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

Zhiyuan Li

January 19, 2021

logo-eps-converted-to.pdf

Contents

1 Flavour tagging at ALTAS

1.1 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 the di-Higgs to $bb\tau\tau$ searches, as well as the recent observation of the Higgs boson decay into bottom quarks [**HIGG-2018-04**] and of its production in association with a top-quark pair [**HIGG-2018-13**], and plays a important role in a large number of Standard Model (SM) precision measurements, studies of the Higgs boson properties, and searches for new phenomena[**SUSY-2014-08**], [**ATLAS-CONF-2018-043**] [**Interpreting Higgs result**].

The ATLAS Collaboration uses various algorithms to identify b-jets [**PERF-2012-04**], 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 ($< c\tau > \approx 450 \ \mu m$), measurable b-hadrons have a significant mean flight length,$< l >=\beta\gamma c\tau$, 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 [**PERF-2012-04**], have been improved and retuned for Run 2. Secondly, in order to maximise the b-tagging performance, the results of the low-level b-tagging algorithms are combined in high-level algorithms consisting of 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 $\epsilon_c(\epsilon_l)$.

In this chapter, the performance of the algorithms is quantified in terms of c-jet and light-flavour jet rejections, defined as $1/\epsilon_c$ and $1/\epsilon_l$, respectively. The imperfect description of the detector response and physics modelling effects in Monte Carlo (MC) simulations necessitates the measurement of the performance of the b-tagging algorithms with collision data [**PERF-2012-04 ttbar**, **b-jet identification semi ttbar**]. In this chapter, the measurement of the b-jet tagging efficiency of the high-level b-tagging algorithms used in proton–proton (pp) collision data recorded during Run 2 of the LHC at $\sqrt{s} = 13$ TeV is presented. The corresponding measurements for c-jets and light-flavour jets, used in the

measurement of the b-jet tagging efficiency to correct the simulation such that the overall tagging efficiency of c-jets and light-flavour jets match that of the data, are described elsewhere [**ATLAS-CONF-2018-006**], [**cjet**]. 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 is selected by requiring that both W bosons decay leptonically, referred to as di-leptonic $t\bar{t}$ decays in the following. Similarly a source of c-jets can be obtained by requiring that one W boson decay leptonically and the other decay hadronically, referred to as semi-leptonic $t\bar{t}$ decay in the following.

The most performant algorithms presently in use in physics analyses at ATLAS are based on multivariate (MV2) combinations of the available information or a deep feed-forward neural network(DL1)[**tagging**][**ATL-PHYS-PUB-2017-013**], as shown in Figure **??**. 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)[**ATL-PHYS-PUB-2017-013**]. The DL1r tagger is now the default b-tagging algorithm used for flavour tagging in ATLAS.

1.2   Calibration methods

Monte Carlo (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 match the data in terms of b-tagging efficiency, charm-jet mis-tagging and light-jet mis-tagging rates[**FTAG-2018-01**]. The calibration is done 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. More details of the calibrations of different flavour are given in the following subsections.

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 a di-leptonic (both W bosons from the top decay decay into leptons) $t\bar{t}$ sample of highly enriched in $t\bar{t}$ events. Events with 2 jets and 2 opposite signs leptons are selected from di-leptonic $t\bar{t}$ samples. 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[**FTAG-2018-01**].

For the light jet mis-tagging calibration, two methods are used to measure the mistagging rate from the data: 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

MV2_DL1-eps-converted-to.pdf

Figure 1: Performance of MV2c10 and DL1 tagger[**FTAG-2018-01**].

criteria used in the nominal identification algorithm, and 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[**ATLAS-CONF-2018-006**].

The charm-jet mis-tagging calibration utilises semi-leptonic $t\bar{t}$ events. The events kinematics are shown by the diagram in Figure **??**, where the $t\bar{t}$ pair decays to a pair of b and $\bar{t}$ quark, circled in red. One of the W boson, circled in blue, decays hadronically to quarks, and the other W boson, decays leptonically to either electron or muon and the corresponding neutrinos, circled in purple and green. The lepton in the final state is used for triggering, and a combined likelihood fit is used to extract the c mis-tagging efficiency (more details in Section **??**).
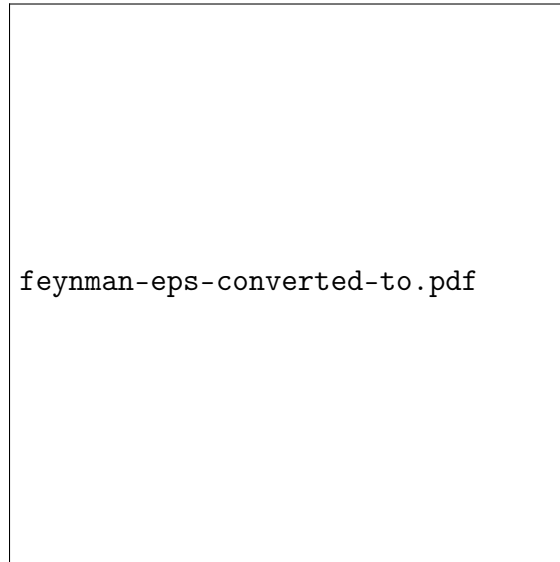

Figure 2: Feynman diagram of the semi-leptonic $t\bar{t}$ events.

2 Charm mis-tagging calibration

2.1 Introduction

As determined by the CKM matrix[**CKM1**][**CKM2**], the W boson decay to light-quark pair or light quark c quark pair dominantly, and very rarely decay to pairs containing b-quarks. More specifically, the branching ratio of a W boson decays to a pair of light quarks or a light quark and a c quark is 33.7%, and of pairs containing a b quark is only 0.058%[**PDG**]. Therefore, b-tagged jets from the W decay are most likely to be mis-tagging of c-jets or light-

jets. Given the ratio between the DL1 light-jet rejection and the corresponding charm-jet rejection ranges from 10 to 40 (see ref[**ATL-PHYS-PUB-2017-013**]), the c-jet in c-jet-light-jet pairs is the most likely source for the b-tagged jet. A kinematic likelihood technique, referred to as KLFitter[**ERDMANN201418**], is used to assign jets to the proper $t\bar{t}$ decay product without using any b-tagging information (more details in Section **??**). The charm-jet efficiency is then extracted by a combinatorial likelihood fit applying to the pair of jets from W decays, where the main floating parameter is the c-jet efficiency[**cjet**].

It is worth mentioning that my 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. The calibration uses the full Run2 data of 139 fb^{-1} integrated luminosity collected at $\sqrt{s} = 13$ TeV by the ATLAS detector. In the task, 4 rounds of calibration has been conducted, each round with different jet collection and b-tagging algorithm. To be concise only the result of the last round is presented in this chapter. During the task the calibration range has been extended down to 20 GeV (which was 25 GeV) and a new selection category has been developed to increase the stats of the scale factors in the high p_T (p_T greater than 70 GeV) region.

2.2 Data and Monte Carlo samples

The data analysed in this study correspond to 139 fb^{-1} [**DAPR-2010-01**, **DAPR-2011-01**, **DAPR-2013-01**, **LUCID2**], of pp collision data collected by the ATLAS detector between 2015 and 2018 with a centre-of-mass energy of 13 TeV and a 25 ns proton bunch crossing interval. The data sample was collected using a set of single-muon [**Aad:2020uyd**] and single-electron triggers [**TRIG-2018-05**]. The single-muon triggers had p_T thresholds in the range of 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 in the range 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.

Dedicated Monte Carlo simulated samples (MC) are used to model SM processes and to estimate the expected signal yields. All samples were produced using the ATLAS simulation infrastructure [**SOFT-2010-01**] and GEANT4 [**Agostinelli:2002hh**]. A subset of samples use a faster simulation based on a parameterisation of the calorimeter response and GEANT4 for the other detector systems [**SOFT-2010-01**]. 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 [**Sjostrand:2014zea**] using the NNPDF2.3LO set of PDFs [**Ball:2012cx**] and a set of tuned parameters called the A3 tune [**ATL-PHYS-PUB-2016-017**].

The events that are used in this study originate mostly due to ttbar production. This process is modelled using the POWHEGBOX v2 [**Frixione:2007nw**, **Nason:2004rx**, **Frixione:2007vw**, **Alioli:2010xd**] generator at NLO with the `NNPDF3.0nlo` parton distribution function (PDF) set and the h_{damp} parameter[1] set to 1.5 m_{top} [**ATL-PHYS-PUB-2016-020**]. 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 [**EvtGen**].

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` [**Ball:2014uwa**] parton distribution function (PDF) set. The associated production of top quarks with W bosons (tW) is modelled using the POWHEGBOX v2 [**Re:2010bp**, **Nason:2004rx**, **Frixione:2007vw**, **Alioli:2010xd**] generator at NLO in QCD using the five-flavour scheme and the `NNPDF3.0nlo` set of PDFs [**Ball:2014uwa**]. The diagram removal scheme [**Frixione:2008yi**] 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 [**Bothmann:2019yzt**] 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 [**Gleisberg:2008fv**] and OPENLOOPS [**Buccioni:2019sur**, **Cascioli:2011va**, **Denner:2016kdg**] libraries. They are matched with the SHERPA parton shower [**Schumann:2007mg**] using the MEPS@NLO prescription [**Hoeche:2011fd**, **Hoeche:2012yf**, **Catani:2001cc**, **Hoeche:2009rj**] using the set of tuned parameters developed by the SHERPA authors. The `NNPDF3.0nnlo` set of PDFs [**Ball:2014uwa**] is used and the samples are normalised to a next-to-next-to-leading order (NNLO) prediction [**Anastasiou:2003ds**].

Samples of diboson final states (VV) are simulated with the SHERPA v2.2.1 or v2.2.2 [**Bothmann:2019y** 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

---

[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.

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 [**Gleisberg:2008fv**, **Schumann:2007mg**] using the MEPS@NLO prescription [**Hoeche:2011fd**, **Hoeche:2012yf**, **Catani:2001cc**, **Hoeche:2009rj**]. The virtual QCD correction are provided by the OPENLOOPS library [**Buccioni:2019sur**, **Cascioli:2011va**, **Denner:2016kdg**]. 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 [**Alwall:2014hca**] generator at NLO with the `NNPDF3.0nlo` [**Ball:2014uwa**] parton distribution function (PDF). The events are interfaced to PYTHIA8.210 [**Sjostrand:2014zea**] using the A14 tune [**ATL-PHYS-PUB-2014-021**] and the `NNPDF2.3lo` [**Ball:2014uwa**] PDF set. The decays of bottom and charm hadrons are simulated using the EVTGEN v1.2.0 program [**Lange:2001uf**].

