Search for Higgs boson pair-production in the $bb\tau\tau$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV data with the ATLAS detector

Zhiyuan Li

February 10, 2021

Contents

1 Introduction

2 Theory and Motivation

2.1 The Standard Model and the Higgs boson

2.2 Beyond the Standard Model

3 The ATLAS experiment at the Large Hadron Collider

3.1 The Large Hadron Collider

3.2 TThe ATLAS Detector

4 Data and Monte Carlo samples

5 Physics Objects Reconstruction

5.1 Track and vertex

5.2 Electron

5.3 Muon

5.4 Jet

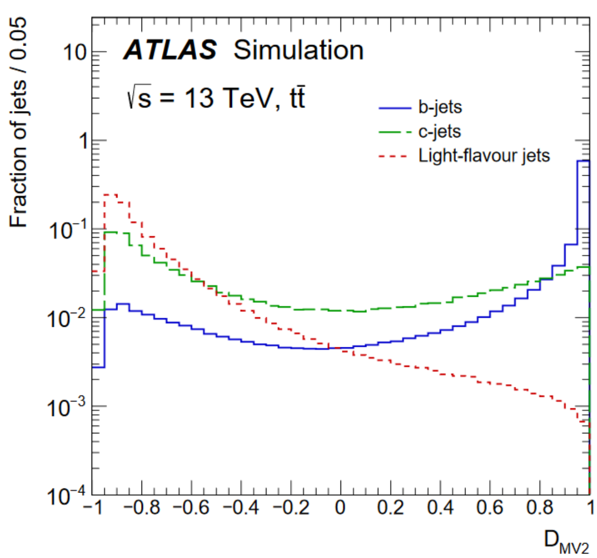
5.5 b-Jets

5.5.1 Flavour tagging

The identification of jets containing b-hadrons (b-jets) against the large background of jets containing c-hadrons (c-jets) or coming from the hadronization of light (u,d,s) quarks or gluons is of major importance in many areas of the physics programme of the ATLAS experiment at the LHC. It is crucial in a large number of Standard Model (SM) precision measurements, studies of the Higgs boson properties, and searches for new phenomena [1, 2, 3]. It also plays an important role in the di-Higgs to $bb\tau\tau$ searches presenting in Chapter 7.

The ATLAS Collaboration uses various algorithms to identify b-jets [4], referred to as b-tagging algorithms, when analysing data recorded during Run 2 of the LHC. These algorithms exploit the long lifetime, high mass and high decay multiplicity of b-hadrons, as well as the properties of the b-quark fragmentation. Given a lifetime of the order of 1.5 ps,

b-hadrons have a significant mean flight length ($\langle c\tau \rangle \approx 450\ \mu m$), in the detector before decaying, generally leading to at least one vertex displaced from the hard-scatter collision point. The strategy developed by the ATLAS Collaboration is based on a two-stage approach. Firstly, low-level algorithms reconstruct the characteristic features of the b-jets via two complementary approaches, one that uses the individual properties of charged-particle tracks, later referred to as tracks, associated with a hadronic jet, and a second which combines the tracks to explicitly reconstruct displaced vertices. These algorithms, first introduced during Run 1 [4], have been improved and retuned for Run 2 [5]. Secondly, in order to maximise the b-tagging performance, the results of the low-level b-tagging algorithms are combined into high-level algorithms via multivariate classifiers. The performance of a b-tagging algorithm is characterised by the probability of tagging a b-jet (b-jet tagging efficiency, ϵ_b) and the probability of mistakenly identifying a c-jet or a light-flavour jet as a b-jet, labelled $\epsilon_c(\epsilon_l)$. The most performant algorithms presently in use in physics analyses at ATLAS are based on multivariate combinations of the available information(MV2) or addditionally using a deep feed-forward neural network(DL1) [6, 7], as shown in Figure 1. Depending on the low-level algorithm, the DL1 tagger can be further separated into two taggers: DL1 and DL1r, where the DL1 tagger uses traditional track-based impact parameter taggers IP2D and IP3D and DL1r tagger uses a Recurrent Neural Network tagger (RNNIP) [7]. The DL1r tagger is now the default b-tagging algorithm used for flavour tagging in ATLAS. The performance of the algorithms is quantified in terms of c-jet (light-jet) rejections, defined as $1/\epsilon_c$ and $1/\epsilon_l$.



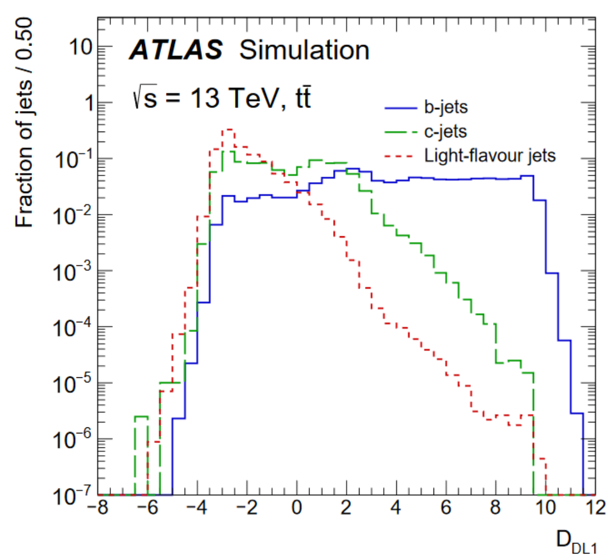


Figure 1: Performance of MV2c10(left) and DL1(right) tagger [5].

5.6 Missing transverse energy

5.7 Hadronically decaying τ lepton

6 Charm-jet mis-tagging calibration

6.1 Calibration methods for b-jet and light-jet

MC simulations are not able to model very well the performance of the b-tagging algorithms in data. For this reason calibration is required, i.e. correcting MC to recover the data in terms of b-tagging efficiency, charm-jet mis-tagging and light-jet mis-tagging rates [5]. The calibration is performed for all supported working points, which are cuts in the b-tagging algorithm output identifying the tagging efficiencies, and jet collections(TODO: refer back to the object definition chapter). In general, the efficiency is calculated with data and simulations, and scale factors are then calculated to match the efficiency extracted from simulations to the data. The production of $t\bar{t}$ pairs at the LHC provides an abundant source of b-jets by virtue of the high cross-section and the $t \to Wb$ branching ratio being close to 100%. A very pure sample of $t\bar{t}$ events can be selected by requiring that both W bosons decay leptonically, referred to as di-leptonic $t\bar{t}$ decays in the following. For the b-jet calibration, the performance of the b tagging algorithms is evaluated in the simulation and the efficiency with which these algorithms identify jets containing b-hadrons is measured in collision data. The measurement uses a likelihood-based method in the di-leptonic $t\bar{t}$ sample, where events with exactly 2 jets and 2 opposite signs leptons are selected. The data b-jet efficiency is then extracted from a combined likelihood fit, and subsequently compared with that predicted by the simulation. Scale factors are then calculated to match the performance of the algorithms to the data [5].

For the light-jet mis-tagging calibration, two methods are used to measure the mis-tagging rate from the data. The first is the negative tag method, which uses a high statistics data sample enriched in light-jets with the application of a modified algorithm which reverses some of the criteria used in the nominal identification algorithm. The second is the adjusted Monte Carlo (adjusted-MC) method, which adjusts the characteristic track observables in the simulation to match the data, and then compares the adjusted simulation to the "standard" simulation. The scale factors are then calculated using the these two methods. The scale factors of the two different methods are in good agreement within the systematics uncertainties [8].

6.2 Calibration method for charm-jet

It is worth mentioning that the author's qualification task to become an ATLAS author is to calibrate the rate of a charm jet being mis-identified as a b-jet, which is a part of the calibration of the b-tagging algorithm. During the task the calibration range has been extended down to 20 GeV (previously 25 GeV) of jet p_{T} and a new selection category has been developed to increase the data statistics of the scale factors in the high-p_T ($p_T^{jet} >$ 70 GeV) region.

As determined by the CKM matrix [9, 10], the W boson decay dominantly to a pair of light-quark (u-quark and d-quark) or to a s-quark and a c quark. The W boson decays very rarely to pairs containing a b-quark. More specifically, the branching ratio of a W boson decays to the u-quark and d quark pair or the s-quark and c quark is 33.1%, and to pairs containing a b quark is only 0.057% [11]. Therefore, b-tagged jets from the W decay are most likely to be mis-tagged c-jets or light-jets. Furthermore, given the ratio between the DL1 light-jet rejection and the corresponding charm-jet rejection ranges from 10 to 40 (see ref [7]), the c-jet is much more likely to be mis-tagged than the light-jet. This allows for a source of mistagged c-jets to be obtained in the $t\bar{t}$ events, requiring that one W boson decay leptonically and the other decay hadronically (referred to as semi-leptonic $t\bar{t}$ decay in the following), where the b-tagged jets from the W decay are candidates of miss-tagged c-jet. Requiring a W boson decaying leptonically reduces the number of combinations of jets of different flavour, and allows triggering with the lepton.

The events kinematics are shown by the diagram in Figure 2, where the $t\bar{t}$ pair decays to a b- and a $\bar{b}$- quark, circled in red. One of the W bosons, circled in blue, decays hadronically to quarks, and the other W boson decays leptonically to either an electron or a muon and the corresponding neutrinos, circled in green and purple, respectively. The lepton in the final state is used for triggering, and a combined likelihood fit is used to extract the c mis-tagging efficiency.

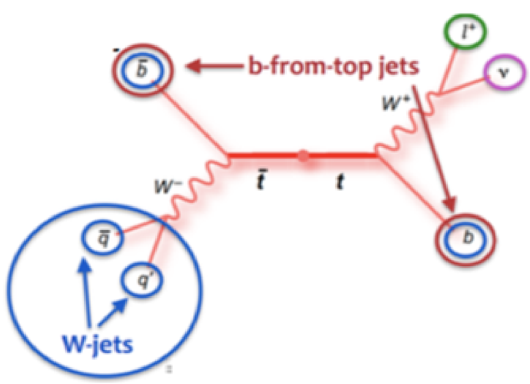


Figure 2: Feynman diagram of the semi-leptonic $t\bar{t}$ events.

A kinematic likelihood technique, referred to as KLFitter [12], is used to assign jets to the proper $t\bar{t}$ decay product (more details in Section 6.4). The charm-jet efficiency is then extracted by applying a combinatorial likelihood fit to the pair of jets from the hadronic W decay, where the main floating parameter is the c-jet efficiency. The calibration is given as scale factors in bins of p_{T} for 4 fixed-cut working points (WP) that scale the simulation shape to reproduce that of the data.

6.3 Data and Monte Carlo samples

The data analysed in this study correspond to 139 fb^{-1} [13, 14, 15, 16], of pp collision data collected by the ATLAS detector between 2015 and 2018 with a centre-of-mass energy of 13 TeV. The data sample was collected using a set of single-muon [17] and single-electron triggers [18]. The single-muon triggers had p_{T} thresholds in the range 20–26 GeV for isolated muons and 50 GeV for muons without any isolation requirement. The single-electron triggers employed a range of p_{T} thresholds varying between 24–300 GeV and a combination of quality and isolation requirements depending on the data-taking period and the p_{T} threshold. All detector subsystems were required to be operational during data taking and to fulfil data quality requirements.

TODO: remove the overlap between this section and the Data MC chapter in the thesis Dedicated MC are used to model SM processes. All samples were produced using the ATLAS simulation infrastructure [19] and Geant4 [20]. A subset of samples use a faster simulation based on a parameterisation of the calorimeter response and Geant4 for the other detector systems [19]. The simulated events are reconstructed with the same algorithms as used for data, and contain a realistic modelling of pile-up interactions. The pile-up profiles in the simulation match those of each dataset between 2015 and 2018, and are obtained by overlaying minimum-bias events, simulated using the soft QCD processes of Pythia 8 [21] using the NNPDF2.3LO set of PDFs [22] and a set of tuned parameters called the A3 tune [23].

The events that are used in this study originate mostly due to $t\bar{t}$ production. This process is modelled using the PowhegBox v2 [24, 25, 26, 27] generator at NLO with the `NNPDF3.0nlo` parton distribution function (PDF) set and the h_{damp} parameter[1] set to 1.5 m_{top} [28]. The events were interfaced to Pythia 8.230 to model the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune and using the `NNPDF2.3lo` set of PDFs. The decays of bottom and charm hadrons were performed by EvtGen v1.6.0 [29]. The simulated $t\bar{t}$ events are split according to the origin of W-jets. The notation "$t\bar{t}$, ll" denotes that both W-jets are light flavour jets. Similarly, "$t\bar{t}$, cl" ("$t\bar{t}$,

[1]The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of Powheg matrix elements to the parton shower and thus effectively regulates the high-p_{T} radiation against which the $t\bar{t}$ system recoils.

bl") indicates that one of the W-jets is a c-jet (b-jet) whereas the other is a light flavour jet. W-jets with origin other than what is discussed above fall into the category denoted by "$t\bar{t}$, other". This category includes events in which at least one of the W-jets comes from a hadronically decaying τ-lepton.

In addition to $t\bar{t}$ production, there are some minor backgrounds that contribute to the final event sample that is used for the calibration. These backgrounds consist mostly of single-top and diboson production, the production of $t\bar{t}$ in association with a vector boson and the production of a vector boson in association with jets. The details of the modeling of these samples are given in the following.

Single-top s-channel production is modelled using the POWHEGBOX v2 generator at NLO in QCD in the five-flavour scheme with the `NNPDF3.0nlo` [30] parton distribution function (PDF) set. The associated production of top quarks with W bosons (tW) is modelled using the POWHEGBOX v2 [31, 25, 26, 27] generator at NLO in QCD using the five-flavour scheme and the `NNPDF3.0nlo` set of PDFs [30]. The diagram removal scheme [32] is used to remove interference and overlap with $t\bar{t}$ production. The events for both single-top s-channel and tW production are interfaced to PYTHIA8.230using the A14 tuneand the `NNPDF2.3lo` set of PDFs.

