbonn erc

ATL-PHYS-PUB-2015-039 and CMS DP-2015/045
Commissioned input variables & final discriminants performance with early Run2 data.



Tagger calibrations

Residual discrepancies between the tagging efficiency in data and simulation are corrected in tagger calibrations.

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}}$$

(also derived for mis-tag rates and c-jets).

Calibration strategies:

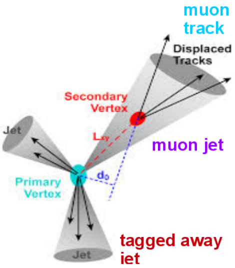
- general idea: measure b-jet content of a b-enriched jet sample before and after the b-tagging requirement

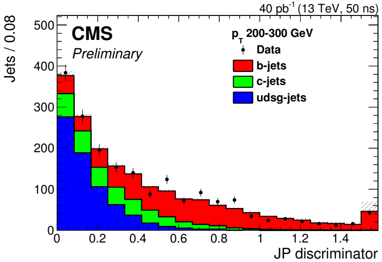
b-enriched sample can contain e.g.:

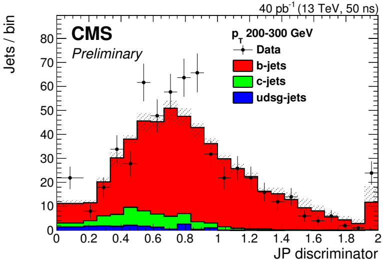
- jets with a soft muon:
  - independent from top analysis datasets
  - in Run1: uncertainties larger than in $t\bar{t}$ calibrations
- $t\bar{t}$ events:
  - single lepton and dilepton decay channels provide two orthogonal datasets
  - most precise calibrations in Run1, expected to (eventually) have more weight than muon-in-jet calibrations in Run2.

Calibration: jets with soft muon

- CMS DP-2015/045
- selecting a b-enriched sample by requiring a muon in a jet
- the sample can be further enriched in $b\bar{b}$ by requiring a tagged away jet
- pretag and posttag plots of the discriminant variable (JP distribution) for one of the calibration methods (life-time)

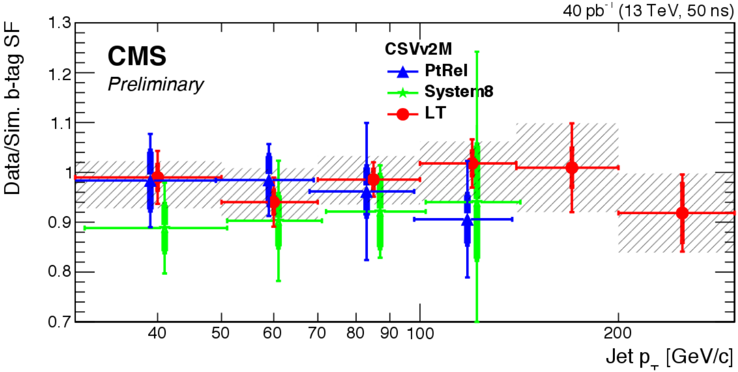






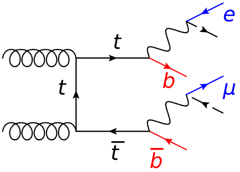
Calibration: jets with soft muon

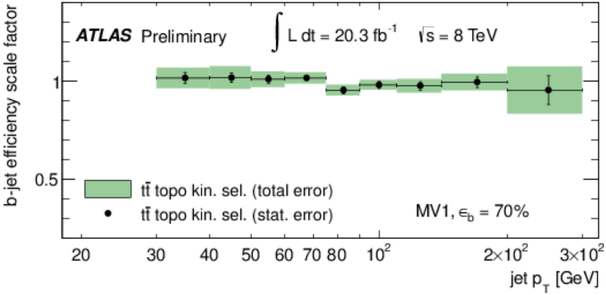
- CMS DP-2015/045
- scale factors of Data/Simulation tagging efficiency as a function of jet p_T
- figure: scale factors from three calibration methods (Ptrel,System8,LT) and their combination for medium light jet mis-tag probability working point (0.01)