2.3   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. [**ERDMANN201418**]. 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 b-jet from the leptonically decaying top-quark ($t \to Wb \to \ell\nu b$) one other 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.

2.3.1   Maximising likelihood

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 jets (from light or c-quarks) resulting from the hadronic W decay 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. 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
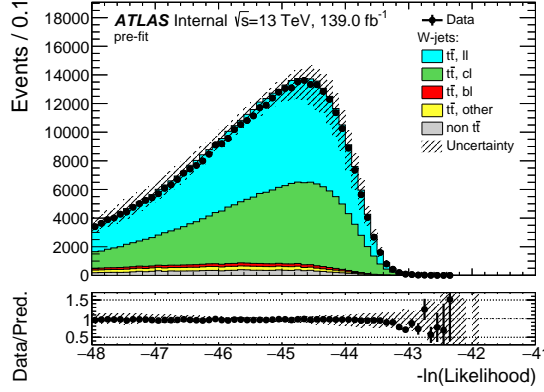
Figure 3: Distribution of the negative logarithm of the likelihood that is used to reconstruct the $t\bar{t}$ decay.

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[**cjet**]. To improve the purity of the selected events further, 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 **??**. 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, which comes from other processes like W or Z production in association with jets or single-top production. 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" ("$t\bar{t}$, 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.

2.4 Event selection

2.4.1 Standard selection

Events are required to contain exactly one trigger-matched lepton with p_T above 27 GeV and exactly four jets with p_T above 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 above 27 GeV are rejected. In addition, events are required to have

a minimum of 20 GeV missing transverse momentum, which is assumed to be the result of the neutrino from the leptonically decaying W boson. The events are also required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and 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 $E_{\mathrm{T}}^{\mathrm{miss}}$. For each event, the kinematic likelihood fitter is applied which determines the jet assignment to the various top decay products assumed from the $t\bar{t}$ decay products. A requirement of the log of KLFitter output > -48 is placed on the chosen event permutation. Due to the requirement on jet p_T and binning strategy, the calibration result can be applied to jets with p_T between 25 to 200 GeV. The inclusive yields of the standard selection and the low p_{T} selection (defintion in Section **??**)of the data/MC are given in Table **??**, an example of the p_{T} distributions before any tagging or fitting and after the standard selection is shown in Figure **??**. More plots can be found in Appendix, Figure **??**. The meaning of each legend in the plots are described in Section **??**. 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 **??** The data/MC ratio shows good agreement within the systematic uncertainties.

2.4.2 Selection for low-p_T extension

In order to extend the calibration in the low-p_T region so that the calibration can be applied to jets with p_T smaller to 25 GeV, instead of requiring events to have exactly 4 jets, events are required to have exactly 3 jets with p_T greater than 25 GeV and exactly 1 jet with p_T greater than 20 GeV. Other than that, all other requirements for the selection are the same. The distributions of the sub-leading jet are shown in appendix, Figure **??**. 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	287105.0	218351.0
Ttbar	292203.7 ± 202.8	223769.3 ± 177.6
Other	10945.7 ± 123.1	7280.1 ± 103.1
Data/MC	0.947 ± 0.002	0.945 ± 0.002

Table 1: 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, combining with the low p_{T} selection.
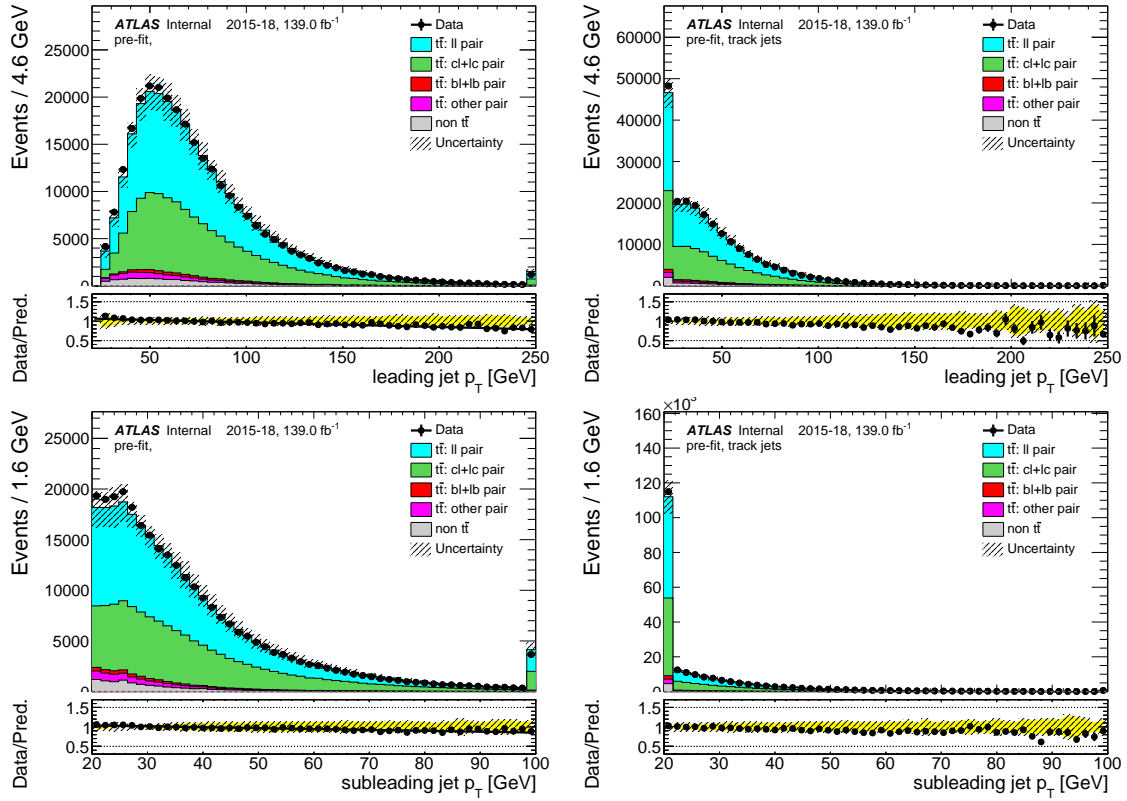
Figure 4: Data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column.
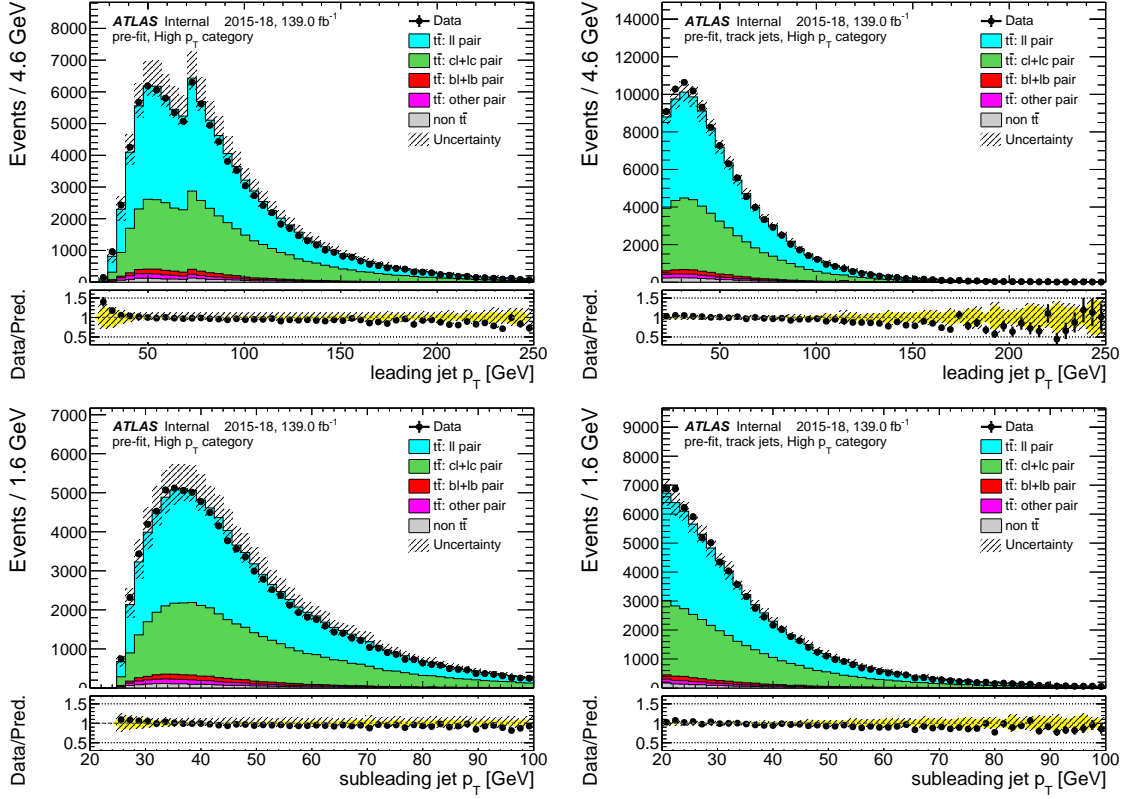
Figure 5: Data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column.

2.4.3 High p_T selection

Instead of requiring events to have exactly 4 jets, events are required to have at least 5 jets with p_T greater than 25 GeV, in which at least 1 jet with p_T greater than 70 GeV. The choice of cut value is based on study shown in the following section **??**. Other than that, all other requirements for the selection remaining the same. The yields of the data/MC are given in Table **??**, an example of the p_T distributions before any tagging or fitting and after the standard selection is shown in Figure **??**. More plots can be found in Appendix, Figure **??**.

	Particle flow jets	Track jets
Data	98273.0	120929.0
Ttbar	99432.4 ± 116.4	117090.1 ± 126.5
Other	1841.5 ± 21.2	2641.2 ± 47.6
Data/MC	0.970 ± 0.003	1.010 ± 0.003

Table 2: Prefit comparison of the number of events in data and in simulation considering particle flow jets and track jets for the high-p_T selection.

2.4.3.1 Optimisation of high p_T cut

In the high-p_T selection, the value to define high-p_T is chosen to be 70 GeV. A study has been done to investigate the effect on the c-jet purity and the potential statistical gain, where the c-jet purity is defined as:

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

where $N_{true \ \ c\text{-}jet}$ 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 c-jet purity and the statistical gain are calculated for 4 different cut values as shown in Figure **??**, comparing with the cut value of 0. The value of 70 GeV is chosen because of its relatively small effect on the c-jet purity and relatively high gain in statistics.

2.4.4 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 **??**, an example of the p_T distributions before any tagging or fitting and after the standard selection is shown in Figure **??**. More plots can be found in Appendix, Figure **??**.