The production of Z+jets and W+jets is simulated with the SHERPA v2.2.1 [33] generator using next-to-leading order (NLO) matrix elements (ME) for up to two partons, and leading order (LO) matrix elements for up to four partons calculated with the Comix [34] and OPENLOOPS [35, 36, 37] libraries. They are matched with the SHERPA parton shower [38] using the MEPS@NLO prescription [39, 40, 41, 42] using the set of tuned parameters developed by the SHERPA authors. The `NNPDF3.0nnlo` set of PDFs [30] is used and the samples are normalised to a next-to-next-to-leading order (NNLO) prediction [43].

Samples of diboson final states (VV) are simulated with the SHERPA v2.2.1 or v2.2.2 [33] generator depending on the process, including off-shell effects and Higgs-boson contributions, where appropriate. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, are generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Samples for the loop-induced processes $gg \to VV$ are generated using LO-accurate matrix elements for up to one additional parton emission for both cases of fully leptonic and semileptonic final states. The matrix element calculations are matched and merged with the SHERPA parton shower based on Catani-Seymour dipole factorisation [34, 38] using the MEPS@NLO prescription [39, 40, 41, 42]. The virtual QCD correction are provided by the OPENLOOPS library [35, 36, 37]. The `NNPDF3.0nnlo` set of PDFs is used, along with the dedicated set of tuned parton-shower parameters developed by the SHERPA authors.

The production of $t\bar{t}$ in assosiation with a vector boson is modelled using the MAD-

GRAPH5_aMC@NLO v2.3.3 [44] generator at NLO with the `NNPDF3.0nlo` [30] parton distribution function (PDF). The events are interfaced to PYTHIA8.210 [21] using the A14 tune [45] and the `NNPDF2.3lo` [30] PDF set. The decays of bottom and charm hadrons are simulated using the EVTGEN v1.2.0 program [46].

6.4 Kinematic Likelihood Fitter

The four-vectors of the four highest p_{T} jets, the lepton and the event $E_{\mathrm{T}}^{\mathrm{miss}}$ are used as inputs to a likelihood-based $t\bar{t}$ event reconstruction algorithm, which is described in more detail in Ref. [12]. This algorithm uses a likelihood function to assign the four jets to the $t\bar{t}$ decay topology. In particular, the algorithm assigns one jet to be the b-jet from the leptonically decaying top-quark ($t \to Wb \to \ell\nu b$), another to the b-jet from the hadronically decaying top-quark ($t \to Wb \to qq'b$, where qq' are the quarks in which the W boson decays) and the remaining two jets to the jets that come from the hadronic W boson decay. The jet assignment does not use any b-tagging information. The following notation will be used: the jets that are assigned as the decay products of the W boson are referred to as W-jets and the remaining two jets are referred to as top-jets.

6.5 Maximising likelihood

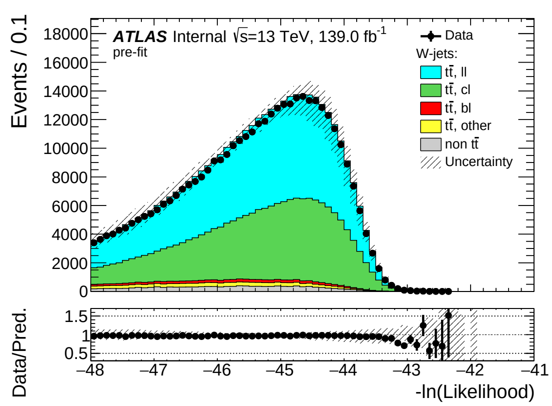


Figure 3: Distribution of the negative logarithm of the likelihood that is used to reconstruct the $t\bar{t}$ decay.

Taking only four jets in the event limits the total number of possible jet orderings (permutations) in the event. In the semi-leptonic channel, four jets can be permuted a total number of times equal to $4! = 24$. However, the two W-jets are kinematically indistinguishable. This reduces the possible number of permutations to 12. Furthermore, no b-tagging information is used in the kinematic likelihood to limit the possible number of permutations as this would bias the result. For every combination of jet ordering, the likelihood is maximised over its free parameters, the energy of the four jets, the lepton energy and the three components of the momentum of the neutrino, and provides a value based on how closely the kinematic information from the reconstructed objects for a specific jet ordering resembles the expected kinematic behaviour of the decay of a Standard Model semi-leptonic $t\bar{t}$ event. The likelihood therefore distinguishes the possible permutations on an event-by-event basis. The best permutation, given by the largest log-likelihood value, is adopted as the jet ordering for the event. An additional requirement of log-likelihood $>$ -48 is placed on the output of the likelihood value for the chosen event permutation. An example of the distribution of log-likelihood of the best permutations is shown in Figure 3. In this figure, the data events are compared against the simulation. The majority of the events come from $t\bar{t}$ production. There is only a very small fraction of events, which is denoted as "non $t\bar{t}$" on the figure, that come from other processes like W or Z production in association with jets or single-top production.

6.6 Event selection

6.6.1 Standard selection

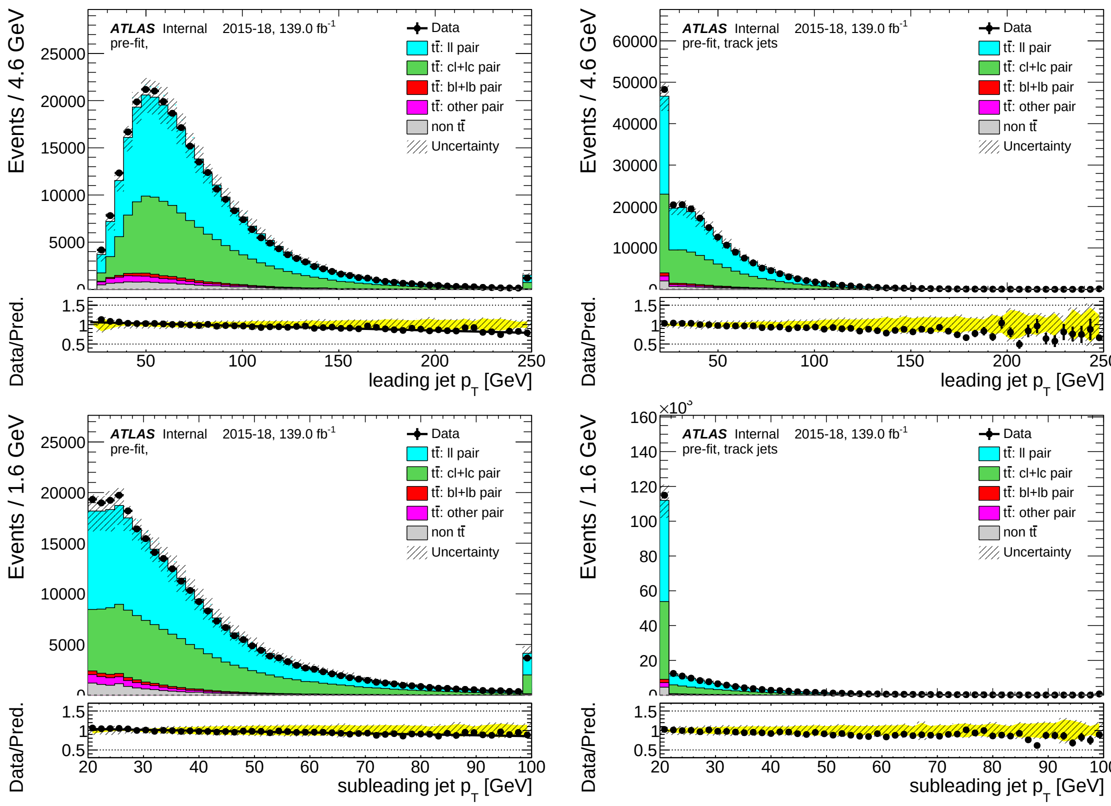


Figure 4: Standard selection: data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column, inclusive of the low-p_{T} selection. The leading jet and sub-leading jet refer to the highest p_{T} W-jet and the second highest p_{T} jet respectively. The 'non $t\bar{t}$' background indicates background comes from non-$t\bar{t}$ processes like W or Z production in association with jets or single-top production.

Events are required to contain exactly one trigger-matched lepton with $p_T > 27$ GeV and exactly four jets with $p_T > 25$ GeV. Leptons are required to have p_T above 27 GeV in order to avoid the turn-on curve for the single lepton triggers. Events which contain an additional lepton with $p_T > 27$ GeV are rejected. The events are also required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV, which is assumed to be the result of the neutrino from the leptonically decaying W boson. The transverse mass m_T between the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, is constrained as follows:

$$m_T = \sqrt{2p_T^{\ell} E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi)} > 40 \text{ GeV},$$

where $\Delta\phi = \phi(E_{\mathrm{T}}^{\mathrm{miss}}) - \phi(\ell)$ is the azimuthal difference between the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$. An example of the p_{T} distributions before any tagging or fitting and after the standard selection is shown in Figure 4. More plots can be found in Appendix A.1.1, Figure 15. The yellow band in the lower pad shows the overall systematics uncertainties combining the experimental uncertainties and the $t\bar{t}$ modelling uncertainties, as described in Section 6.7. The data/MC ratio shows good agreement within the systematic uncertainties.

6.6.2 Selection for low-p_T extension

The author has developed an othorgonal selection to extend the calibration in the low-p_T region so that the calibration can be applied to jets with $p_{\mathrm{T}} < 20$ GeV. Instead of requiring events to have exactly 4 jets $p_{\mathrm{T}} > 25$ GeV, events are required to have exactly 3 jets with $p_T > 25$ GeV and exactly 1 jet with 25 GeV $> p_T >$ 20 GeV. Other than that, all other requirements for the selection are the same. This additional cut provides candidates for the W-jet that is used for calibration in the $20 - 25$ GeV region. The inclusive yields of the standard selection and the low-p_{T} selection of the data/MC are given in Table 1. The distributions of the sub-leading jet are shown in Appendix A.1.2, Figure 16. Good agreement between MC and data is shown in these distributions, and the p_T range of the sub-leading has gone down to 20 GeV.

	Particle flow jets	Track jets
Data	287105	218351
$t\bar{t}$	292200 ± 200	223770 ± 180
Non $t\bar{t}$	10950 ± 120	7280 ± 100
Data/MC	0.947 ± 0.002	0.945 ± 0.002

Table 1: Standard selection: prefit comparison of the number of events in data and in simulation considering particle flow jets and track jets for events with exactly 4 jets, inclusive of the low-p_{T} selection.

6.6.3 High-p_T selection

It has been observed that in the previous calibrations, the statistics is relatively low for the high-p_{T} region (e.g. jet $p_{\mathrm{T}} > 100$ GeV). Therefore, the author has worked on an othorgonal selection to improve this situation. Instead of requiring events to have exactly 4 jets, events are required to have at least 5 jets with $p_T > 25$ GeV, in which at least 1 jet with $p_{\mathrm{T}} > 70$ GeV. Other than that, all other requirements for the selection remaining the same. The choice of cut value 70 GeV is based on the study shown in the following. The effect on the c-jet purity and the potential statistical gain is investigated, where the c-jet purity is defined as:

$$c\text{-jet purity} = \frac{\mathrm{N}_{\mathrm{true}\ c\text{-jet}}}{\mathrm{N}_{\mathrm{all}}}, \tag{1}$$

where $N_{\mathrm{true}\ c\text{-jet}}$ stands for the number of events with a true c-jet from the W decay, and N_{all} stands for the number of all events. The ideal situation is the high-p_{T} selection will maximally increase the statistics while mininally decrease the c-jet purity, therefore a figure of merit P^{Cut} is defined as:

$$P^{\mathrm{Cut}} = \frac{\sum_i \mathrm{Gaininstats}_{\mathrm{i}}^2}{\sum_i c\text{-jet purity}_{\mathrm{i}}^2},$$

where i stands for the number of bins, and gain in stats stands for increase in statistics. The c-jet purity and the statistical gain are calculated for 4 different cut values as shown in Figure 5, comparing with the cut value of 0. The value of 70 GeV is chosen as it gives the highest value of P^{Cut}.

	Particle flow jets	Track jets
Data	98273	120929
$t\bar{t}$	99430 $\pm$ 120	117090 $\pm$ 130
Non $t\bar{t}$	1842 $\pm$ 21	2640 $\pm$ 50
Data/MC	0.97 $\pm$ 0.003	1.01 $\pm$ 0.003

Table 2: High-p_{T} selection: prefit comparison of the number of events in data and in simulation considering particle flow jets and track jets.

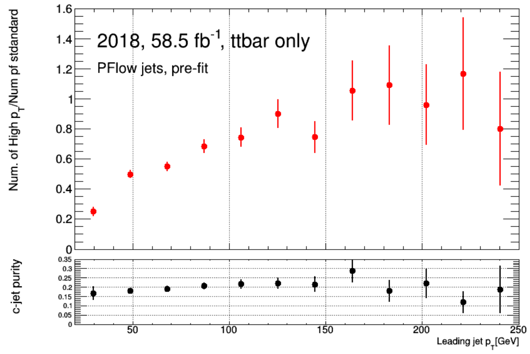
The yields of the data/MC are given in Table 2. An example of the p_{T} distributions before any tagging or fitting applying the high-p_{T} selection is shown in Figure 6. More plots can be found in Appendix A.2, Figure 17.