Calibration: ttbar

- most precise calibrations in Run1
- E.g. : dilepton $t\bar{t}$ events: pure sample of b-jets.
- figure: ATLAS Run1 calibration in $t\bar{t}$ dileptons events
- Run2 $t\bar{t}$ calibrations:
  $\sigma_{t\bar{t}}(13\text{TeV})/\sigma_{t\bar{t}}(8\text{TeV}) \sim 830\ \text{pb}/250\ \text{pb} \sim 3.3$
  $\mathcal{L}_{\text{int}}\ (13\text{TeV})/\mathcal{L}_{\text{int}}\ (8\text{TeV}) \sim 4.3\ \text{fb}^{-1}/23\ \text{fb}^{-1} \sim 0.2$.
- $\Rightarrow$ statistical precision: $\sqrt{1/(3.3 \cdot 0.20)} \sim 1.2 \cdot X$
  worse than in Run1.
  $X >\sim 1$ accounts for difference in fractions of good data between Run2 and Run1





Uncertainty Correlations

b-tagging uncertainties can be split in categories according to whether:

- they are correlated with uncertainties evaluated in the top analyses
  correlations handled within experiment, not discussed in the following
- they are correlated between ATLAS and CMS
  correlations handled during ATLAS and CMS combination, discussed in the following

b-tagging uncertainty category	**Correlated with top analysis**	**Correlated between ATLAS and CMS**
General physics modelling (parton shower etc.)	**YES**	**YES**
Specific physics modelling (B hadron energy spectrum, μ p_{T} modelling etc.)	**NO**	**YES**
Detector modelling (calorimeter response, pile-up etc.)	**YES**	**NO**
Calibration method specific	**NO**	**NO**

NB: many systematic uncertainties related to physics object reconstruction also contain physics modelling uncertainties. Eg JES . . .

Combination in Run1

Recommendations for b-tagging combination in Run1 were developed, documented and discussed:

- LHCTopWG wiki pages
- presentation by L. Scodellaro, M. z. Nedden and J. M. Keaveney at the TOP LHC WG meeting open session, 21-23 May 2014

Correlated uncertainties for Run1 top analyses:

Group	source	size at ATLAS	size at CMS
soft muon calibration	b/c prod.	low pT: 0.1% - 0.2%, high pT for b-prod.: 1.2% - 2.0%	low pT: 0.1% - 0.3%, high pT: 0.5% - 1.3%
soft muon calibration	mu pT	first pT bin: 2.5%, 0.2% - 0.9% elsewhere	low pT: 0.1% - 1.1%, high pT: 0.1 - 0.9%
soft muon calibration	c/l ratio	<0.1% - 0.2%	<0.1% - 0.2%
soft muon calibration	b-frag	0.2% - 2.7%	0.2% - 0.8%
ttbar calib.	PS	0.1% - 1.5%	0.3% - 0.6%
ttbar calib.	IFSR	0.3% - 1.4%	0.3% - 0.6%

(Charm-to-light systematics was merged with b-frag.)

Combination in Run1 cont'd

Example of information provided for a specific algorithm (from L. Scodellaro's talk):

pT bin	30-40	40-50	50-60	60-70	70-80	80-100	100-120	120-160	160-210	210-260	260-320	320-400	>400
Uncor.	1.8	1.7	1.9	1.7	1.7	2.1	2.0	2.1	1.5	2.1	2.2	4.6	4.6
b/c prod.	0.1	<0.1	0.2	0.3	0.1	0.4	0.5	0.6	1.0	0.7	1.3	0.9	1.0
muon pT	0.1	0.1	0.7	0.6	0.4	1.0	0.8	0.5	0.1	0.1	0.1	0.8	1.0
b-frag.	0.7	0.7	0.4	0.6	0.5	0.5	0.3	0.2	0.2	0.3	0.2	0.4	0.4
PS	0.4	0.4	0.5	0.4	0.5	0.6	0.5	0.4	0.3	0.3	0.3	0.3	0.3
IFSR	0.4	0.4	0.5	0.4	0.5	0.6	0.5	0.4	0.3	0.3	0.3	0.3	0.3