2.5 Systematic uncertainties

The systematic uncertainties considered and propagated in this calibration can be broadly categorised into experimental systematics and modelling uncertainties. Supporting material for this section can be found in the appendix, Tab.**??**.

2.5.1 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 the electron trigger, identification and
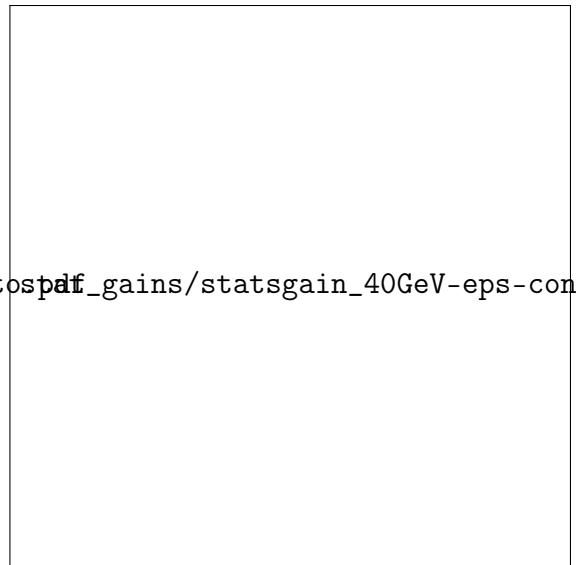
	Particle flow jets	Track jets
Data	385378.0	302308.0
Ttbar	383517.0 ± 229.4	302690.5 ± 203.9
Other	12424.0 ± 123.8	8566.8 ± 104.1
Data/MC	0.973 ± 0.002	0.971 ± 0.002

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

stat_gains/statsgain_0GeV-eps-converted-to.pdf

stat_gains/statsgain_40GeV-eps-converted-to.pdf

(a) Gain in stats and the c-jet purity with cut
value 0 GeV.

(b) Gain in stats and the c-jet purity with cut
value 40 GeV.

stat_gains/statsgain_70GeV-eps-converted-to.pdf

stat_gains/statsgain_90GeV-eps-converted-to.pdf

(c) Gain in stats and the c-jet purity with cut
value 70 GeV.

(c) Gain in stats and the c-jet purity with cut
value 90 GeV.

stat_gains/statsgain_140GeV-eps-converted-to.pdf

(d) Gain in stats and the c-jet purity with cut
value 140 GeV.
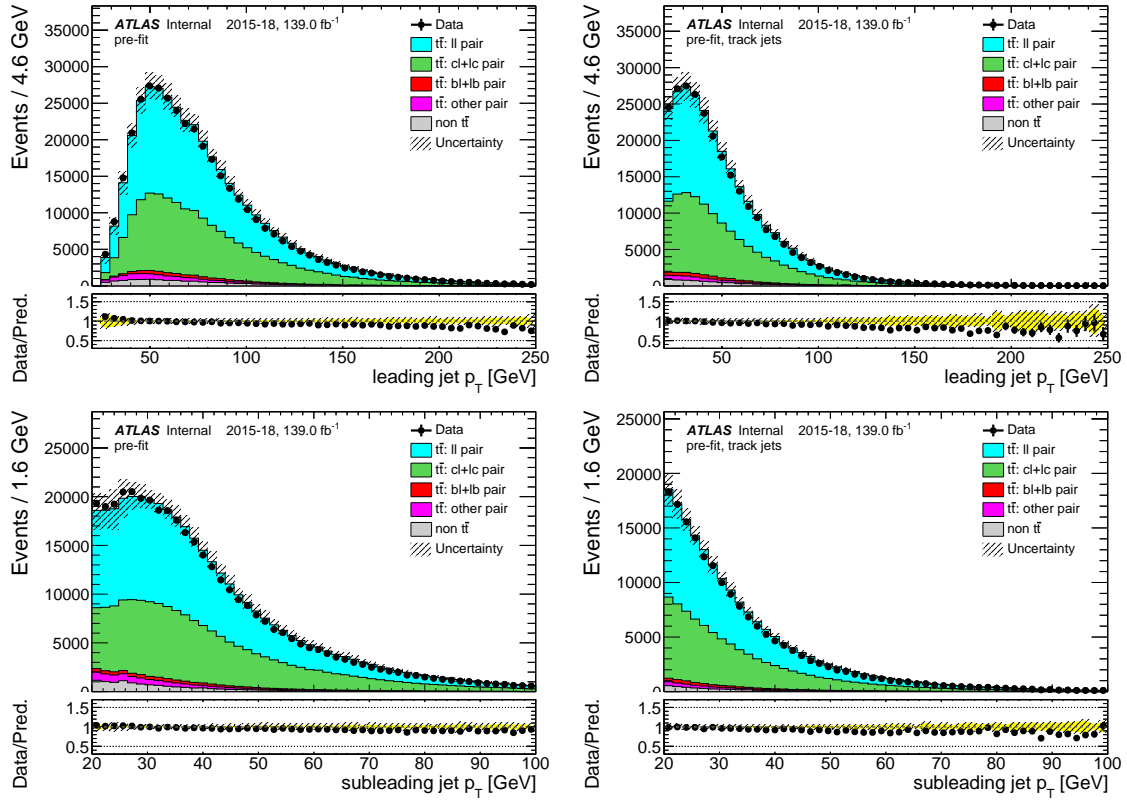
Figure 7: Data versus simulation for various variables used in the analysis for particle flow jets in the left column and for track jets in the right column.

reconstruction efficiencies, and for uncertainties associated with the isolation requirements. Scaling and smearing corrections are applied to the p_T of simulated muons in order to minimise the differences in resolution between data and MC events, and the uncertainties of the corrections are considered. Differences in the identification efficiency and in the efficiency of the trigger selection are also taken into account. 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_T^{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 top decay and for the jets associated to the hadronically decaying W boson.

2.5.2 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 matrix element generator, parton shower, and initial and final state radiation. In all cases, MC-to-MC SFs are taken into account. The uncertainty in the choice of NLO matrix element generator is derived by comparing two alternative predictions, POWHEG and aMC@NLO_MG5, each of which is showered with HERWIG++. An uncertainty due to the choice of parton shower and hadronisation model is derived by comparing the prediction from POWHEG interfaced either to PYTHIA8 or HERWIG++. Finally, the uncertainty on modelling of initial and final state radiation 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 comparison of the nominal $t\bar{t}$ and the systematics is shown in Table **??**.

Comparison of $t\bar{t}$ with systematic samples		
TtbarAF2	386261.9 ± 254.5 0.716%	304858.1 ± 226.1 0.716%
DATA/MC(AF2)	0.967 ± 0.002	0.965 ± 0.002
TtbarRadHi	377129.7 ± 224.0 -1.665%	297963.8 ± 200.3 -1.562%
DATA/MC(RADHI)	0.989 ± 0.002	0.986 ± 0.002
TtbarPH	331961.5 ± 217.0 -13.443%	259942.3 ± 193.4 -14.123%
DATA/MC(PH)	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.

17

Data	1589.0
$t\bar{t}$	1102.7 ± 12.5
Other	82.7 ± 5.5
Data/MC	1.341 ± 0.037

Table 5: Yields of the 2018 data/MC, combining the standard and the high p_T selection with at least 1 PFlow W jet with DL1r > 8 to reject most of the light and c jets.

Data	1336.0
$t\bar{t}$	943.5 ± 11.6
Other	69.1 ± 4.5
Data/MC	1.319 ± 0.040

Table 6: Yields of the 2018 data/MC, combining the standard and te high p_T selection with at least 1 VR-Track W jet with DL1r > 8 to reject most of the light and c jets.

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

The b-jets are rarely found in the two W decay jets. There are two main sources of the b-jets, first is a W boson decaying to pair of b and c quark. The other source is when the $t\bar{t}$ plus a gluon process is selected, and the gluon is split into a pair a b quarks. If we require no c-jet in either top-decay jets or W-jets, the only main source of the b-jets in the W-jets is the $t\bar{t}$ + heavy flavour process. It's shown in for the both PFlow and VR-Track jets collections, this process is underestimated by the MC by about 30%, as shown in Table **??** and Figure **??**. For this reason, events in the simulation in which the top-jets and at least one of the W-jets are b-jets, are scaled by 1.25 ± 0.25, where the value is measured in reference[**TOPQ-2017-12**]. Hence a 1.25 ± 0.25 scale factor is applied to all events with 3 truth b jets, and all results shown in this chapter have this scale factor implemented. The full difference between the simulation before applying this scale factor and after is taken as a systematic error. All the plots in this chapter have the uncertainty due to the scale factor implemented.

2.6 Results

2.6.1 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, the calibration includes the EMPFlow jet and VR-Track jet collection, and MV2c10, DL1 and DL1r taggers. Each calibration is carried out with all supported jet collections and taggers, with the full data set collected from 2015 to 2018 in the ATLAS detector. 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.

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 stats uncertainties.