6.6.4 Combined selection

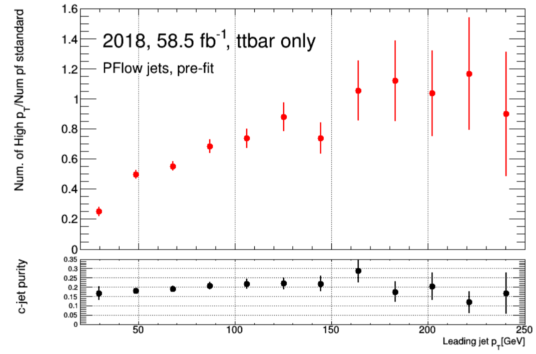
As the standard selections, low-p_{T} selection and high-p_{T} selection are othorgonal to each other, all the selections are combined to provide the maximum range and statistics for the calibration. The yields of the data/MC are given in Table 3, an example of the p_{T} distributions before any tagging or fitting and after the standard selection is shown in Figure 7. More plots can be found in Appendix A.2.1, Figures 18-23.

6.7 Systematic uncertainties

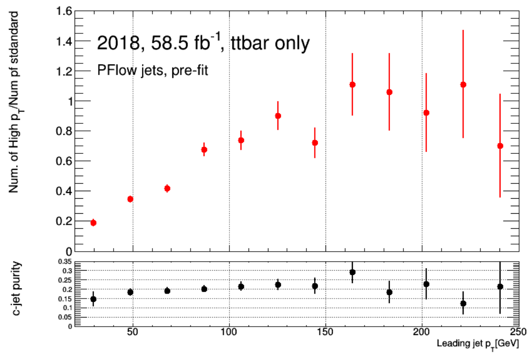
The systematic uncertainties considered and propagated in this calibration can be broadly categorised into experimental and modelling systematics uncertainties.



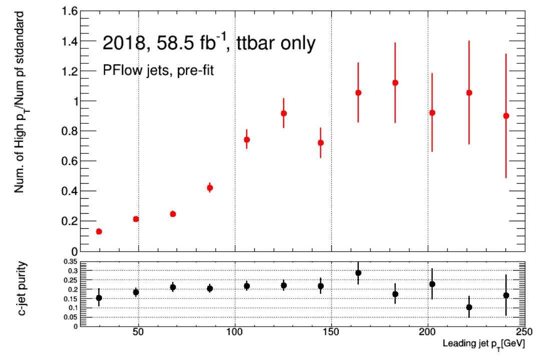
(a) No cut.



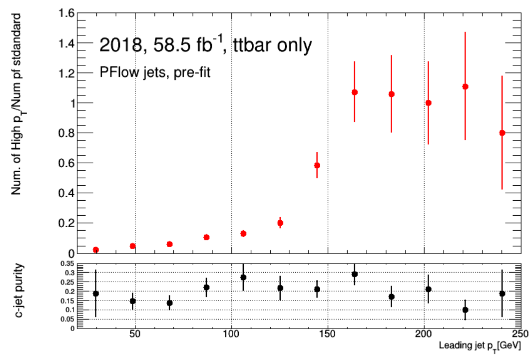
(b) p_{T} > 40 GeV.



(c) p_{T} > 70 GeV.



(c) p_{T} > 90 GeV.



(d) p_{T} > 140 GeV.

Figure 5: Comparison of different cut values in terms of gain in stats and c-jet purity.

6.7.1 Experimental uncertainties

Experimental uncertainties are related to the detector and estimated using data-driven methos or MC simulations. The electron energy scale and resolution are corrected to provide better agreement between MC predictions and data, uncertainties due the corrections are considered. Uncertainties are taken into account on the electron and muon trigger, identification and reconstruction efficiencies, and for uncertainties associated with the isolation requirements. Scaling and smearing corrections are applied to the p_T of simulated electrons and muons in order to minimise the differences in resolution between data and MC

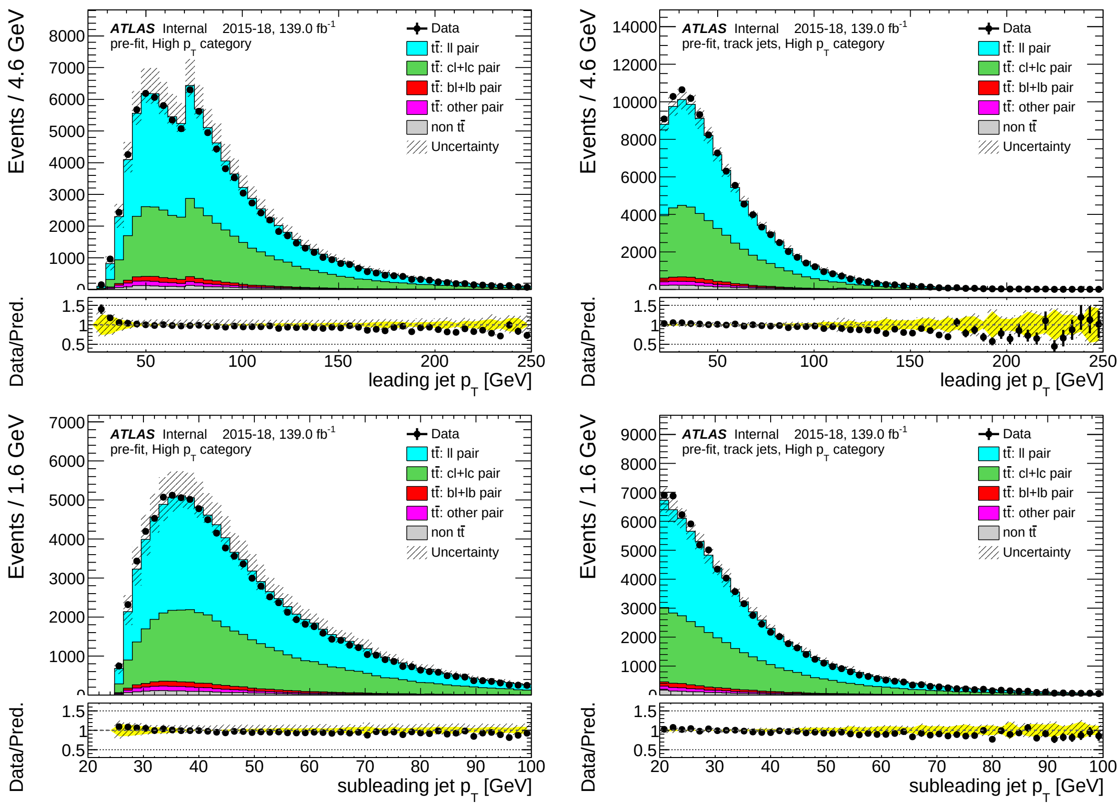


Figure 6: High-p_{T} selection: data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column.

events, and the uncertainties of the corrections are considered. The jet energy scale (JES) uncertainty depends on p_T and η and takes into account uncertainties due to pile-up effects. Uncertainties on the jet energy resolution (JER) is taken into account. Uncertainties on the energy scale and resolution of the electrons, muons, jets and taus are propagated to the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$, which also has additional dedicated uncertainties on the scale, resolution, and reconstruction efficiency of tracks not associated to any of the reconstructed objects, along with the modelling of the underlying event. Uncertainties on the b-tagging probabilities for b- and light-jets are considered both for the tagging jets assigned to the b-quark from the top decay and for the jets associated to the hadronically decaying W boson. Supporting material for this section can be found in the appendix, Tab.6.

6.7.2 Modelling uncertainties

Uncertainties on the modelling of the inclusive $t\bar{t}$ background are estimated by replacing the nominal MC sample by alternative MC samples. The nominal sample is also replaced by variations of the parton shower, initial and final state radiation. In all cases, MC-to-MC SFs are taken into account. The dominating modelling uncertainty is due to the choice

	Particle flow jets	Track jets
Data	385378	302308
$t\bar{t}$	383520 ± 230	302690 ± 200
Non $t\bar{t}$	12420 ± 120	8570 ± 100
Data/MC	0.973 ± 0.002	0.971 ± 0.002

Table 3: Combined selection: prefit comparison of the number of events in data and in simulation considering particle flow jets and track jets for an inclusive selection.

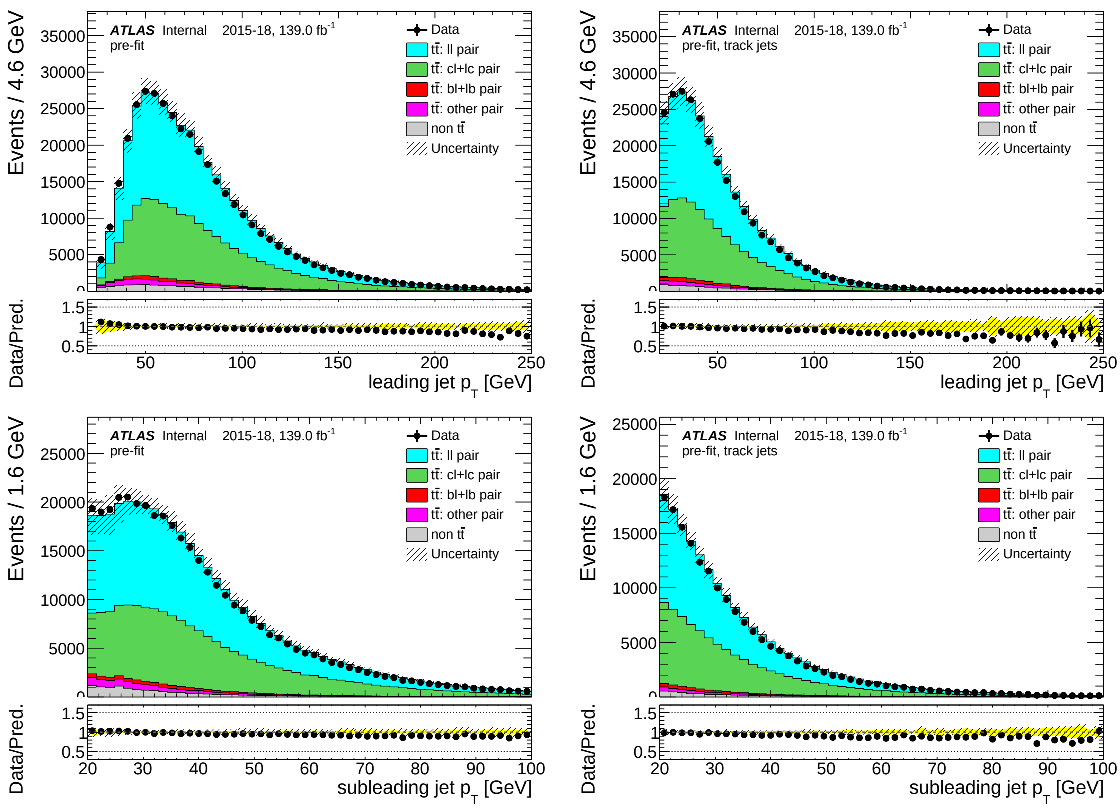


Figure 7: Combined selection: data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column.

of parton shower and hadronisation model, which is derived by comparing the prediction from POWHEG interfaced either to PYTHIA8 or Herwig++.The uncertainty on modelling of initial state radiation (ISR) is assessed with two alternative POWHEG+PYTHIA8 samples. The samples include one with an increase in radiation which has the re-normalisation and factorisation scales decreased by a factor of two and the *hdamp* parameter doubled, while the sample with a decrease in radiation has the scales increased by a factor of two. The comparisons of the nominal $t\bar{t}$ sample and the samples with each systematic uncertainty are shown in Table 4.

	Particle flow jets		Track jets	
	Yields	Ratio of difference to nominal sample	Yields	Ratio of difference to nominal sample
$t\bar{t}$ Nominal	385378 ± 230		302690 ± 200	
Data/MC	0.973 ± 0.002		0.971 ± 0.002	
$t\bar{t}$ AF2	386260 ± 250	0.716%	304860 ± 230	0.716%
DATA/MC(AF2)	0.967 ± 0.002		0.965 ± 0.002	
$t\bar{t}$ ISR	377130 ± 220	-1.665%	297960 ± 200	-1.562%
DATA/MC(ISR)	0.989 ± 0.002		0.986 ± 0.002	
$t\bar{t}$ Herwig	331960 ± 220	-13.443%	259940 ± 190	-14.123%
DATA/MC(Herwig)	1.119 ± 0.002		1.126 ± 0.002	

Table 4: Comparison of the number of events in data and in simulation considering particle flow jets and track jets for an inclusive selection. The uncertainty due to the variations samples being produced by fast simulation is included in the table as $t\bar{t}$ AF2.

6.8 Under-estimation of $t\bar{t}$ + Heavy flavour background

Depsite the fact that the true nature of most of the reconstructed W-jets are either c-jets or light-jets, there is still a very small amount of them are true b-jets. There are two main sources of these true b-jets. The first is a W boson decays to a b- and a c-quark. The second is when the $t\bar{t}$ plus a gluon process (referred to as $t\bar{t}$ + heavy flavour process) is selected, and the gluon is split into a pair a b quarks and one of them is assigned as a W-jet. We can exclude the first source by requiring no c-jets in the W-jets, leaving the true b-jet in the W-jets can only come from the $t\bar{t}$ + heavy flavour process. This process is underestimated by the MC by about 30% for both the PFlow and VR-Track jets collections, as shown in Table 5 and Figure 8, where an extra cut requiring at least one W-jet with DL1r > 8 is added to the combined selection. This cut is to reject most of the true c-jets and true light-jet. For this reason, events in the simulation in which the top-jets and at least one of the W-jets are b-jets(referred to as 3 true b-jetsevents), are scaled by 1.25 ± 0.25, where the value is measured in Ref. [47]. All results shown in this chapter have this scale factor implemented, and the full difference between the simulation before applying this scale factor and after is taken as a systematic uncertainty. This uncertainty has been added in quadrature to the systematic uncertainties described in section 6.7 in all the plots in this chapter.