- Analysers : propagate all the six of sets of uncertainty.
- Using the same technique as for the global uncertainty used so far.
- Resulting uncertainties on the final results can be combined taking into account their correlations between the ATLAS and CMS.

Run2 Uncertainty Correlations

Top analyses to be combined for the first Run2 combinations of results available on the time-scale of the Moriond conference are expected to be either of:

- **relying predominantly on calibrations in which Run2 $t\bar{t}$-based calibrations have the largest weight (ATLAS and CMS):**
  - candidate analyses: single top and $t\bar{t}$ l+jets channel measurements
  - **our proposal: take correlations into account for General physics modelling categories**:
  $t\bar{t}$ generator choice, parton shower, I/FSR and ren/fac scale variations

- **rely on predominantly on Run2 muon-in-jet calibration (CMS)**
  - uncertainty correlation cannot be taken into account for 1st combinations
  - later on- correlations can likely be taken into account for Specific physics modelling categories:
  b/c production, soft muon p_{T} spectrum modelling and b-fragmentation.

- **rely on Run1 extrapolation**
  - uncertainty correlation will not be taken into account.

- **dilepton $t\bar{t}$ measurements that employ in-situ b-tagging efficiency calibrations**
  - different calibration strategies relevant for these analyses on Moriond time-scale by ATLAS and CMS
  - treat b-tagging uncertainties as uncorrelated for first combinations

$t\bar{t}$ generator choice

Uncertainty due to the choice of the matrix element generator used in the flavour tagging calibrations for the $t\bar{t}$ signal modelling.
Size in $t\bar{t}$ calibrations performed at $\sqrt{s} = 8$ TeV:

- $\sim$ 10-50% contribution to the total systematics uncertainty

ATLAS	CMS	Correlated
Powheg-BOX+ HERWIG++ vs MG5_aMC@NLO+ HERWIG++	Powheg-BOX+PYTHIA8 vs MG5_aMC@NLO+PYTHIA8	YES

Comments:

- table: strategies as used in first Run2 $t\bar{t}$ cross-section measurements
- suggest correlated treatment also in case different parton shower generator (ATLAS: HERWIG++, CMS:PYTHIA8) is used in case of ATLAS and CMS
- further synchronisation of strategies would be of benefit and is likely in Run2

$t\bar{t}$ parton shower

Uncertainty due to the choice of the parton shower and hadronisation generator used in the flavour tagging calibrations for the $t\bar{t}$ signal modelling
Size in $t\bar{t}$ calibrations performed at $\sqrt{s} = 8$ TeV:

- $\sim$ 10-50% contribution to the total systematics uncertainty

ATLAS	CMS	Correlated
Powheg-BOX+ HERWIG++ vs Powheg-BOX+PYTHIA6(8?)	Powheg-BOX+ HERWIG++ vs Powheg-BOX+PYTHIA8	YES(?)