2.6.2 Kinematic distributions

The kinematic distributions of the MC and the data of the latest calibration (October 2020) are shwon in Figure **??** for the PFlow jets and Figure **??** for the VR-Track jets, combining the standard selection and the highest 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 events, which is denoted as "non $t\bar{t}$" on the figure, which comes from other processes like W or Z production in association with jets or single-top production. 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" ("$t\bar{t}$, 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.
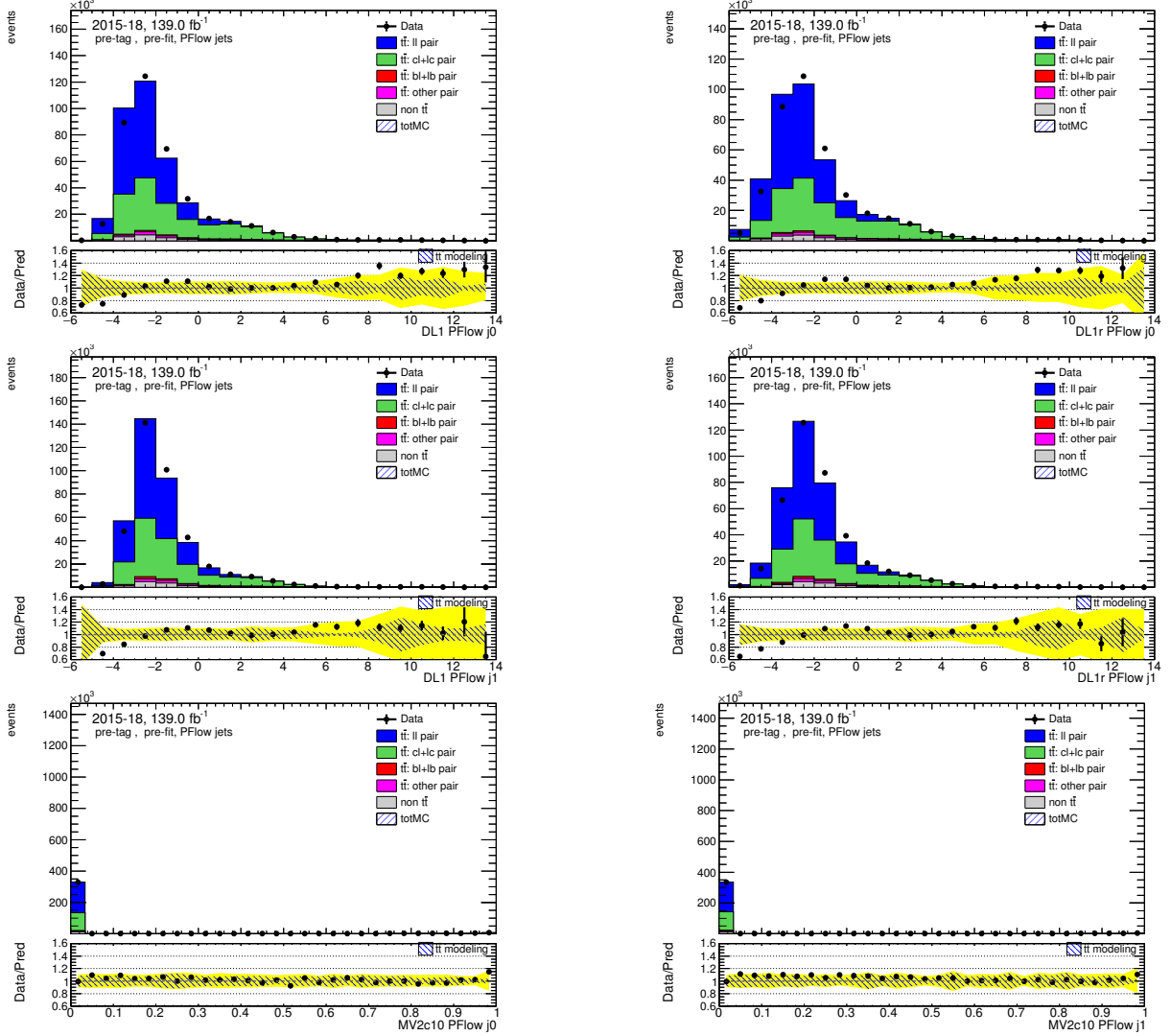
Figure 9: Distributions of the DL1, DL1r and MV2c10 taggers output of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

2.6.3 Efficiencies and Scale Factors

The Dl1 and DL1r c-jet efficiencies and scale factors with systematics uncertainties are calculated with 4 different fixed cut working points: 60%, 70%, 77%, and 85%, for the p4060 derivation, in December 2020. The results are shown in Figure **??