6.9 Results

6.9.1 Overview

Four rounds of calibrations have been carried out, containing different jet collections, Monte Carlo samples, analysis framework and b-jet identification algorithm. In the latest round,

	Particle flow jets	Track jets
Data	1589	1336
$t\bar{t}$	1100 ± 13	940 ± 12
Non $t\bar{t}$	83 ± 6	69 ± 5
Data/MC	1.34 ± 0.04	1.32 ± 0.04

Table 5: Yields of the 2018 data/MC of the combined selection, requiring at least 1 PFlow or track W-jet with DL1r > 8 to reject most of the light- and c-jets.

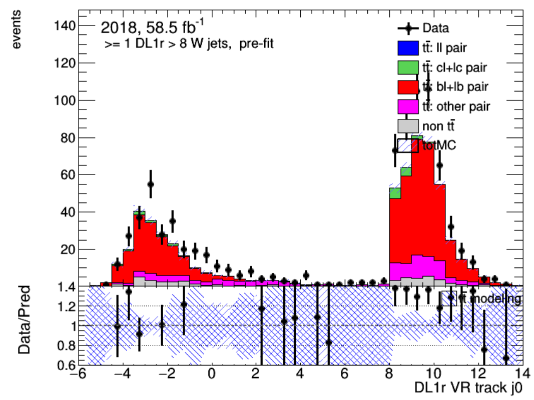


Figure 8: The DL1r score distribution of the leading VR-Track jet, requiring at least 1 VR-Track jets have DL1r > 8 to reject most of the light and the c jets, with $t\bar{t}$ modelling and statistical uncertainties.

the calibration includes the EMPFlow jet and VR-Track jet collection, and MV2c10, DL1 and DL1r taggers. The low-p_T selection and the standard selection are carried out for all four calibrations, while the high-p_T selection is only implemented in the latest calibration.

6.9.2 Kinematic distributions

The kinematic distributions of the MC and the data of the latest calibration (October 2020) are shown in Figure 9 for the PFlow jets and Figure 10 for the VR-Track jets, combining the standard selection, low p_{T} and the high-p_T selection. In these figure, the data events are compared against the simulation. The majority of the events come from $t\bar{t}$ production. There is only a very small fraction of non $t\bar{t}$ events. The W-jets pairs are mostly light-jets pairs and c-jet light-jet pairs, and a very small fraction of the pairs are b-jet light-jet pairs or pairs containing one or more τ hadron(s). The yellow band in the lower pad indicates the overall systematics uncertainties and the black band represents the $t\bar{t}$ modelling systematics uncertainty, which dominates at low b-tagging discriminant (DL1 or DL1r < 4). At high b-tagging discriminant (DL1 or DL1r > 4), the uncertainty due to the 1.25 ± 0.25 scale factor becomes more important.

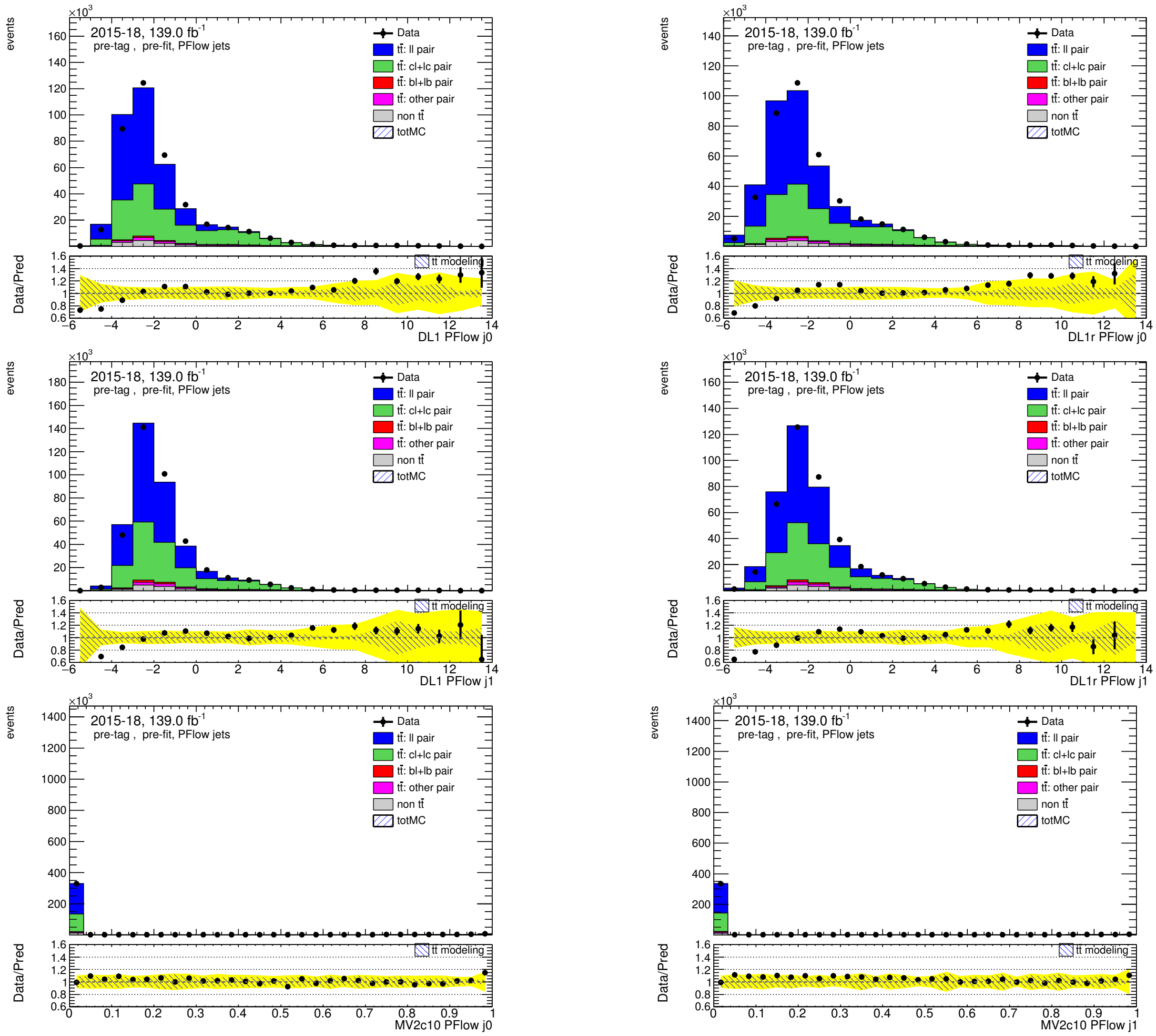


Figure 9: Distributions of the DL1, DL1r and MV2c10 taggers output of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

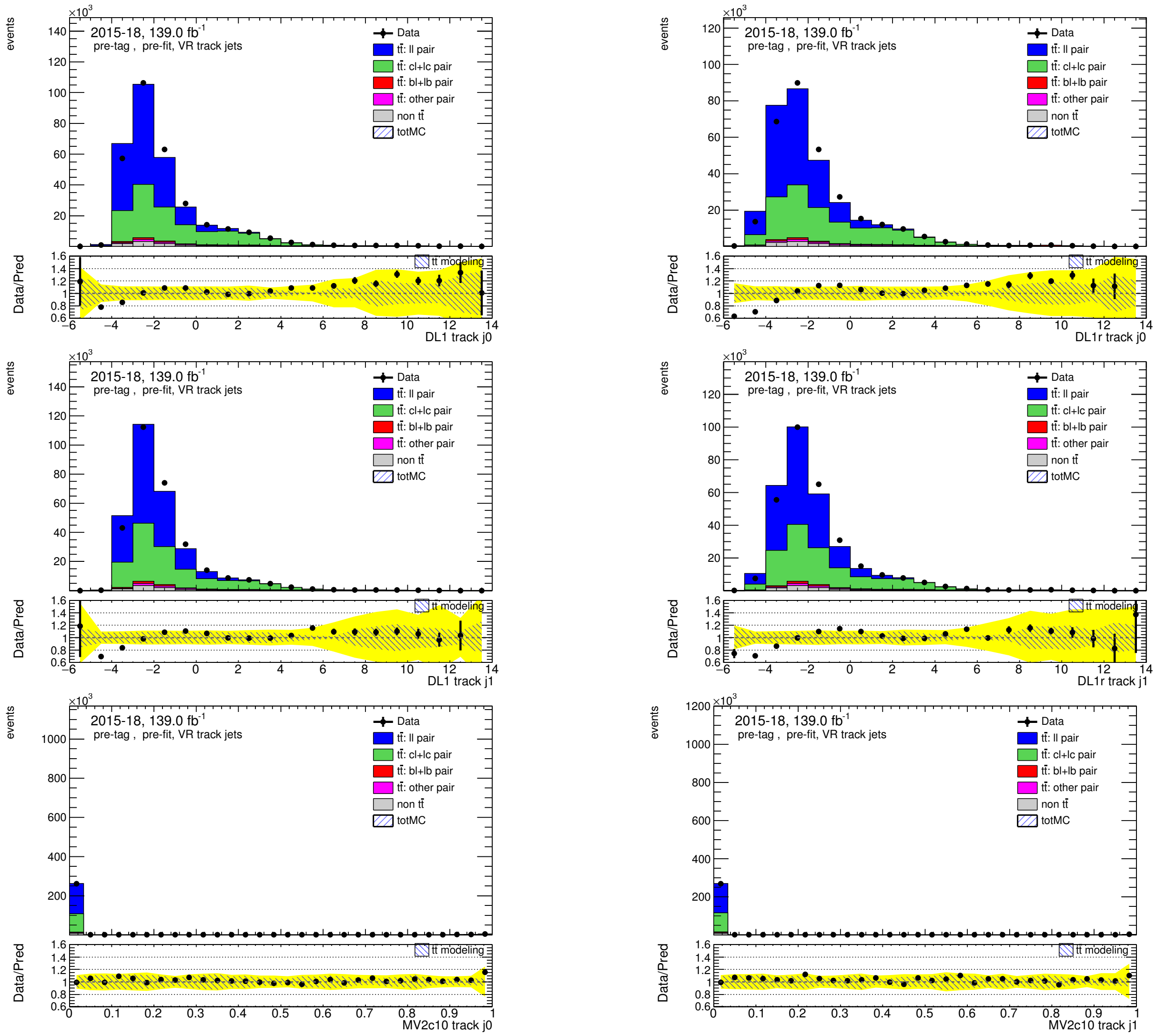


Figure 10: Distributions of the DL1, DL1r and MV2c10 taggers output of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

6.9.3 Efficiencies and Scale Factors

The Dl1 and DL1r c-jet efficiencies and scale factors with systematics uncertainties are calculated with 4 fixed cut working points: 60%, 70%, 77%, and 85% for the PFlow and VR track jets collection in the latest derivation in December 2020. The working point corresponds to the efficiency of a true b-jet passing the b-tagging algorithms for $t\bar{t}$ events. The efficiencies are shown in Figure 11 and 12 for the PFlow jet collections and the VR Track jets respectively. These results combine the standard selection and the high-p_T selection, and a 1.25 ± 0.25 scale factor is applied on events with 3 true b-jets. In the efficiencies plots, the black line represents the data efficiency and the red and green line represent DL1 and DL1r efficiencies respectively. The overall uncertainties are shown in the red and green band, where the systematics uncertainties dominate in the low-p_{T} bins ($p_{\mathrm{T}} < 150$) and the statistical error, represented by the error bars on the markers, become more visible in the last bin. The DL1 efficiencies are in good agreement with the data efficiency within the uncertainty band for most of the bins in all WPs, while the DL1r efficiencies show a bigger discrepancy to the data in general. The scale factors are calculated bin-by-bin as the ratio of the data efficiencies to the tagger efficiencies, as shown in 13, 14. The tighter working points (60%, 70%) show larger uncertainties and bigger deviation from 1, while the looser working points (77%, 85%) have much smaller uncertainty and the simulation is able to recover the data well.