Comments:

- table: strategies in first Run2 $t\bar{t}$ cross-section measurements
- comparison between same shower and hadronisation generators (Powheg-BOX+ HERWIG++vs Powheg-BOX+PYTHIA8) feasible

Assuming the same shower generators and hadronisation are used, remaining smaller caveats are:

- different tunes will likely be used by ATLAS and CMS
- ATLAS uses EVTGEN for b- and c- hadron decays

$t\bar{t}$ radiation and scale uncertainty

Uncertainty due to the choice of the renormalisation/factorisation (μ_R/μ_F) scales in the matrix element generator and additional radiation settings in the flavour tagging calibrations for the $t\bar{t}$ signal modelling.
Size in $t\bar{t}$ calibrations performed at $\sqrt{s} = 8$ TeV:

- $\sim$ 10-50% contribution to the total systematic uncertainty

ATLAS	CMS	Correlated
Powheg-BOX+PYTHIA6 ME: $\mu_R/\mu_F \uparrow$ P2012 radLo, hdamp=mtop vs ME: $\mu_R/\mu_F \downarrow$, P2012 radHi, hdamp=2mt	Powheg-BOX+PYTHIA8 ME: $\mu_R/\mu_F \uparrow$ vs ME: $\mu_R/\mu_F \downarrow$	NEEDS FOLLOW-UP update of I/FSR strategy expected by CMS

Comments:

- table: strategies in first Run2 $t\bar{t}$ cross-section measurements
- further synchronisation of strategies would be of benefit and is likely in Run2

Correlations: next steps

Proposal for flavour tagging uncertainties correlation in 1st Run2 combinations:
for analyses using Run2 $t\bar{t}$ calibrations:

Category	Correlated
$t\bar{t}$ generator choice	YES
$t\bar{t}$ parton shower	YES
$t\bar{t}$ radiation and scale uncertainty	NEEDS FOLLOW-UP

Other analyses: treat flavour tagging uncertainties as uncorrelated (but aim to improve on this).

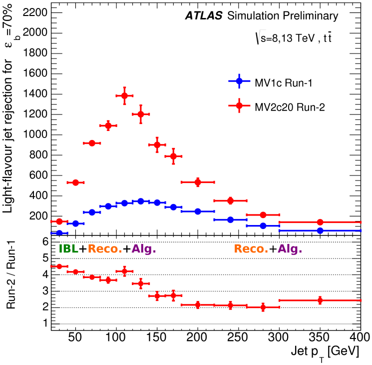
Next steps

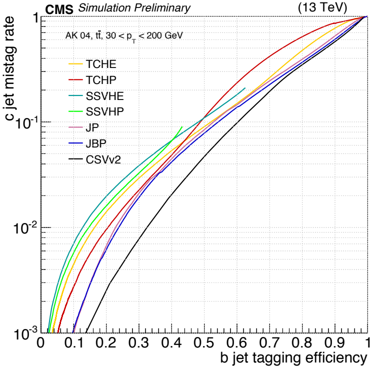
- iterate on the above as MC generator uncertainty recommendations evolve
- check the effects of the variations to see if the correlated treatment is justified.
- follow-up on Run2 soft muon tagger calibrations:
  strategy developed in Run1 TOPLHCWG discussions is a valid baseline for Run2
- consider correlations of the JES MC modelling components. JES expected to be an important source of systematic uncertainty in calibrations.

Extra

Expected performance

ATL-PHYS-PUB-2015-022
CMS DP-2015/038
Both ATLAS and CMS have improved their b-tagging for Run2.

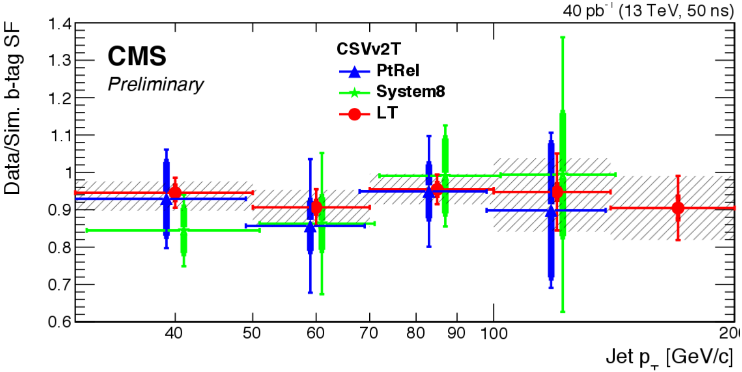




At light jet rejection of 0.01 : b-tagging efficiency of about 70% and c-jet rejection of about 10 (AntiKt0.4 jets in $t\bar{t}$ events).