**, **??**, **??**, **??**, for the charm jets mistagging efficiency and scale factors for the PFlow jet collections and the VR Track jets collections. 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. The c-jets efficiencies and the scale factors are similar between the DL1 and DL1r as expected.
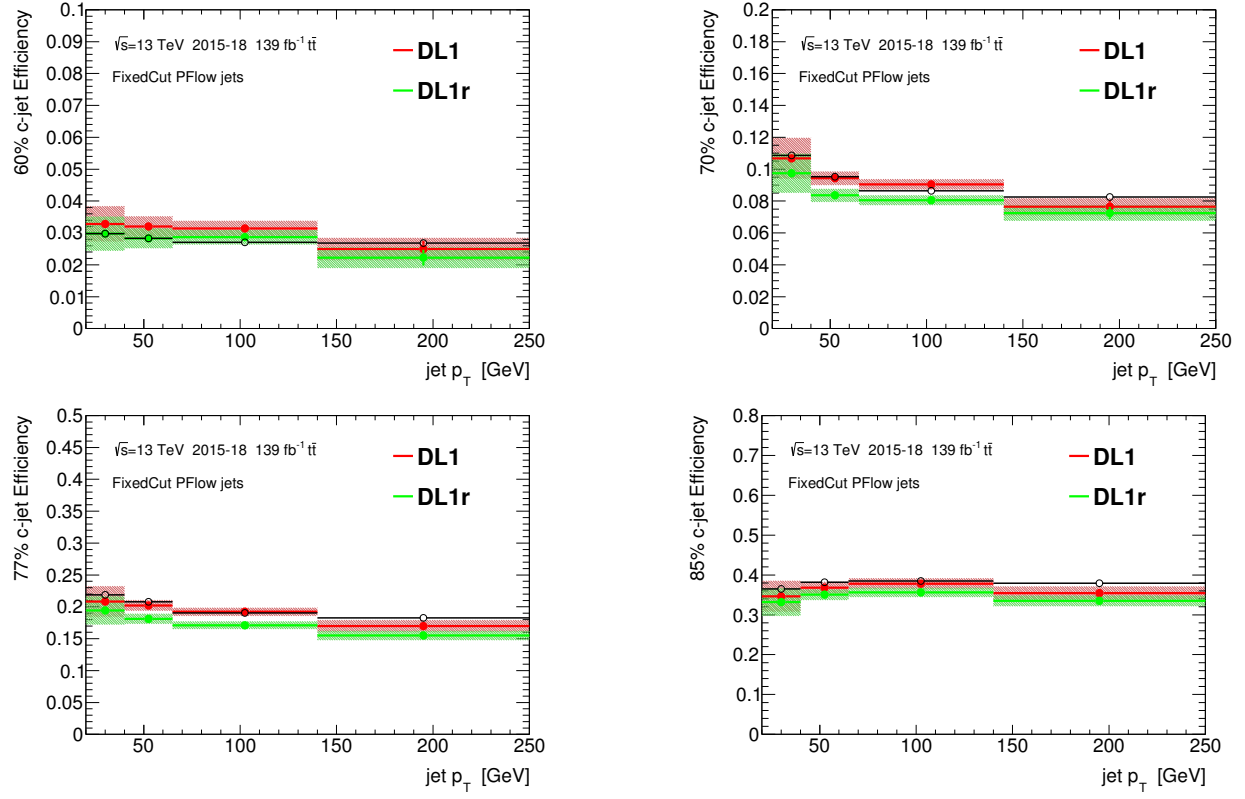
Figure 11: Charm-jet efficiencies of PFlow jets collection of derivation p4060 in December 2020, given for 4 different working points.
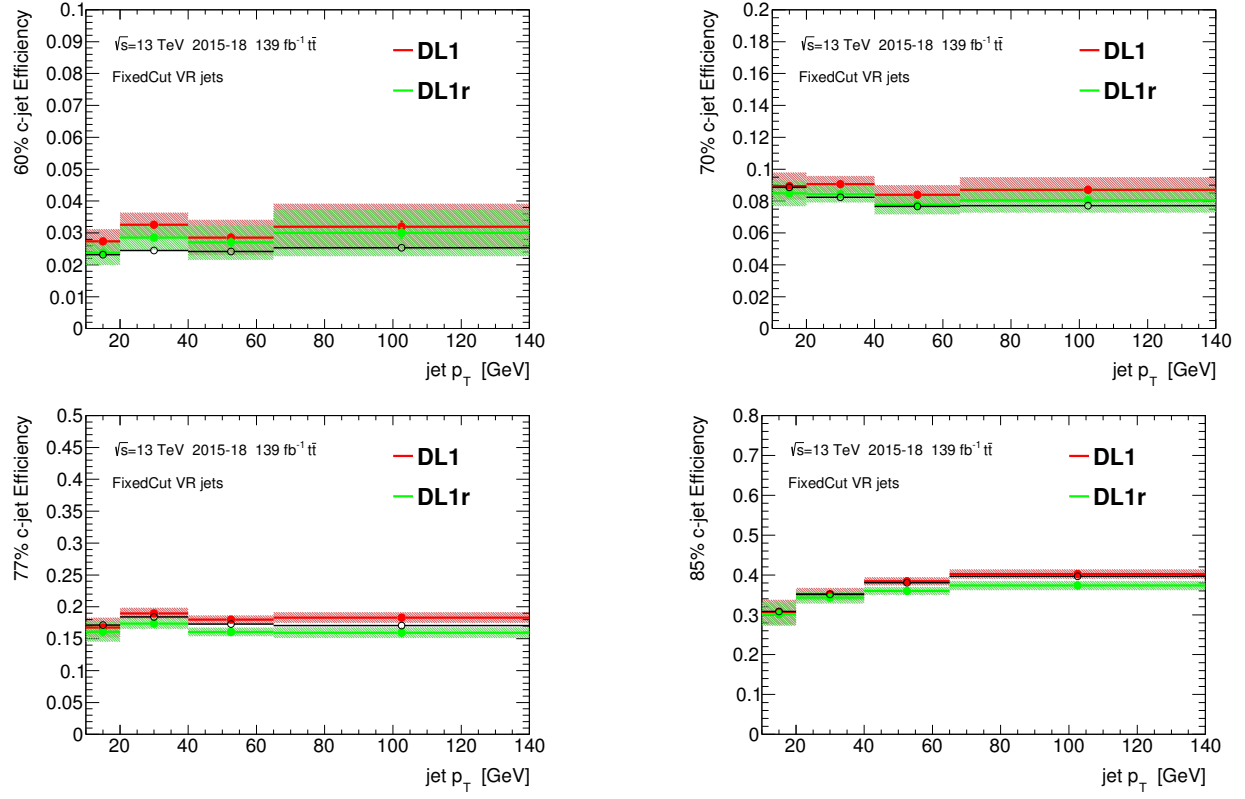
Figure 12: Charm-jet efficiencies of VR-Track jets collection of derivation p4060 in December 2020, given for 4 different working points.
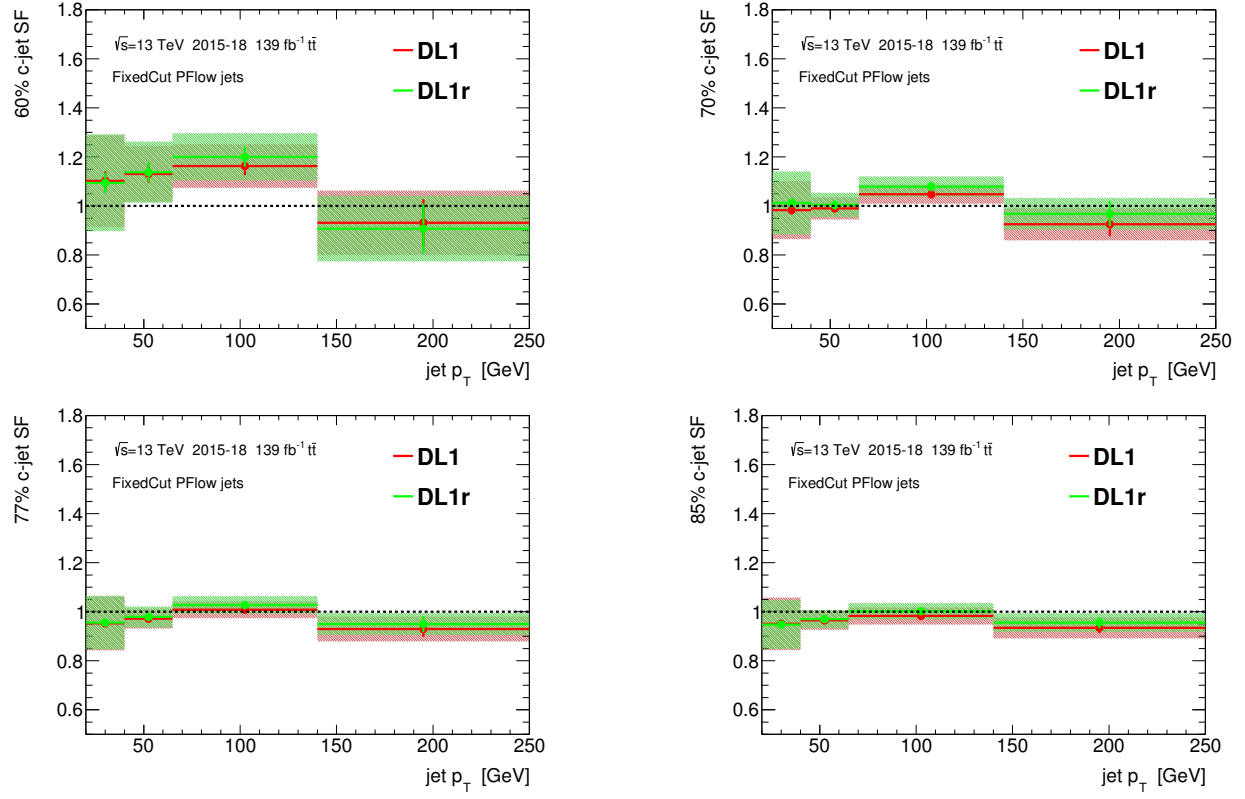
Figure 13: Charm-jet scale factors of PFlow jets collection of derivation p4060 in December 2020, given for 4 different working points.
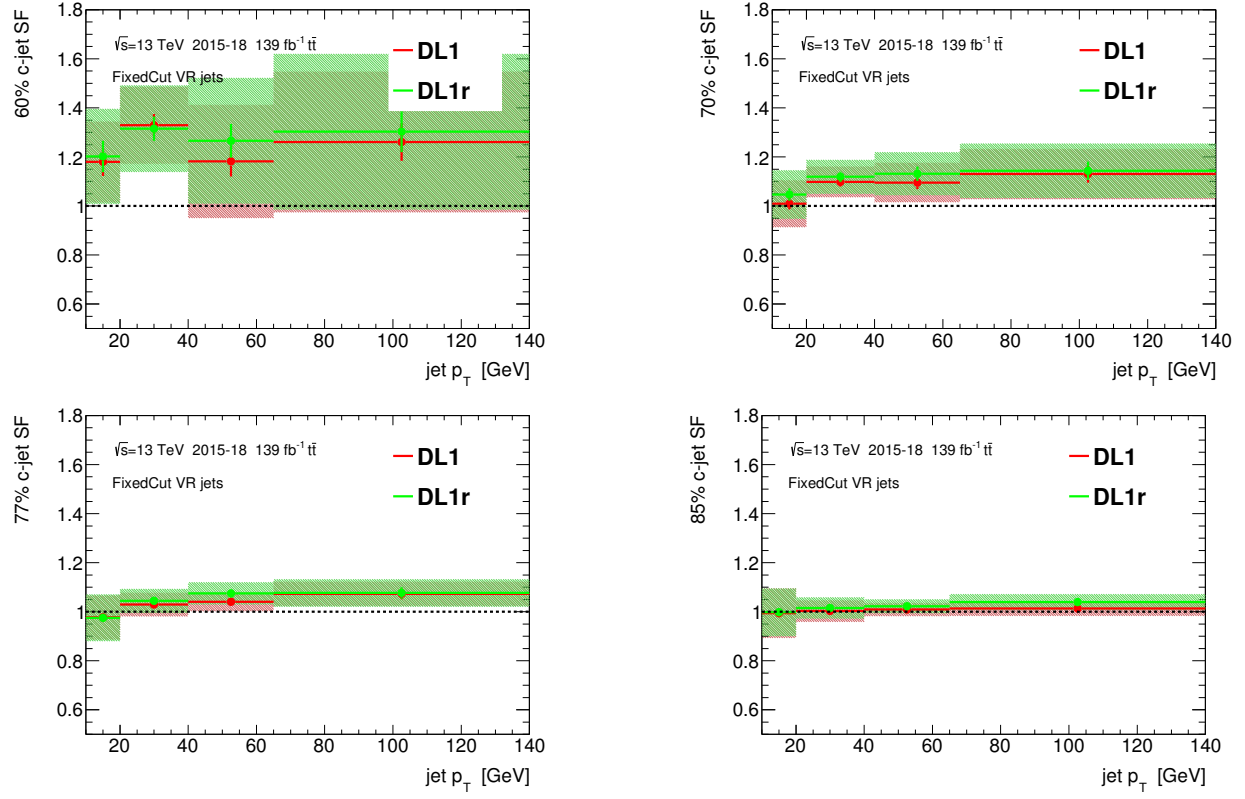
Figure 14: Charm-jet scale factors of VR-Track jets collection of derivation p4060 in December 2020, given for 4 different working points.