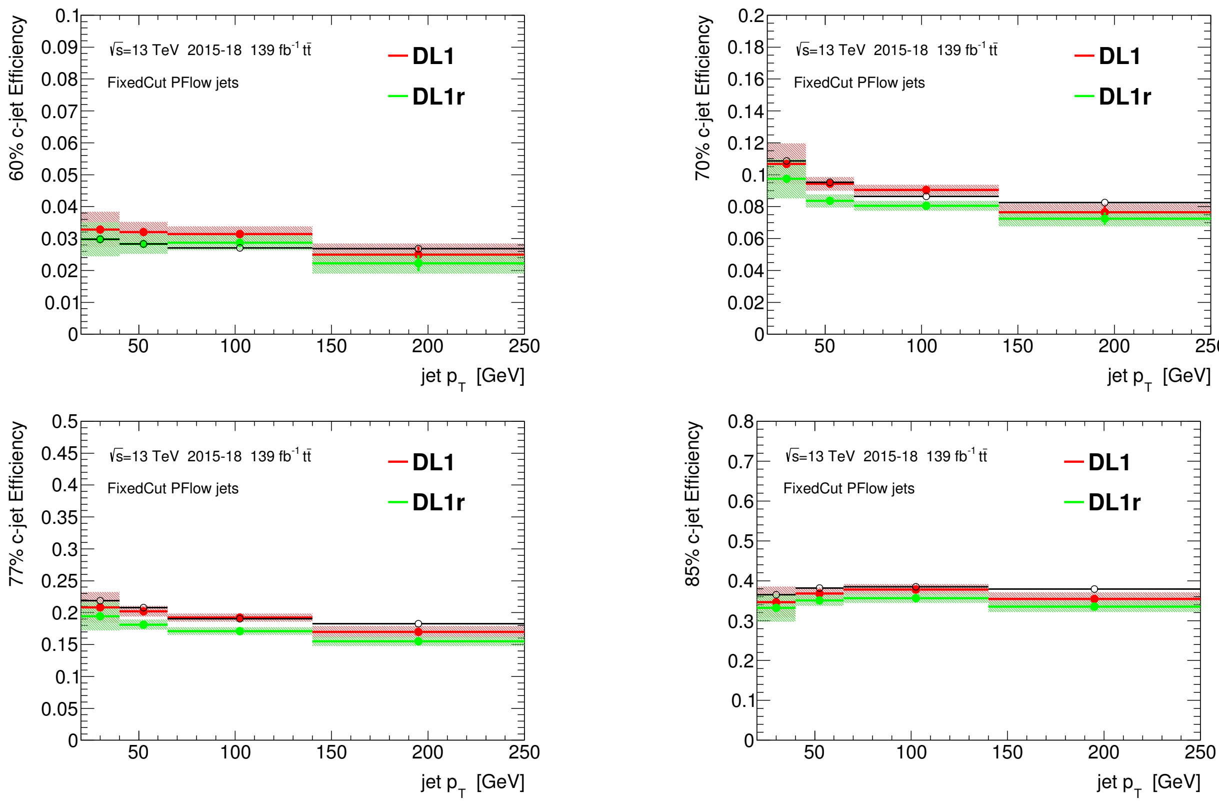


Figure 11: Charm-jet efficiencies of the PFlow jets collection given for 4 different working points.

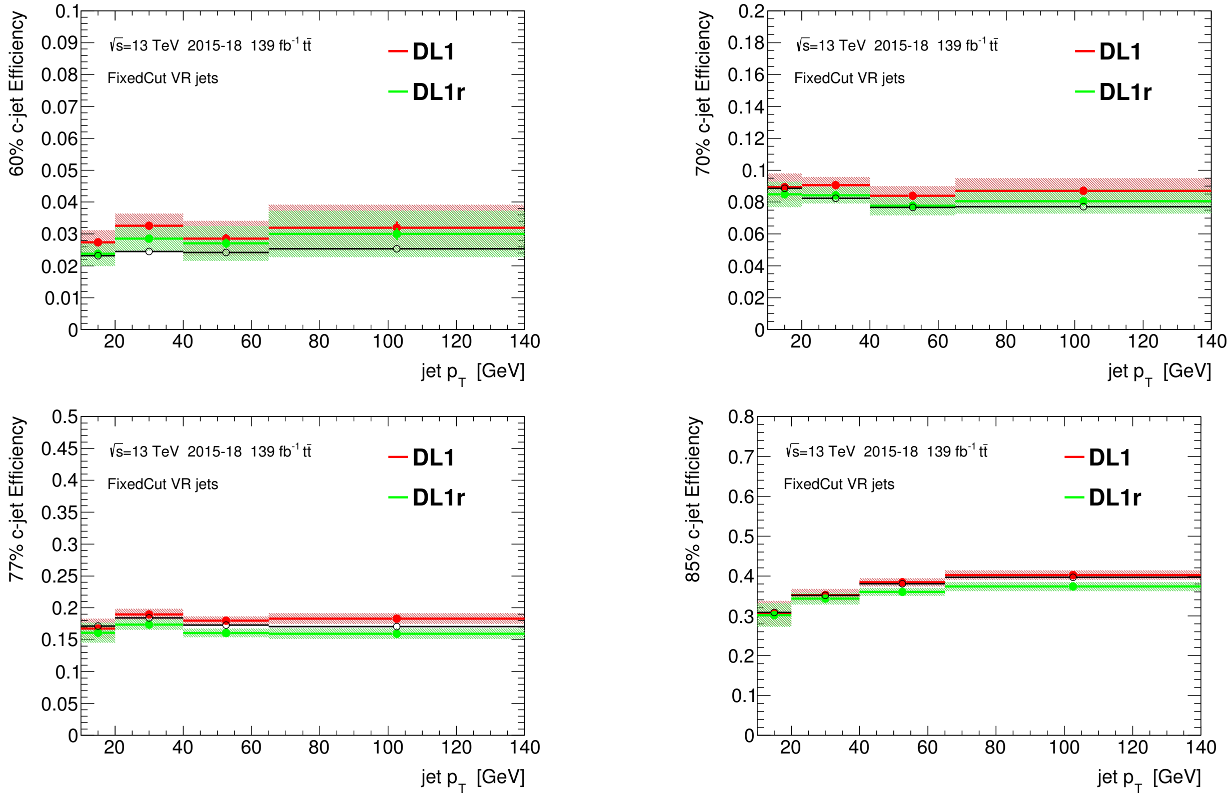


Figure 12: Charm-jet efficiencies of VR-Track jets collection of given for 4 different working points.

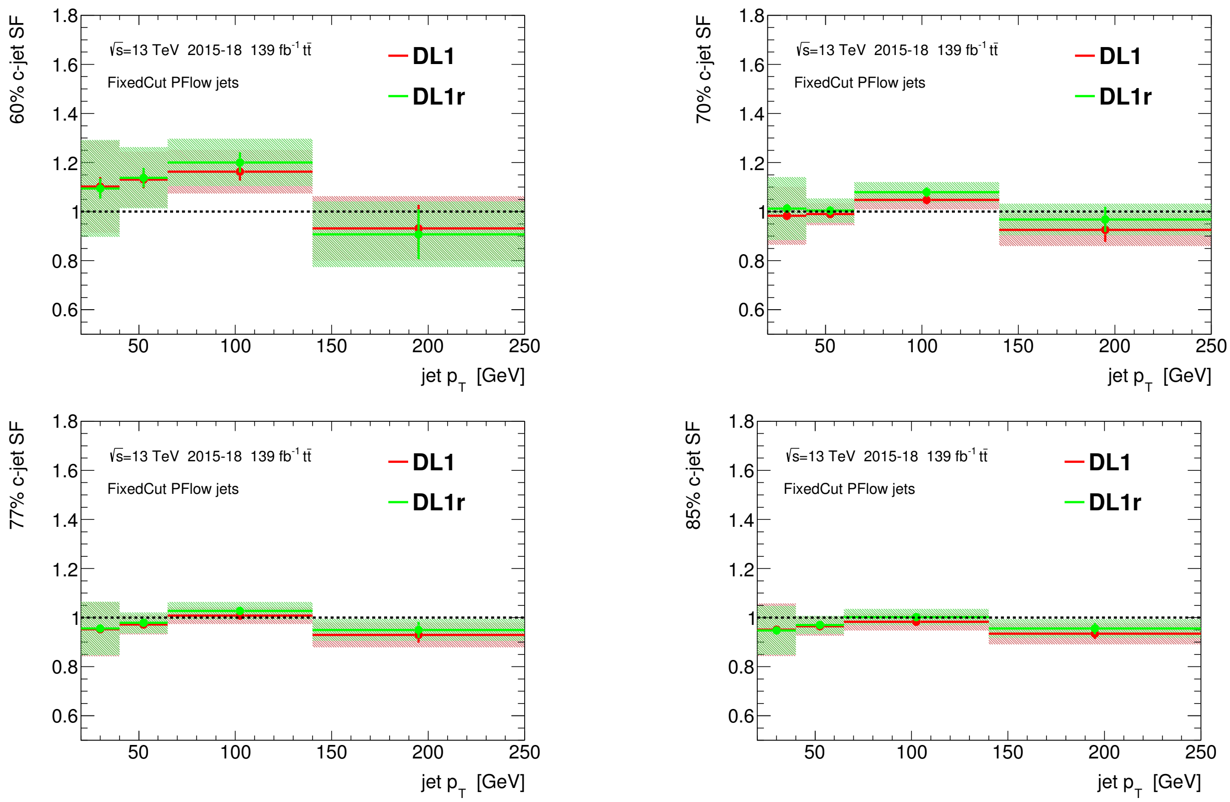


Figure 13: Charm-jet scale factors of the PFlow jets collection of given for 4 different working points.

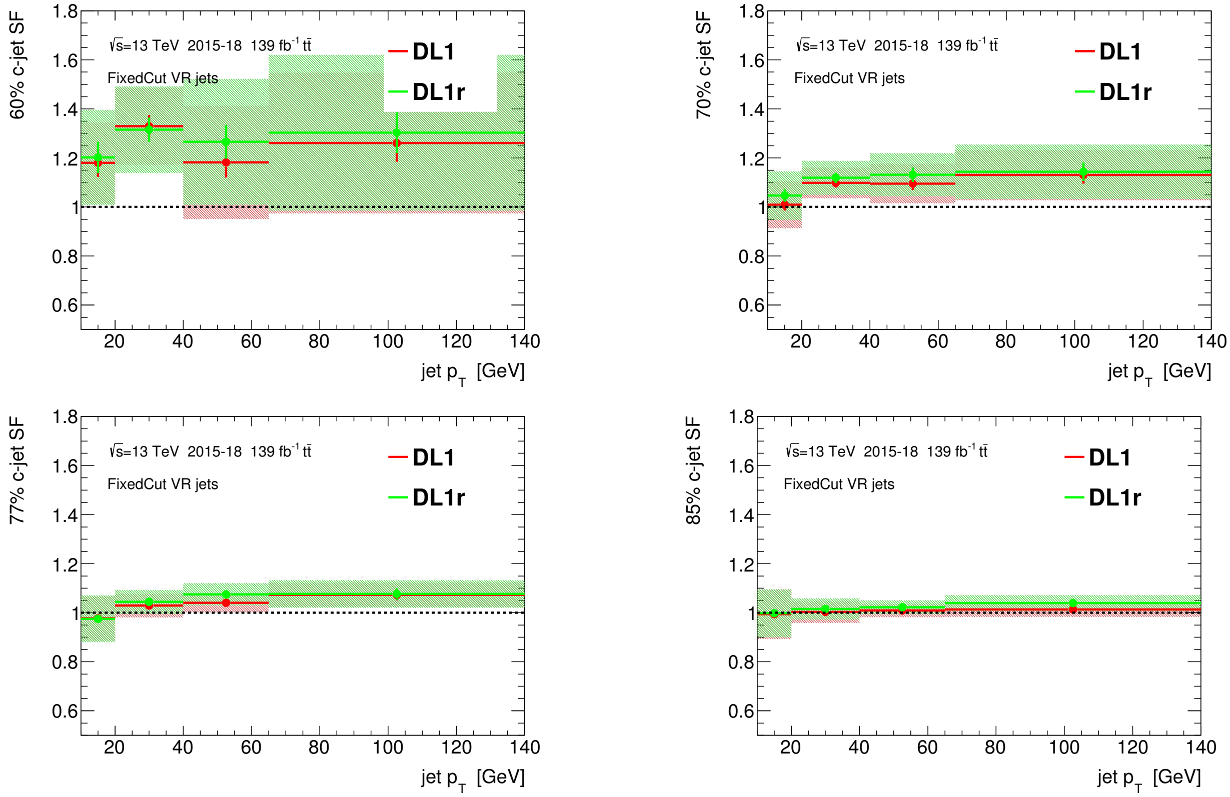


Figure 14: Charm-jet scale factors of VR-Track jets collection given for 4 different working points.

7 Search for Higgs boson pair production in the $b\bar{b}\tau\bar{\tau}$ channel

7.1 Data and Monte Carlo samples

7.2 Trigger and event selection

7.3 Background estimation

7.4 Multivariate analysis

7.5 Systematic uncertainties

7.6 Results

8 Summary

References

[1] ATLAS Collaboration. "Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 — interpreted in the phenomenological MSSM". In: *JHEP* 10 (2015), p. 134. DOI: 10.1007/JHEP10(2015)134. arXiv: 1508.06608 [hep-ex].

[2] ATLAS Collaboration. "Combination of searches for Higgs boson pairs in *pp* collisions at 13 TeV with the ATLAS experiment". In: (2018). URL: https://cds.cern.ch/record/2638212.

[3] Dean Carmi et al. "Interpreting LHC Higgs results from natural new physics perspective". In: *Journal of High Energy Physics* 2012.7 (July 2012). ISSN: 1029-8479. DOI: 10.1007/jhep07(2012)136. URL: http://dx.doi.org/10.1007/JHEP07(2012)136.

[4] "Performance of *b*-jet identification in the ATLAS experiment". In: *Journal of Instrumentation* 11.04 (Apr. 2016), P04008–P04008. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/04/p04008. URL: http://dx.doi.org/10.1088/1748-0221/11/04/P04008.

[5] ATLAS Collaboration. "ATLAS *b*-jet identification performance and efficiency measurement with $t\bar{t}$ events in *pp* collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[6] Luca Scodellaro. "b tagging in ATLAS and CMS". In: (2017). arXiv: 1709.01290 [hep-ex]. URL: https://arxiv.org/abs/1709.01290.

[7] ATLAS Collaboration. "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run". In: (2017). URL: https://cds.cern.ch/record/2273281.

[8] ATLAS Collaboration. "Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV". In: (2018). URL: https://cds.cern.ch/record/2314418.

[9] G. Abbiendi et al. "A measurement of the rate of charm production in W decays". In: *Physics Letters B* 490.1-2 (Sept. 2000), pp. 71–86. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(00)00971-0. URL: http://dx.doi.org/10.1016/S0370-2693(00)00971-0.

[10] K.A. Olive. "Review of Particle Physics". In: *Chinese Physics C* 38.9 (Aug. 2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001. URL: https://doi.org/10.1088/1674-1137/38/9/090001.

[11] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[12] Johannes Erdmann et al. "A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 748 (2014), pp. 18–25. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.02.029.