Calibration: jets with soft muon

- CMS DP-2015/045
- scale factors of Data/Simulation tagging efficiency as a function of jet p_{T}
- figure: scale factors from three calibration methods (Ptrel,System8,LT) and their combination for tight light jet mis-tag probability working point (0.001)



Calibration uncertainties

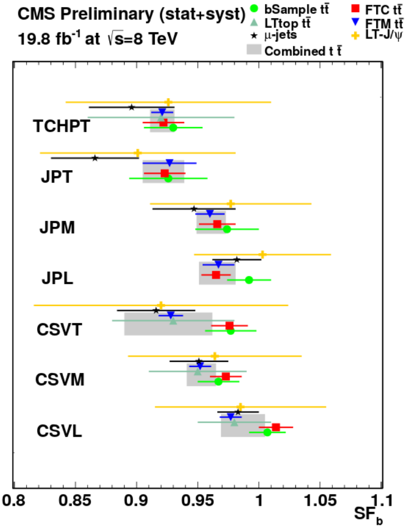
CMS Run1, J. Instrum. 8 (2013) P04013

Table 4: Scale factors SF_b obtained in muon-jet data, J/ψ events and $t\bar{t}$ data for b jets in the p_T range of $t\bar{t}$ events. The overall uncertainties are given.

b tagger	SF_b in muon jets	SF_b in LTJ/ψ events	SF_b in $t\bar{t}$ events
JPL	0.982 ± 0.020	1.003 ± 0.056	0.966 ± 0.015
CSVL	0.983 ± 0.017	0.985 ± 0.070	0.987 ± 0.018
JPM	0.947 ± 0.034	0.977 ± 0.066	0.961 ± 0.012
CSVM	0.951 ± 0.024	0.964 ± 0.071	0.953 ± 0.012
TCHPT	0.896 ± 0.035	0.926 ± 0.084	0.921 ± 0.010
JPT	0.866 ± 0.036	0.901 ± 0.080	0.922 ± 0.017
CSVT	0.916 ± 0.032	0.920 ± 0.104	0.926 ± 0.036

Calibration uncertainties

CMS Run1, J. Instrum. 8 (2013) P04013



Calibration uncertainties

CMS Run1, J. Instrum. 8 (2013) P04013

Table 3: b-tagging efficiency scale factors derived from tt events. The overall uncertainties are given.

Method	bSample	FTC	LTtop	FTM	Combined
JPL	0.992 ± 0.018	0.965 ± 0.012		0.967 ± 0.013	0.966 ± 0.015
CSVL	1.007 ± 0.015	1.014 ± 0.014	0.98 ± 0.03	0.977 ± 0.009	0.987 ± 0.018
JPM	0.974 ± 0.026	0.966 ± 0.015		0.960 ± 0.012	0.961 ± 0.012
CSVM	0.967 ± 0.017	0.973 ± 0.013	0.95 ± 0.04	0.952 ± 0.009	0.953 ± 0.012
TCHPT	0.930 ± 0.024	0.922 ± 0.017	0.92 ± 0.06	0.921 ± 0.009	0.921 ± 0.010
JPT	0.926 ± 0.032	0.923 ± 0.017		0.927 ± 0.022	0.922 ± 0.017
CSVT	0.977 ± 0.021	0.976 ± 0.015	0.93 ± 0.05	0.928 ± 0.010	0.926 ± 0.036