.1 Appendix

A Addition plots for the standard selection

Distribution_March/DataMC_J1_DL1-eps-converted-to.pdf
Distribution_March/DataMC_J1_DL1r-eps-converted-to.pdf

Distribution_March/DataMC_J1_eta-eps-converted-to.pdf
Distribution_March/DataMC_J1_pt-eps-converted-to.pdf

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.

Distribution_March_highpT/DataMC_J0_DL1-eps-convertedtto.pdf

Distribution_March_highpT/DataMC_J0_DL1r-eps-c

Distribution_March_highpT/DataMC_J0_eta-eps-convertedtto.pdf

Distribution_March_highpT/DataMC_J0_pt-eps-cor

Distribution_March_highpT/DataMC_Mbb-eps-convertedtto.pdf

Distribution_March_highpT/DataMC_mjj-eps-conve
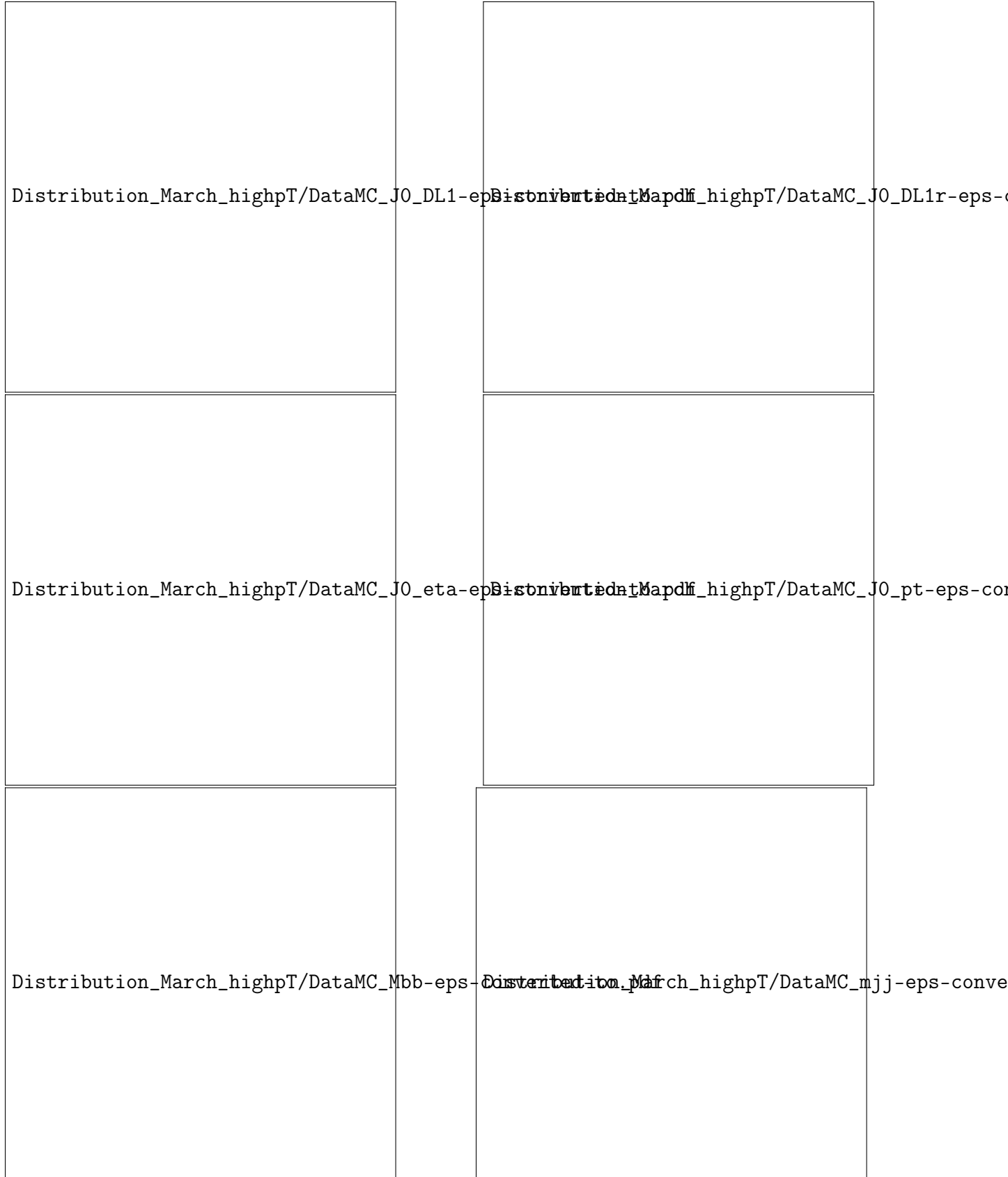
Figure 17: Various kinematic distributions of the combination of the standard selection and the high-p_T selection, before fitting or tagging, with stat-only uncertainties.
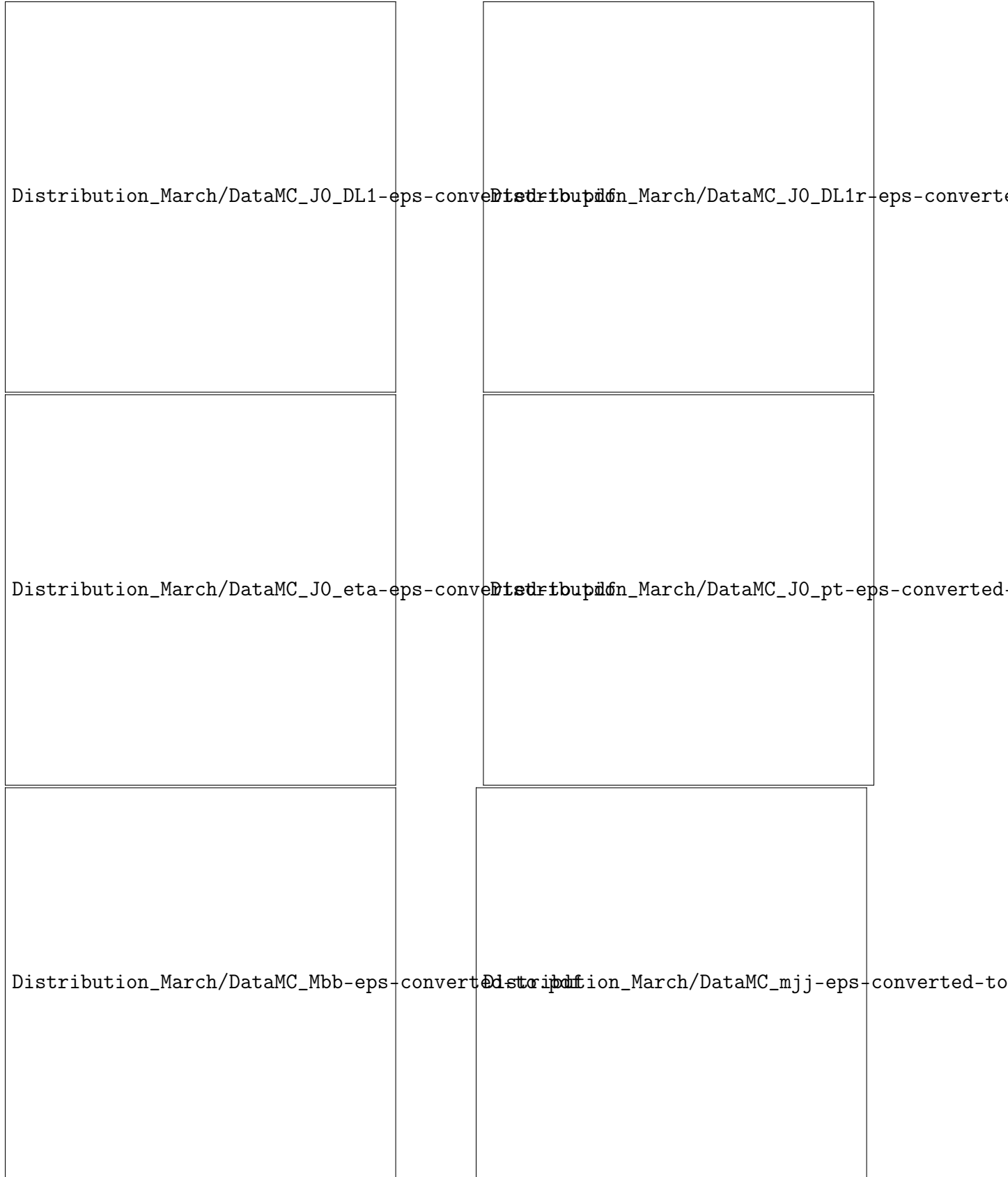
Systematic uncertainty
EG_RESOLUTION_ALL
MUON_ID
MUON_MS
MET_SoftTrk_ResoPara
MET_SoftTrk_ResoPerp
MET_SoftTrk_ScaleDown
MET_SoftTrk_ScaleUp
JET_Pileup_OffsetNPV
JET_Pileup_RhoTopology
JET_EffectiveNP_Modelling1
JET_EffectiveNP_Modelling2
JET_EffectiveNP_Modelling3
JET_EffectiveNP_Modelling4
JET_EffectiveNP_Statistical4
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 7: List of experimental systematics.
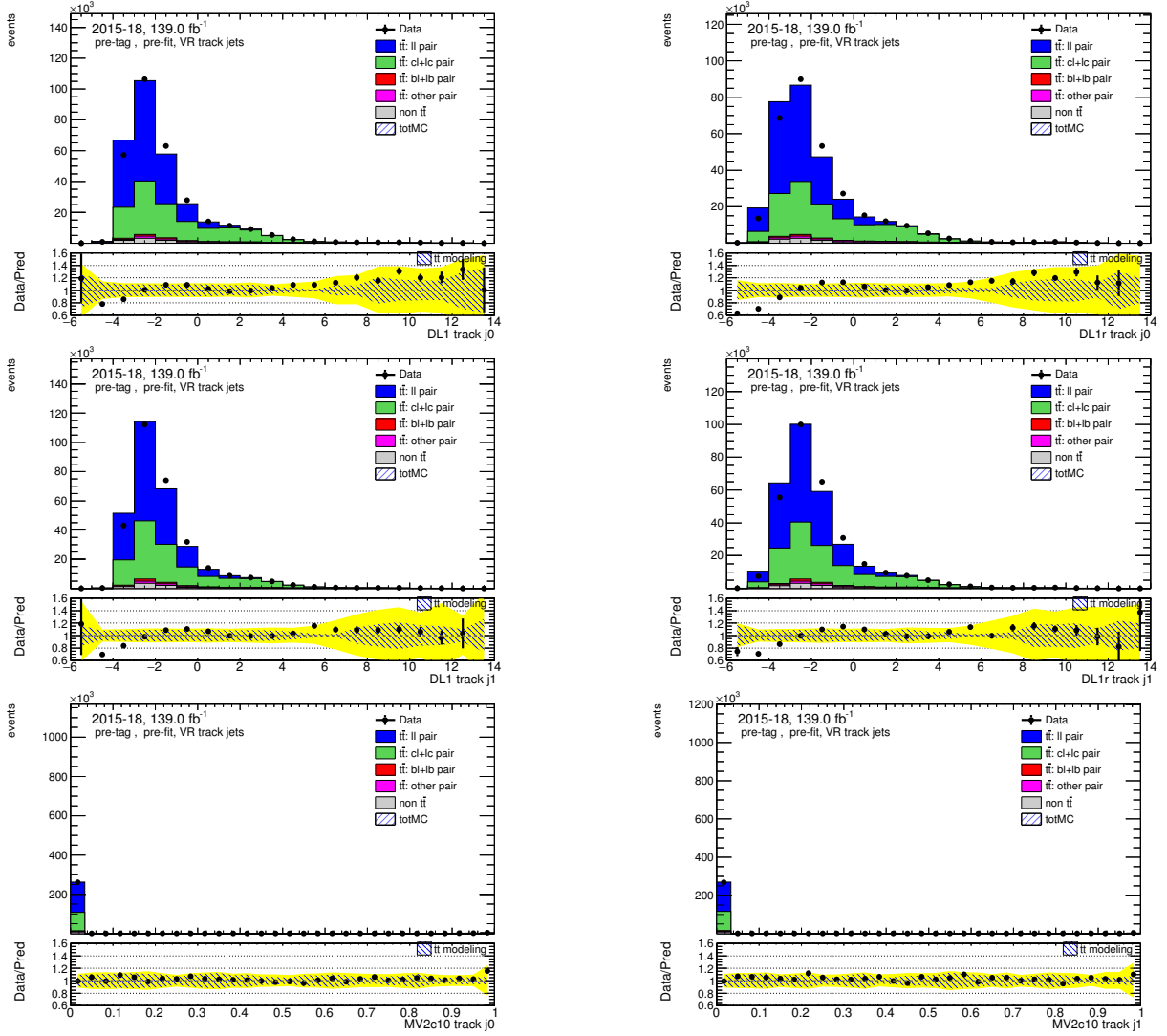
Figure 10: Distributions of the DL1, DL1r and MV2c10 taggers output of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

Distribution_March/DataMC_J0_DL1-eps-converted Distribution_March/DataMC_J0_DL1r-eps-converte

Distribution_March/DataMC_J0_eta-eps-converted Distribution_March/DataMC_J0_pt-eps-converted-

Distribution_March/DataMC_Mbb-eps-converted-to Distribution_March/DataMC_mjj-eps-converted-to

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.
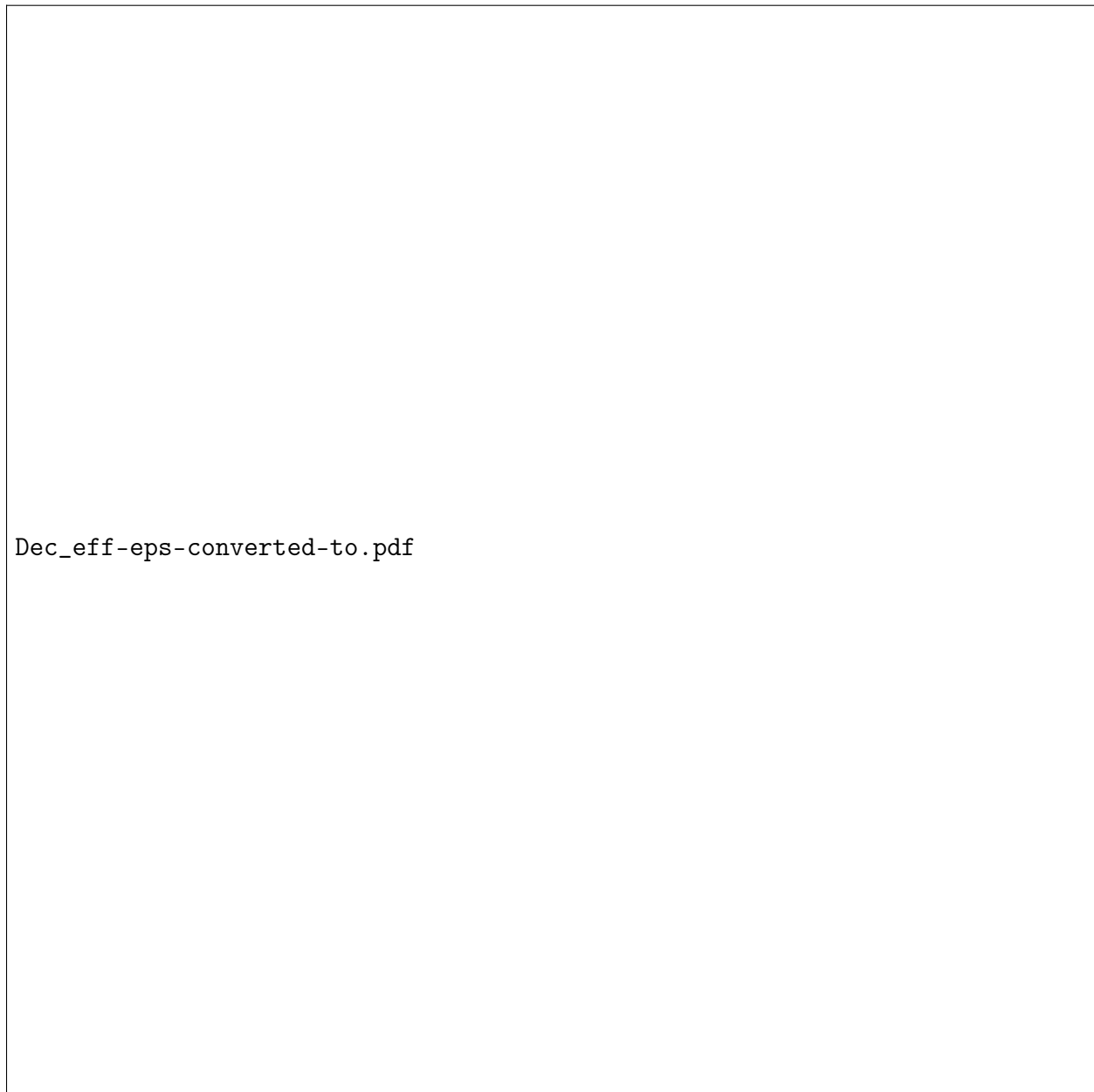
Figure 18: Calibration of derivation p3970 in December 2019, given for 4 different working points.