[13] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 71 (2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv: 1101.2185 [hep-ex].

[14] ATLAS Collaboration. "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 73 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[15] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 76 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[16] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[17] ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2". In: *JINST* 15.09 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447 [hep-ex].

[18] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[19] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[20] S. Agostinelli et al. "GEANT4 – a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[21] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[22] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), p. 244. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[23] ATLAS Collaboration. "The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model". In: (2016). URL: https://cds.cern.ch/record/2206965.

[24] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[25] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[26] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[27] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[28] ATLAS Collaboration. "Studies on top-quark Monte Carlo modelling for Top2016". In: (2016). URL: https://cds.cern.ch/record/2216168.

[29] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[30] Richard D. Ball et al. "Parton distributions for the LHC run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[31] Emanuele Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *Eur. Phys. J. C* 71 (2011), p. 1547. DOI: 10.1140/epjc/s10052-011-1547-z. arXiv: 1009.2450 [hep-ph].

[32] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[33] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[34] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[35] Federico Buccioni et al. "OpenLoops 2". In: *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: 10.1140/epjc/s10052-019-7306-2. arXiv: 1907.13071 [hep-ph].

[36] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[37] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. "COLLIER: A fortran-based complex one-loop library in extended regularizations". In: *Comput. Phys. Commun.* 212 (2017), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph].

[38] Steffen Schumann and Frank Krauss. "A parton shower algorithm based on Catani–Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph].

[39] Stefan Höche et al. "A critical appraisal of NLO+PS matching methods". In: *JHEP* 09 (2012), p. 049. DOI: 10.1007/JHEP09(2012)049. arXiv: 1111.1220 [hep-ph].

[40] Stefan Höche et al. "QCD matrix elements + parton showers. The NLO case". In: *JHEP* 04 (2013), p. 027. DOI: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph].

[41] S. Catani et al. "QCD Matrix Elements + Parton Showers". In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231.

[42] Stefan Höche et al. "QCD matrix elements and truncated showers". In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv: 0903.1219 [hep-ph].

[43] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[44] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[45] ATLAS Collaboration. “ATLAS Pythia 8 tunes to 7 TeV data”. In: (2014). URL: https://cds.cern.ch/record/1966419.

[46] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[47] ATLAS Collaboration. “Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector”. In: *JHEP* 04 (2019), p. 046. DOI: 10.1007/JHEP04(2019)046. arXiv: 1811.12113 [hep-ex].

A Supplementary material for ATLAS flavour tagging

A.1 Additional plots for kinematic variables

A.1.1 Standard selection

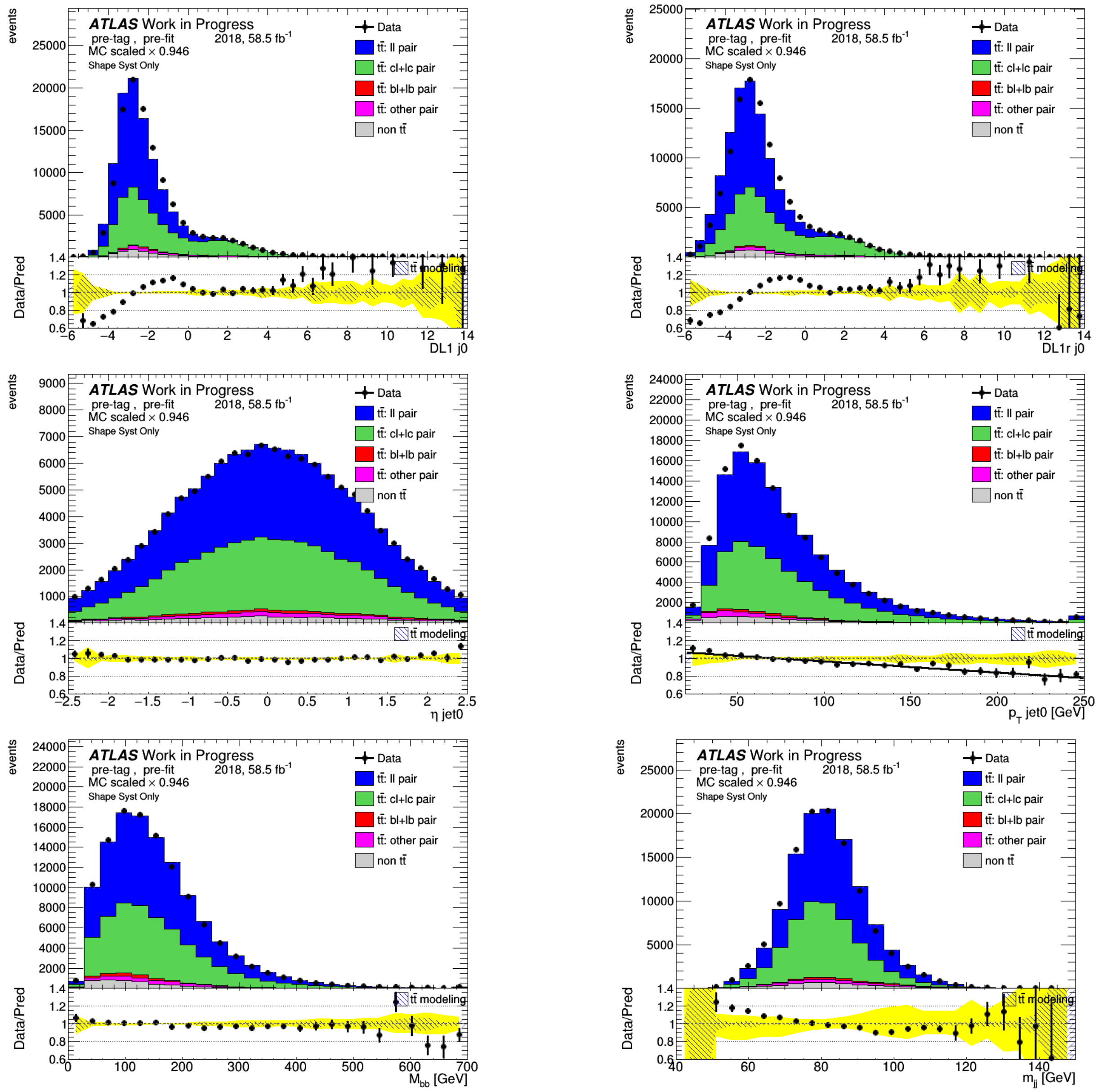


Figure 15: Standard selection: various kinematic distributions of the standard selection before fitting or tagging with systematics. The top two plots are the distributions of the DL1 and Dl1r score distributions of the leading jet.

A.1.2 Low-p_{T} selection

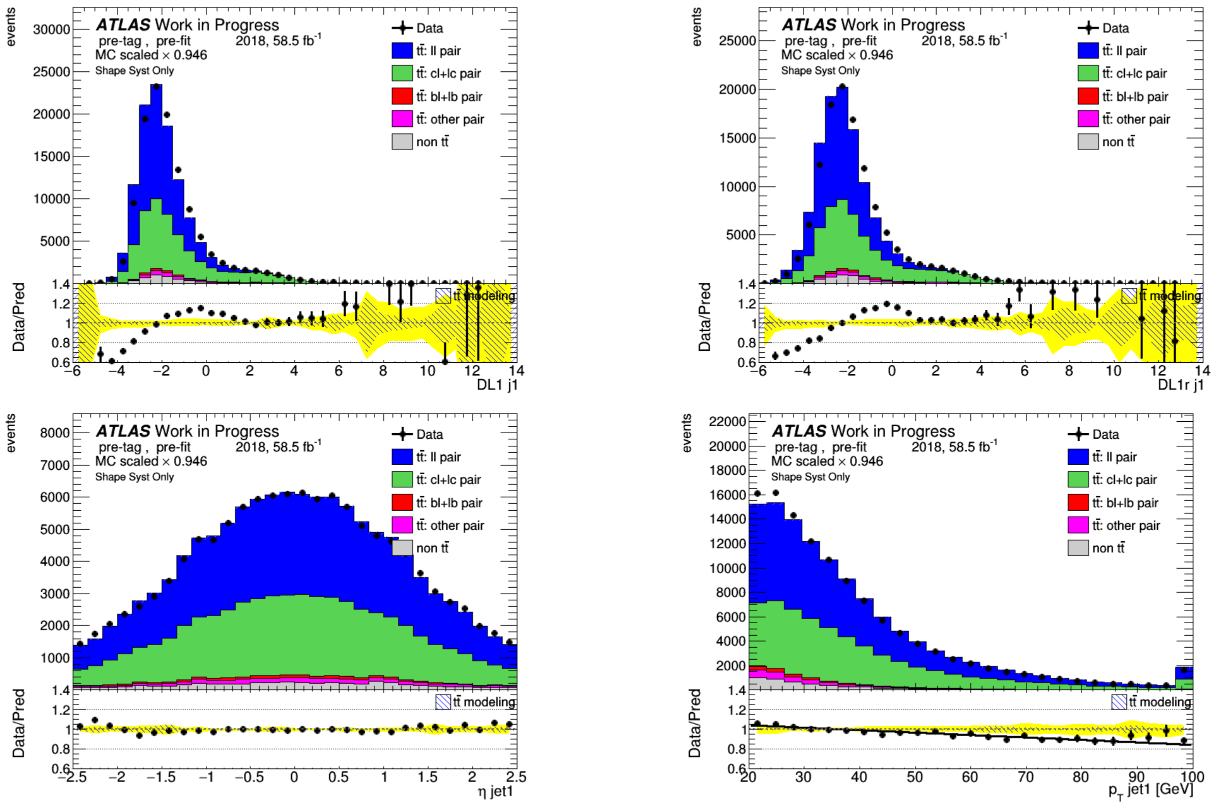


Figure 16: Selection for low-p_T extension: kinematic distributions of the sub-leading W decay jet before fitting or tagging with systematics. The top two plots are distributions of the DL1 and Dl1r score of the sub-leading jet.

A.2 High-p_{T} selection

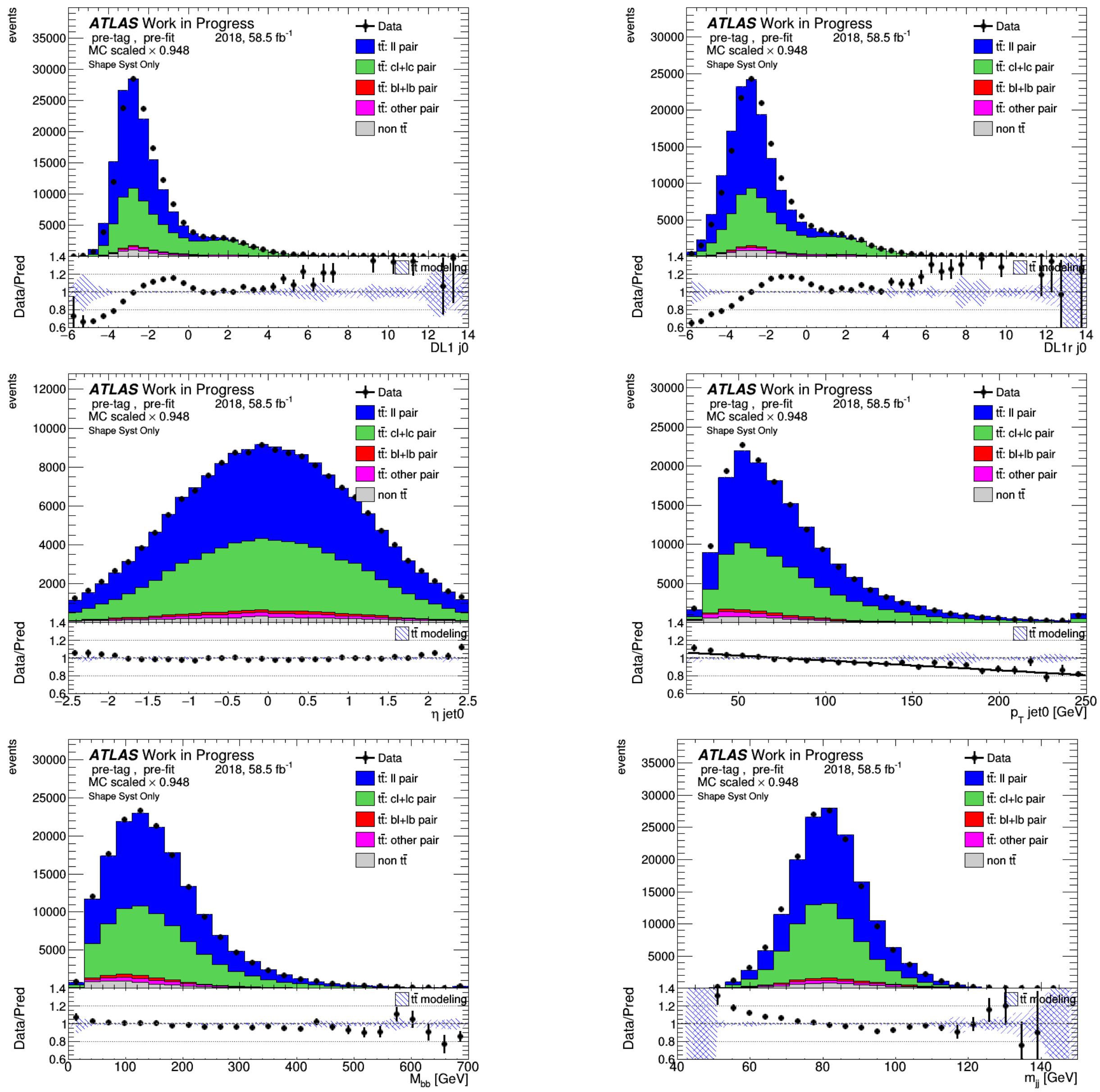


Figure 17: Various kinematic distributions of the combination of the standard selection and the high-p_T selection, before fitting or tagging with stat-only uncertainties.