Calibration uncertainties

ATLAS-CONF-2014-004, ATLAS ttbar dilepton events PDF calibration

p_T interval [GeV]	20-30	30-40	40-50	50-60	60-75	75-90	90-110	110-140	140-200	200-300
SF	0.968	0.979	0.986	0.985	0.971	0.980	0.965	1.000	0.989	1.008
Total error [%]	6.5	3.4	2.8	2.5	2.3	1.8	2.1	2.2	3.6	8.4
Stat. error [%]	2.3	1.2	1.0	1.0	0.9	1.0	1.0	1.0	1.4	3.2
Syst. error [%]	6.1	3.1	2.7	2.3	2.1	1.5	1.9	2.0	3.3	7.6
Systematic Uncertainties [%]										
Hadronisation ($t\bar{t}$)	1.0	0.6	1.5	1.4	1.1	0.5	0.8	0.3	1.0	2.0
Modelling ($t\bar{t}$)	1.1	0.4	1.0	1.1	1.0	0.5	0.7	0.9	0.7	1.7
Top p_T reweighting ($t\bar{t}$)	0.2	0.3	0.3	0.2	0.2	0.1	0.1	0.4	1.4	4.6
More/less PS ($t\bar{t}$)	0.5	0.6	0.9	0.8	0.9	1.0	0.9	0.8	1.4	1.9
More/less PS (single top)	0.2	0.0	0.1	0.1	0.2	0.1	0.2	0.2	0.0	0.0
Modelling (Z+jets)	0.8	0.3	0.2	0.5	0.3	0.2	0.3	0.3	0.9	2.4
Modelling (dibosons)	0.7	0.7	0.6	0.6	0.6	0.6	0.7	0.8	1.3	3.1
Norm. single top	0.5	0.4	0.3	0.2	0.2	0.2	0.2	0.2	0.3	0.0
Norm. Z+jet	0.9	0.6	0.9	0.4	0.7	0.5	0.6	0.7	1.1	1.7
Norm. Z+b/c	0.1	0.1	0.2	0.1	0.1	0.0	0.1	0.0	0.1	0.2
Norm. lepton fakes	0.3	0.3	0.2	0.3	0.2	0.2	0.3	0.3	0.3	0.4
Pile-up reweighting	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.1
Electron eff./res./scale	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Muon eff./res./scale	0.2	0.3	0.2	0.3	0.1	0.2	0.1	0.1	0.0	0.0
E_T^{miss} soft-terms	0.1	0.1	0.2	0.0	0.2	0.0	0.1	0.2	0.3	0.5
Jet energy scale	4.1	2.2	1.2	0.7	0.7	0.3	0.7	0.8	1.2	2.6
Jet energy resolution	2.6	1.0	0.3	0.3	0.1	0.2	0.2	0.0	0.2	0.2
Jet vertex fraction	0.8	0.1	0.0	0.1	0.1	0.1	0.2	0.2	0.2	0.2
Mis-tag rate	2.8	1.1	0.5	0.4	0.3	0.2	0.3	0.4	0.5	1.1

Run1 uncertainties

source: presentation by L. Scodellaro, M. z. Nedden and J. M. Keaveney at the TOP LHC WG meeting open session, 21-23 May 2014

Source	Treatment in ATLAS	Treatment in CMS	Corr.
b/c production	b,c → gg scale by 50%	b,c → gg scale by 50%	Yes
B decay	Reweight according to BR	Neglected (small)	No
b quark frag.	Average B hadron energy fraction varied ±5%	Average B hadron energy fraction varied ±5%	Yes
c2l ratio	c/l ratio scaled by factor 2	l/c ratio scaled by ±20%	Yes
Muon p_T spectrum	p_T spectrum reweighting	Vary cut on muon p_T	Yes
$t\bar{t}$ generator	Compare MC@NLO to POWHEG (with Herwig)	Compare fit to templates from QCD events	No
Parton Showering	Compare Herwig to Pythia	Compare Herwig to Pythia	Yes
ISR and FSR	Using AcerMC+Pythia	Varying Q^2 scale and ME-PS threshold	Yes
Underlying events	Neglegible	Varying parameters	No