Dec-eps-converted-to.pdf

Figure 19: Calibration result of derivation p3970 in December 2019, given for 4 different working points.
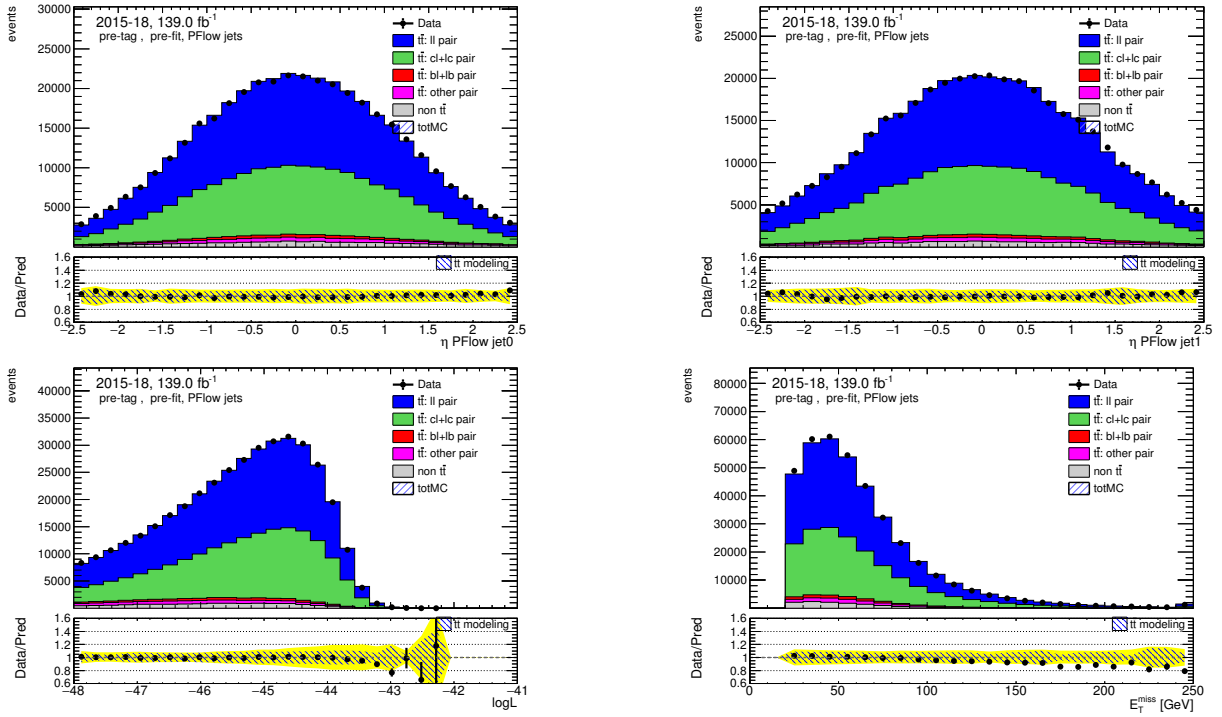
Figure 20: Distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.
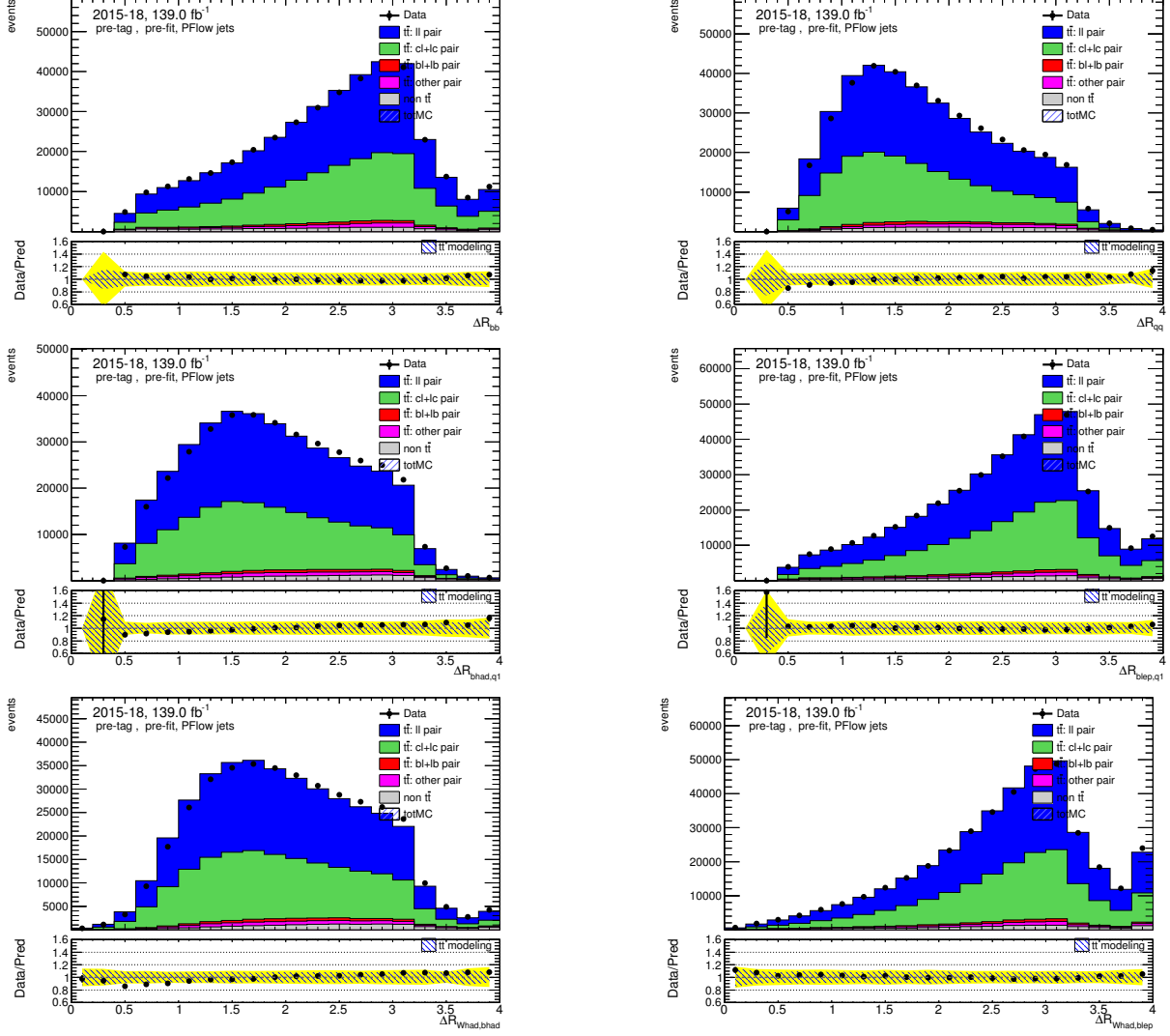
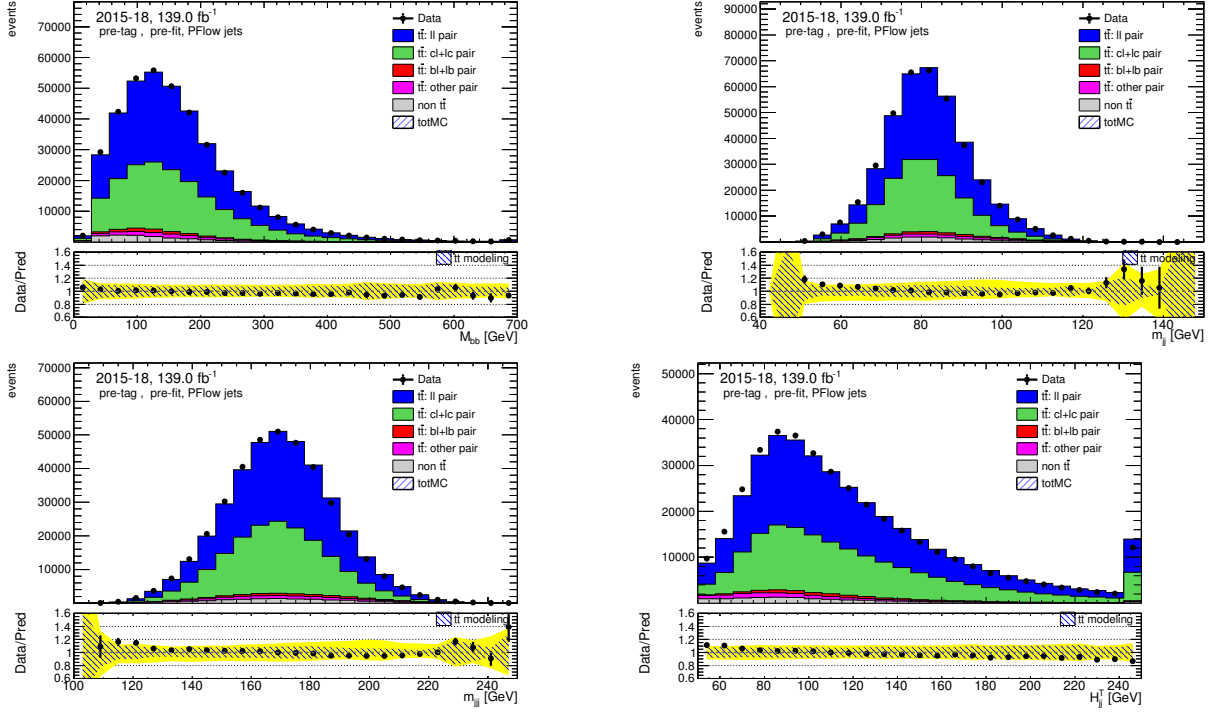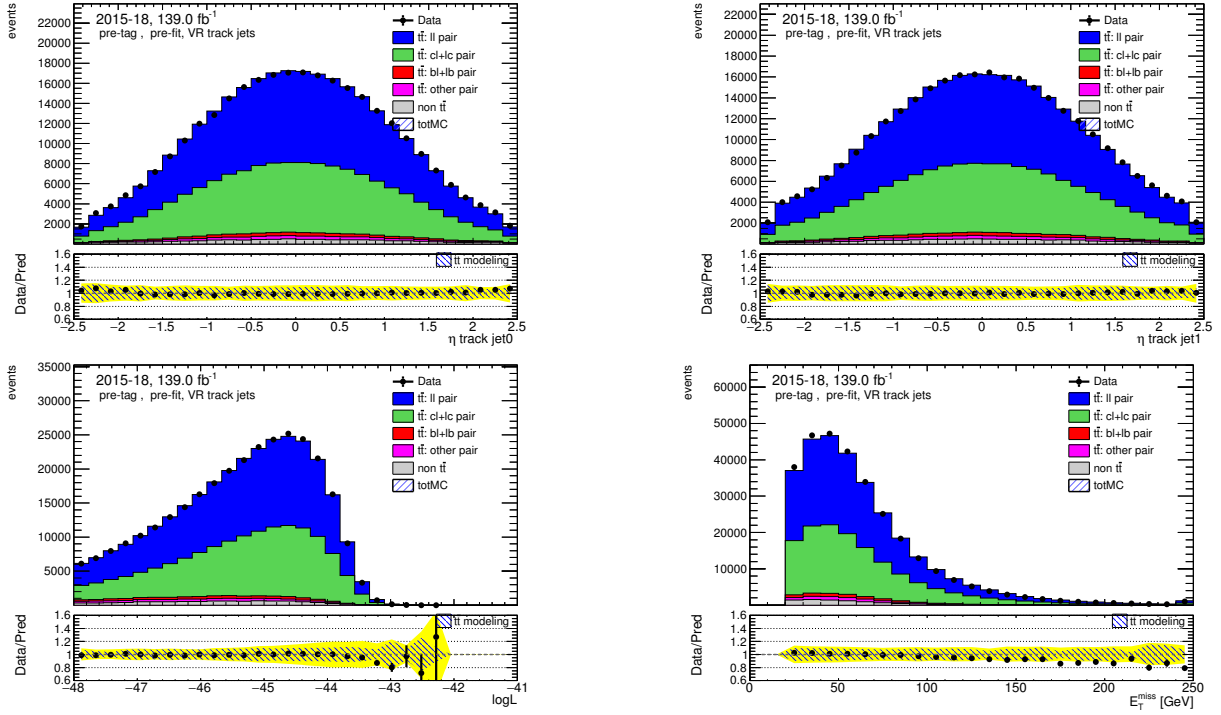Figure 21: Distributions of angle related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.