A.2.1 Combined selection

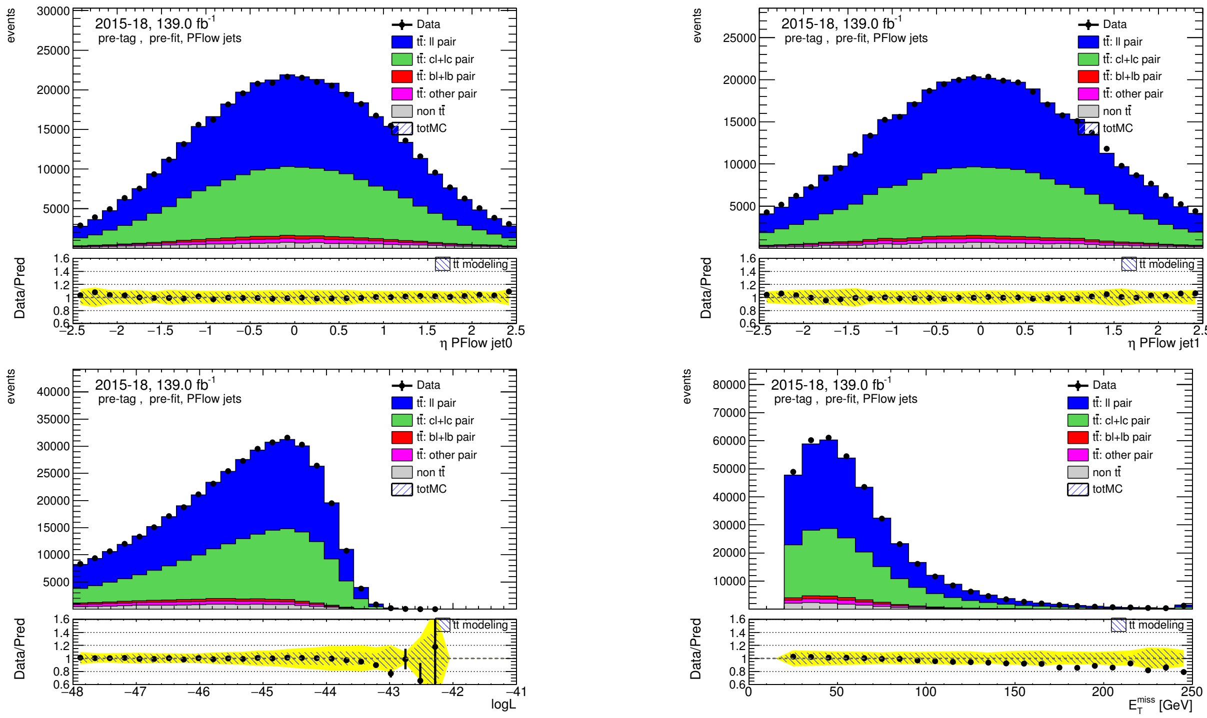


Figure 18: Distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

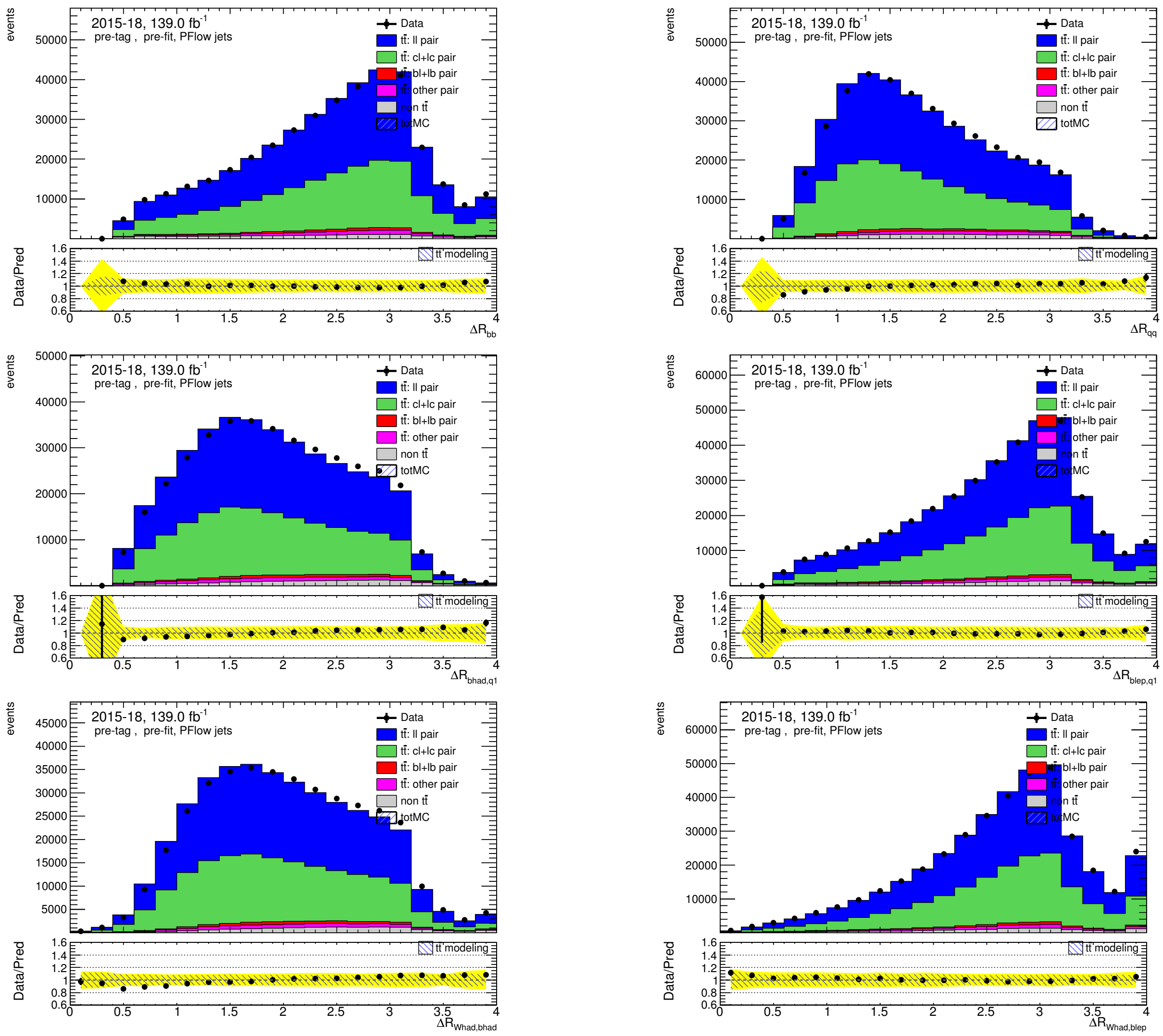


Figure 19: Distributions of angle related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

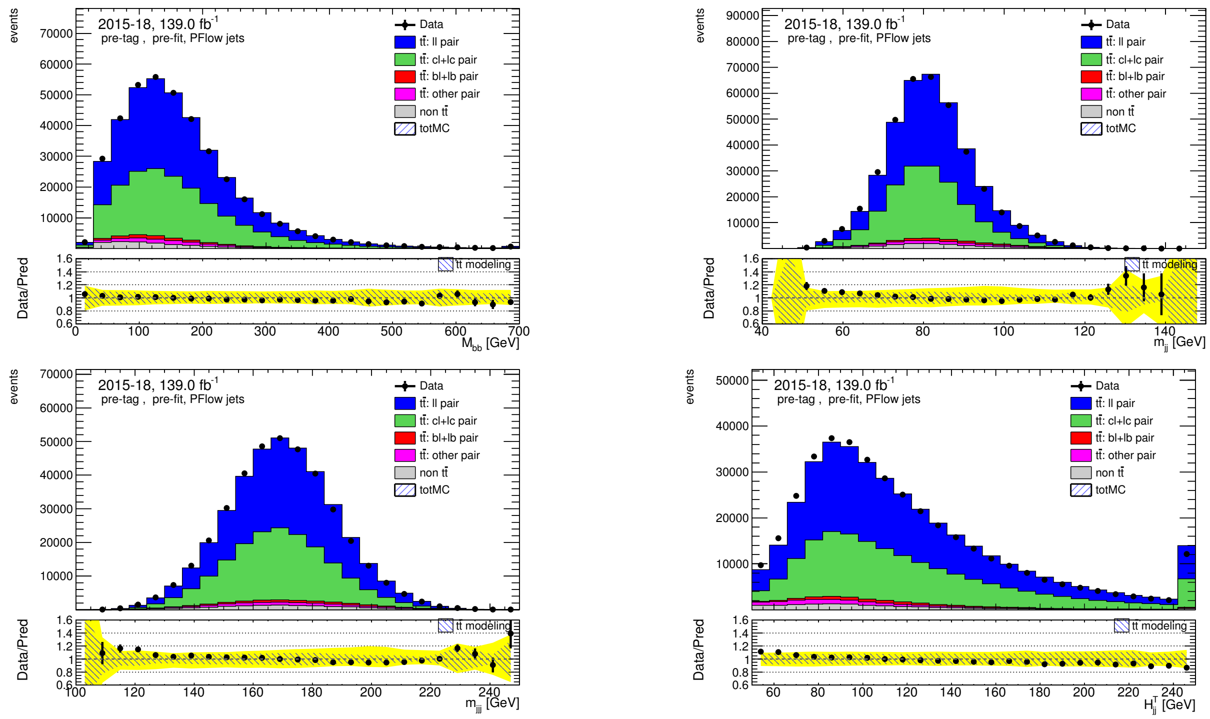


Figure 20: Distributions of mass related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with stat-only uncertainties.

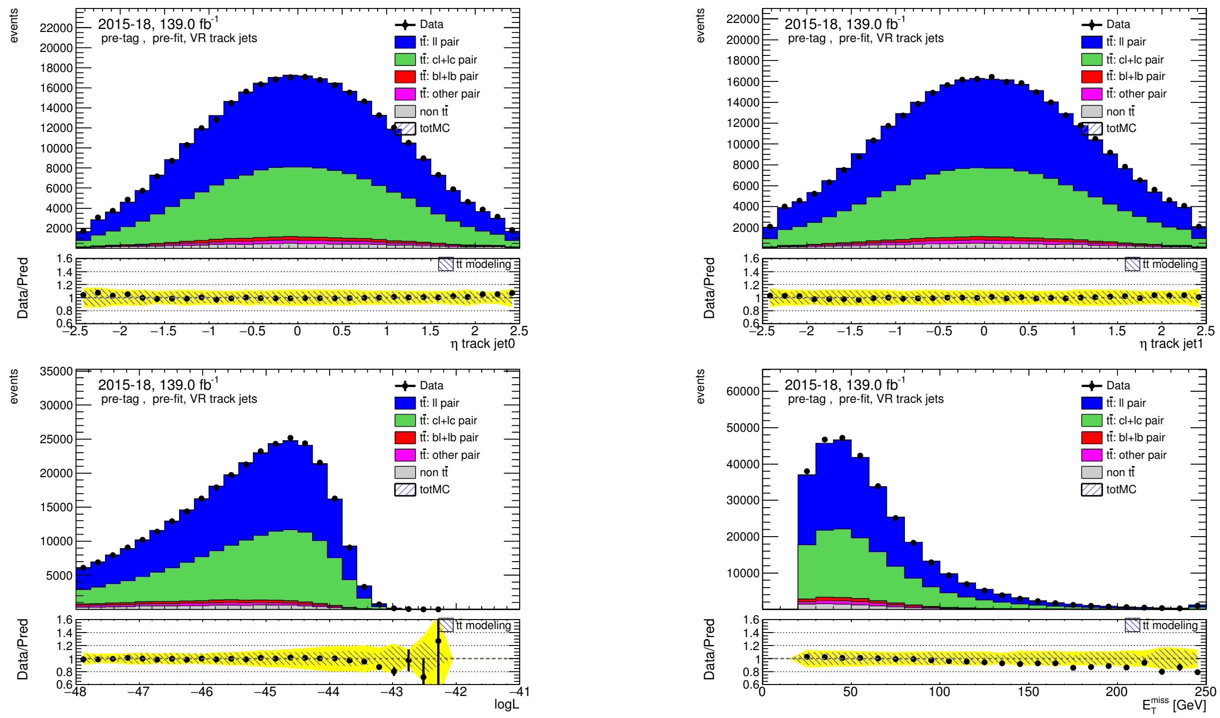


Figure 21: Distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

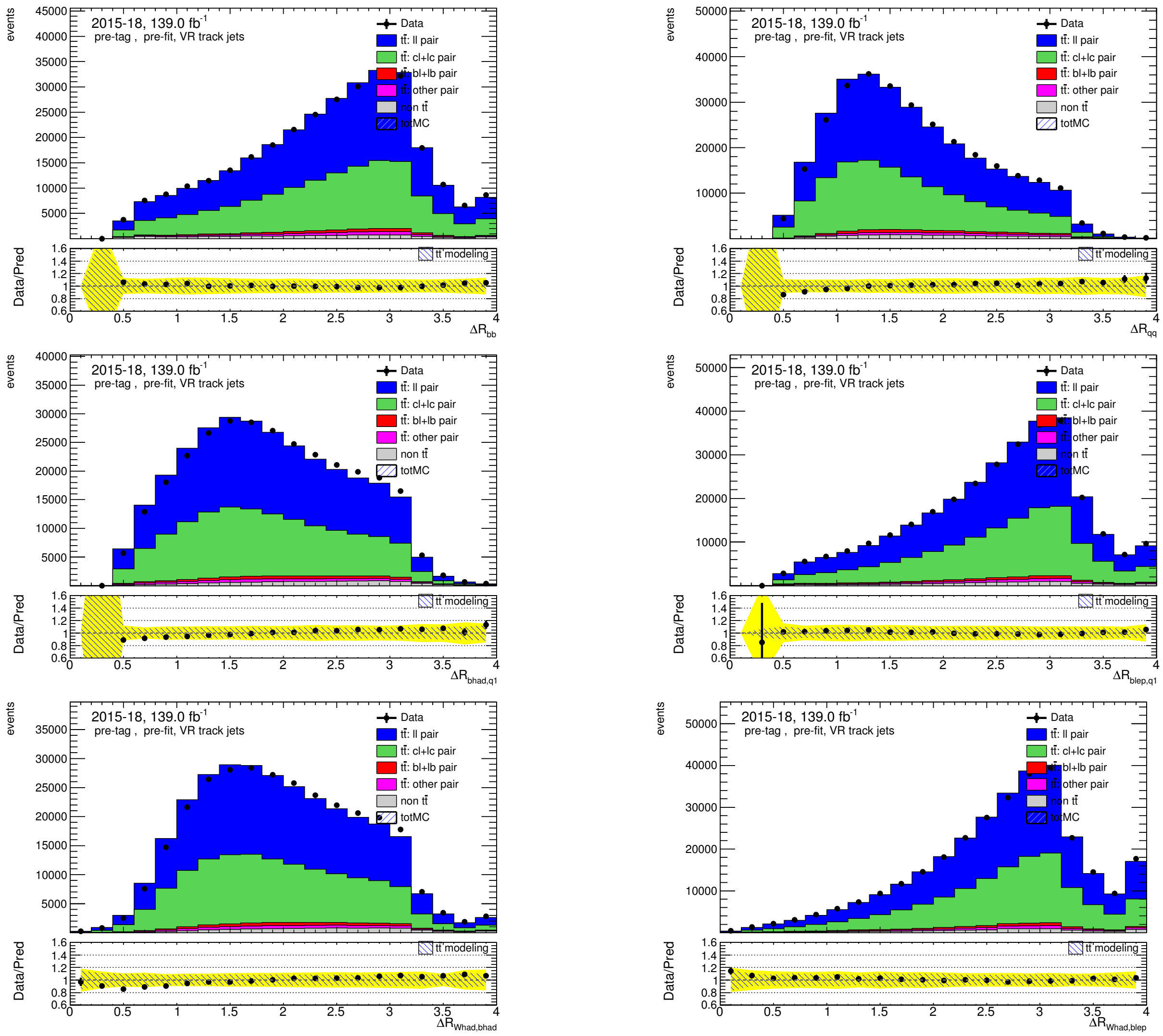


Figure 22: Distributions of angle related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

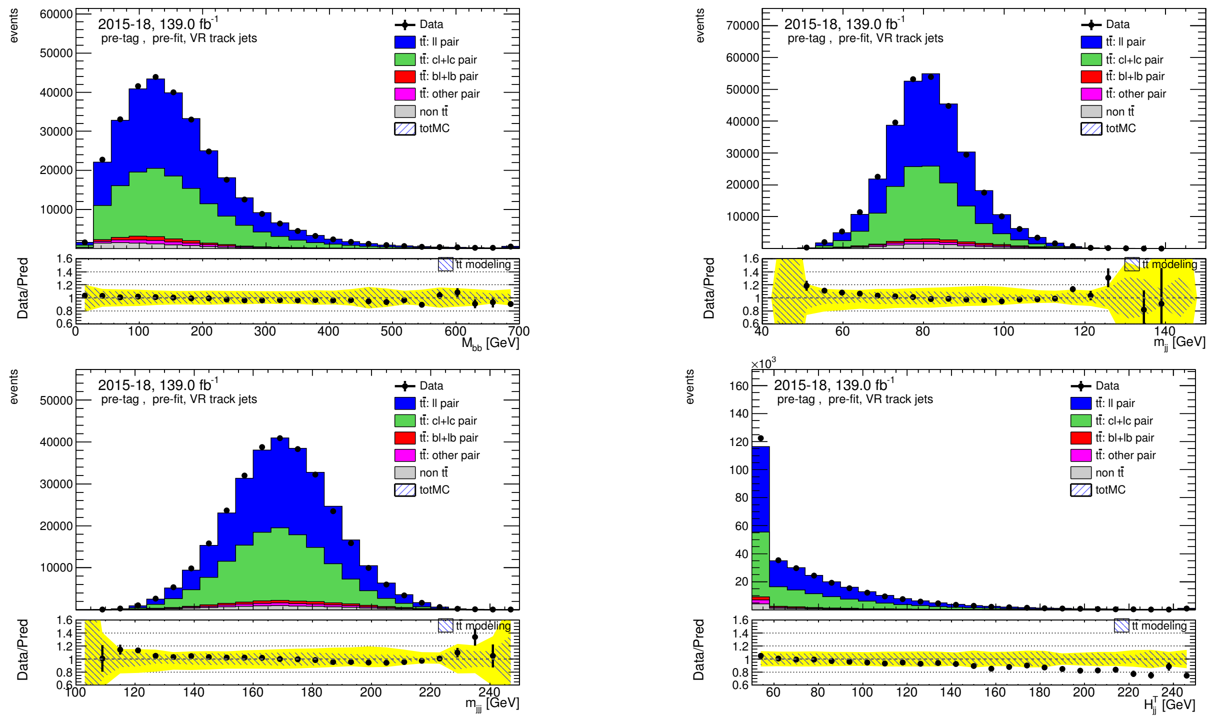


Figure 23: Distributions of mass related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with stat-only uncertainties.

A.3 Plots for previous calibrations

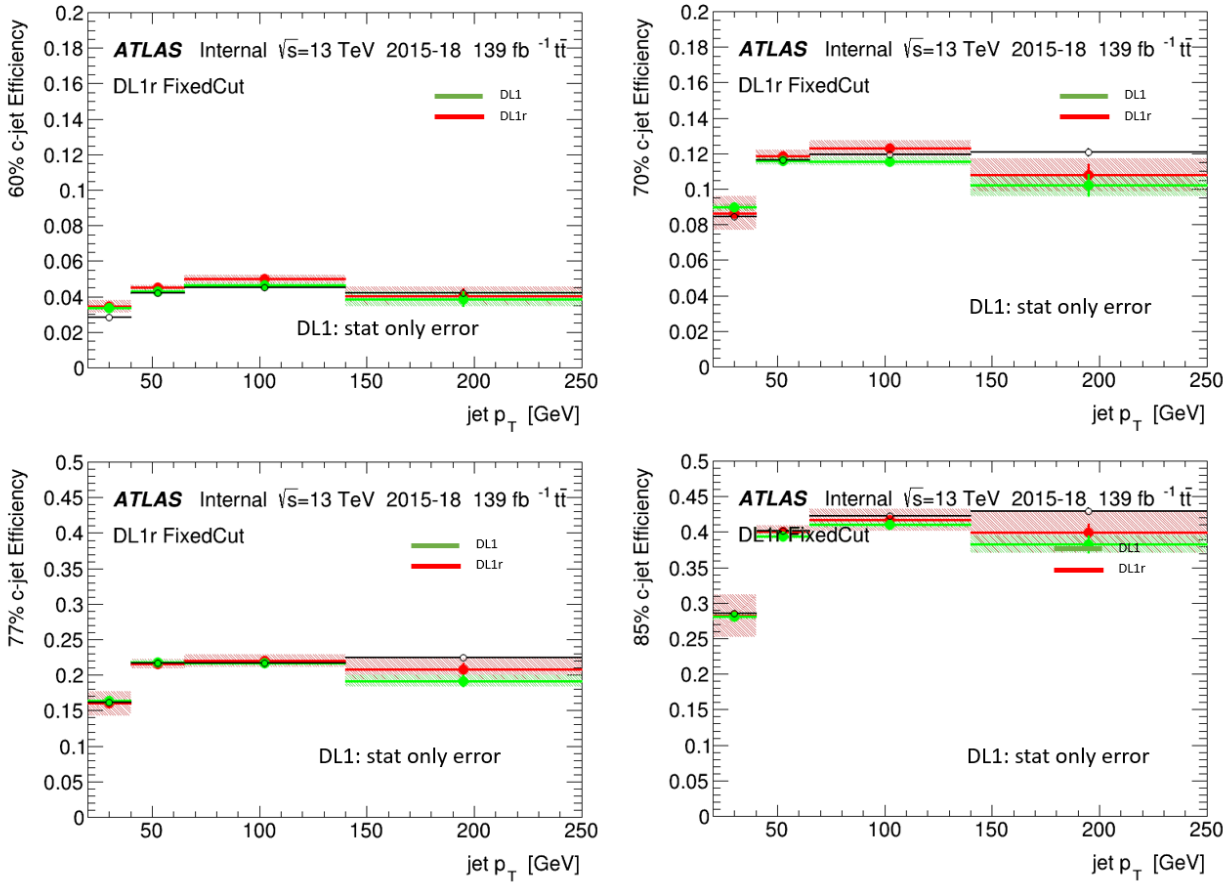


Figure 24: Calibration of derivation p3970 in December 2019, given for 4 different working points.

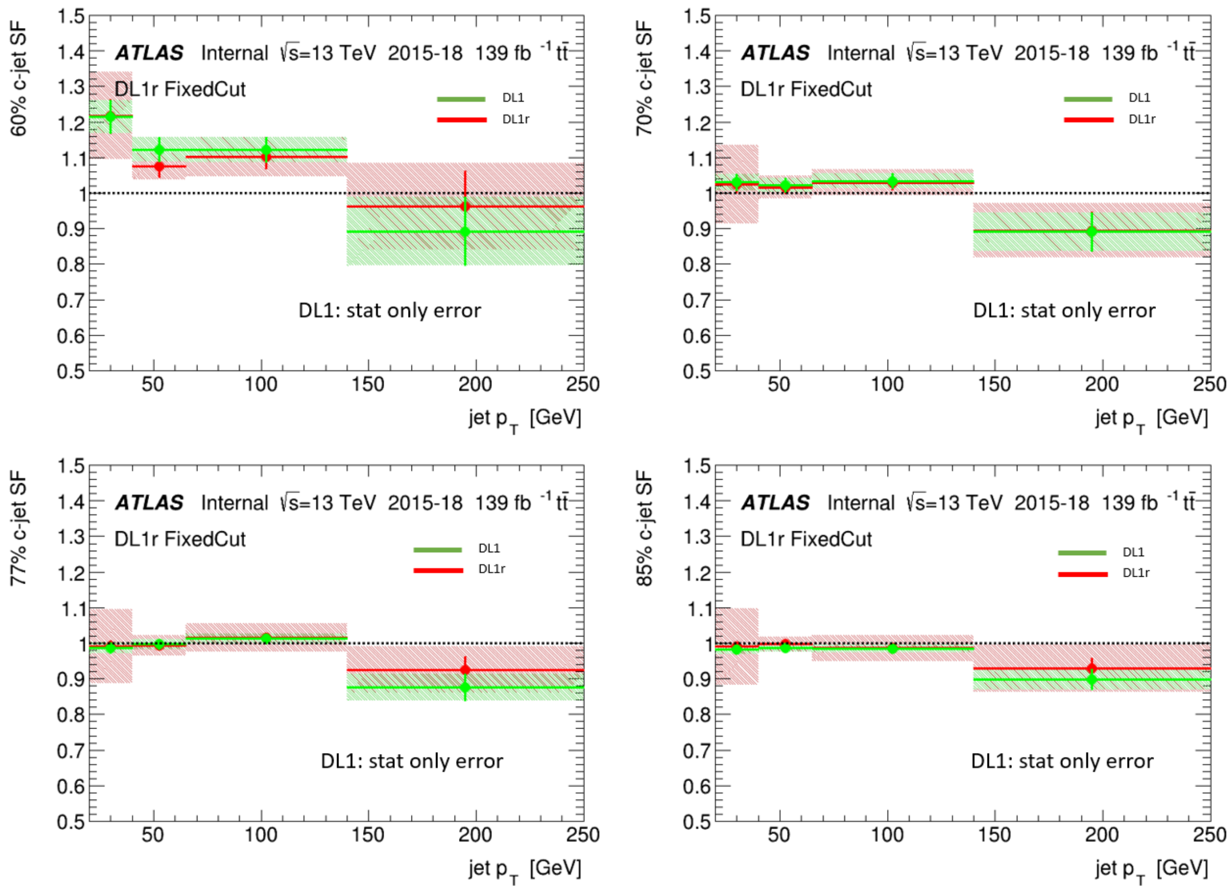


Figure 25: Calibration result of derivation p3970 in December 2019, given for 4 different working points.

A.4 Experimental uncertainties

Systematic uncertainty
EG_RESOLUTION_ALL
MUON_ID
MUON_MS
MET_SoftTrk_ResoPara
MET_SoftTrk_ResoPerp
MET_SoftTrk_ScaleDown
MET_SoftTrk_ScaleUp
JET_Pileup_OffsetNPV
JET_Pileup_RhoTopology
JET_EffectiveNP_Modelling1
JET_EffectiveNP_Modelling2
JET_EffectiveNP_Modelling3
JET_EffectiveNP_Modelling4
JET_EffectiveNP_Statistical4
JET_EffectiveNP_Detector1
JET_JER_EffectiveNP_1
JET_JER_EffectiveNP_2
JET_JER_EffectiveNP_3
JET_JER_EffectiveNP_4
JET_BJES_Response
JET_Flavor_Composition
JET_Flavor_Response

Table 6: List of experimental systematics.