Figure 22: 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 23: 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.
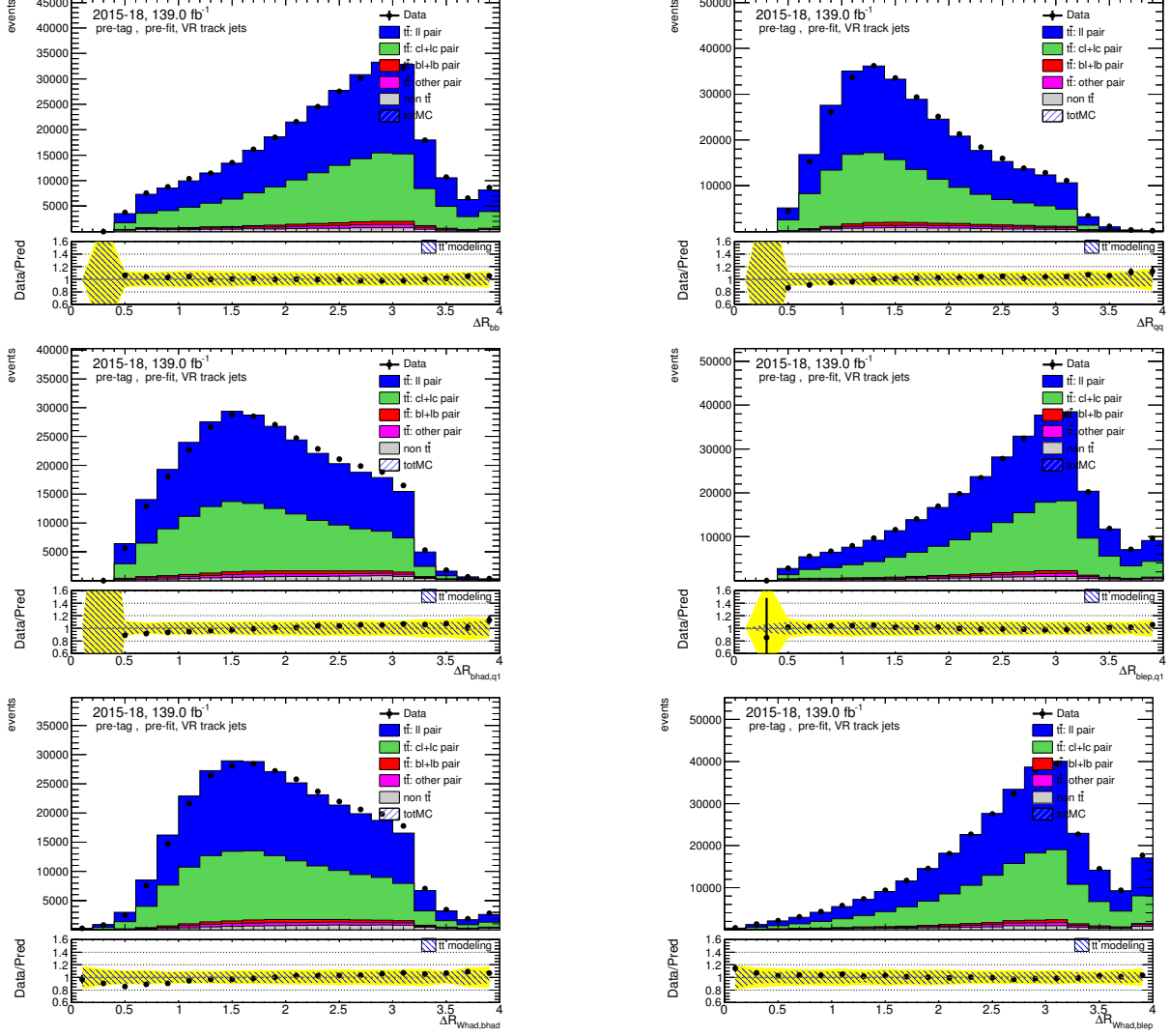
Figure 24: Distributions of angle related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with full uncertainties.
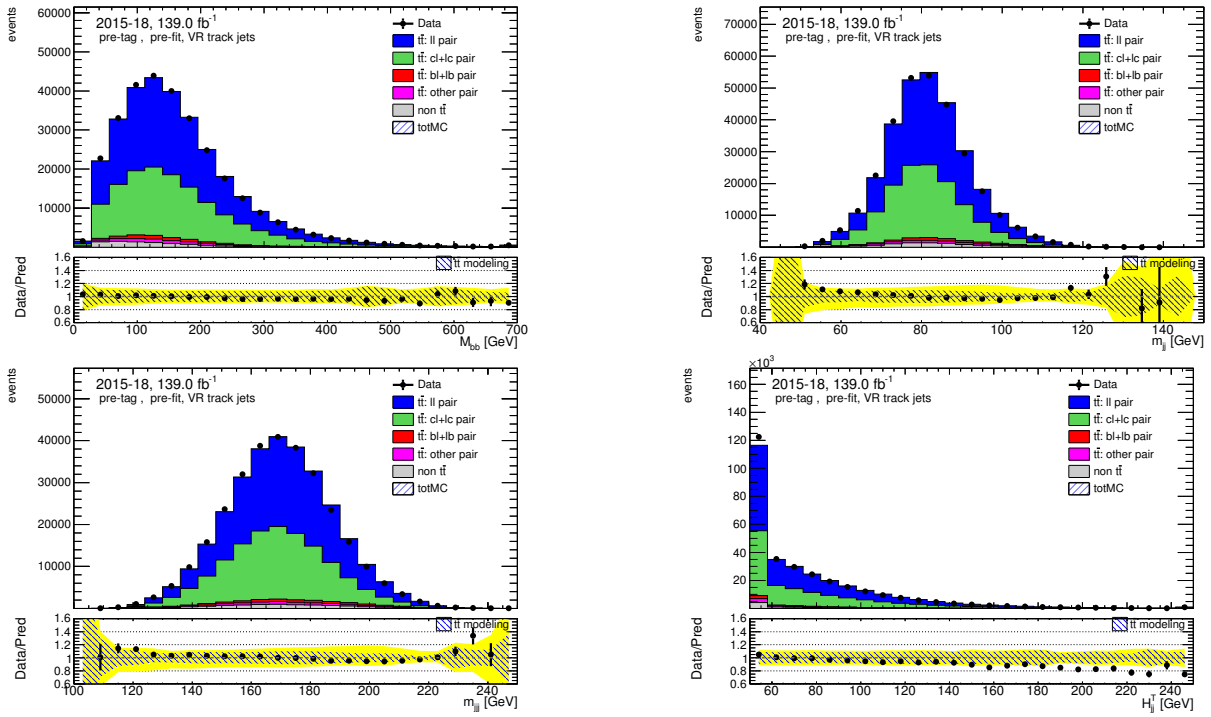
Figure 25: Distributions of mass related variables of the combination of the standard selection and the high-p_T selection, before fitting or tagging with stat-only uncertainties.