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

August 30, 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

The Large Hadron Collider [1] is the world's largest and most powerful particle accelerator. It started in 2008 and remains its crucial role in the many accelerators at CERN and in the world. The main body of the collider consists of a ring tunnel of perimeter of 26.7 km, lies beneath the France-Switzerland near Geneva,with superconducting magnets along the tunnel to keep the particle beam in direction and a large number of accelerating structures to boost the beam to the desired energy.

Inside the tunnel, two beams of particles travelling at close to the speed of light in opposite direction are made to collide. These two beams are kept in separate beam pipes, cooled to $-271.3°C$ ($1.9\ K$) with liquid helium distributed by dedicated system, and ultra-high vacuum, a vacuum thinner than interstellar void, matianed for 48 km of low-temperature section and 6 km of room-temperature section.

Thousands of magnets are used to direct the beams along the beam pipe, either to bend the beams or to focus. The particles are so small that making them collide is akin to firing two needles 10 kilometers away and meet halfway.

All the controls for the accelerator, its services and technical infrastructure are located at the CERN Control Centre. From here, the beams inside the LHC are made to collide at four locations around the accelerator ring, corresponding to the positions of four particle detectors – ATLAS (A Toroidal LHC ApparatuS) [2], CMS [3] (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) [4] and LHCb (b stands for beauty) [5].

3.1.1 Design and performance

The LHC is a two-ring-superconducting-hadron accelerator and collider installed in the existing tunnel that was constructed between 1984 and 1989 for the CERN LEP machine. The LEP tunnel has eight straight sections and eight arcs and lies between 45 m and 170 m below the surface on a plane inclined at 1.4% sloping towards the Léman lake. Approximately 90% of its length is in molasse rock, which has excellent characteristics for this application, and 10% is in limestone under the Jura mountain. There are two transfer tunnels, each approxi-mately 2.5 km in length, linking the LHC to the CERN accelerator complex that acts as injector. As mentioned before, the beam piples are maintained in vacuum for low and high temperature section. For the low temperature section, the vacuum is achieved by pumping in 9000 m^3 of cryogenic gas, which later will be condensed and adhered to the surface of the beampipe. For the room temperature section, the vacuum is achieved by use of non-evaporable getter (NEG) that absorbs residue gas particles when heated. More residue is absorbed by ion pumper.

The proton-proton collider has advantages and disadvantages compared to a proton-anti-proton collider or an electron-positron collider. Two rings are needed to accommodate the two counter-rotation beams, unlike particle-antiparticle colliders that can have both beams sharing the same phase space in a single ring. However it would not be possible to to achieve such high luminosity using anti-proton beams.

In principle, the mass of the proton is much larger than the mass of the electron, the synchrotron radiation losses will be much smaller, and the long straight sections designed for compensate the losses (as designed in the LEP) can be reduced. However these sections are kept as the LEP has as a cost-effective solution. The tunnel in the arcs has a finished internal diameter of 3.7 m.

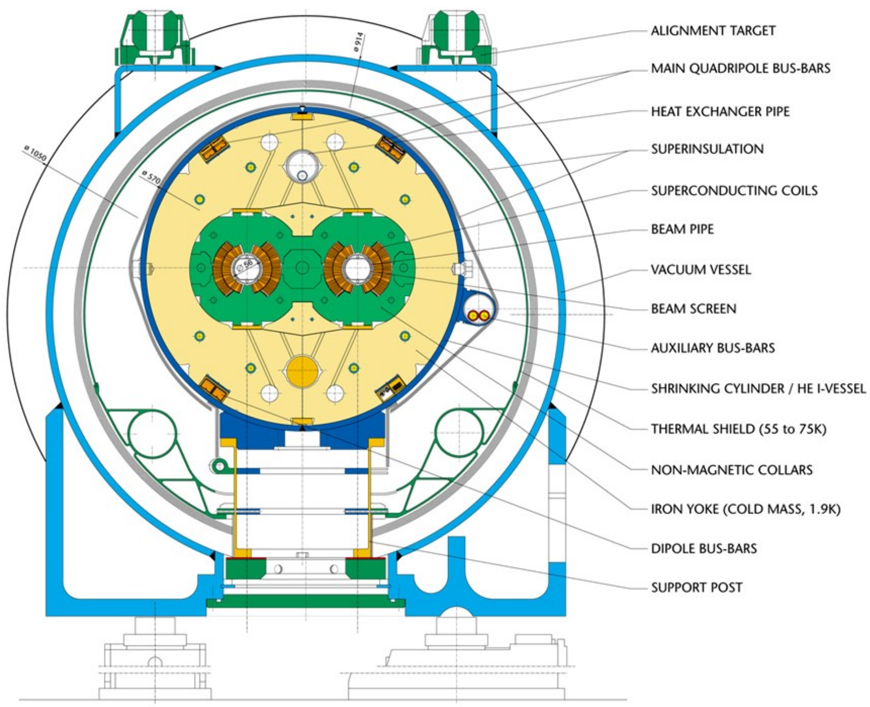


Figure 1: Double-bore magnet configuration of the LHC superconducting magnets [6].

Due to the technical difficulties to install two separate rings in such small space, LHC adopted the twin-bore magnet design [6], as shown in Figure 1. It was proposed by John

Blewett at the Brookhaven laboratory in 1971 first for cost consideration [7], but in the case of the LHC the overriding reason for adopting this solution is the lack of space in the tunnel.

The aim of the LHC is to reveal the physics beyond the Standard Model with centre of mass collision energies of up to 14 TeV. The number of events per second generated in the LHC collisionsis given by:

$$N_{event} = L\sigma_{event},$$

where $\sigma_e vent$ is the cross section for the event under study and L the machine luminosity. The machine luminosity depends on the beam parameters and can be written for a Gaussian beam distribution as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta*} F,$$

where N_b refers to the number of particles per bunch, n_b number of bunches per beam, f_{rev} revolution frequency, γ_r relativistic gamma factor, ϵ_r normalized transverse beam emittance, $\beta*$ beta function at the collision point which describes the size of the beam, and F refers to the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP).

The two high luminosity experiments, ATLAS and CMS are both aiming at a peak luminosity of $L = 10^{34} cm^2 s^1$ for proton operation. The two low luminosity experiments: LHCB for B-physics, is aiming at a peak luminosity of $L = 10^{32} cm^2 s^1$, and the dedicated ion experiment, ALICE, is aiming at apeak luminosity of $L = 10^{27} cm^2 s^1$ for nominal lead-lead ion operation.

The proton beams are designed to carry energy of the order of a few TeV. To reach such high energy, a series of acceleration is required for the beams before entering the LHC ring. The protons are supplied by the injector chain Linac2 — Proton Synchrotron Booster (PSB) — Proton Synchrotron (PS) — Super Proton Synchrotron (SPS), as shown in Figure 2, reaching an energy of 450 GeV when leaving the SPS. Each proton beam contains 2808 "bunches" of approximately 1.15×10^{11} protons, arranged in "trains" of bunches with 72 bunches each "carriage". Inside the "carriage", each beam has a spacing of 24.95 ns and between each "carriage" there is a gap of $\tilde{3}20$ ns. The beams are required to have well defined transverse and longitudinal emittance.

3.1.2 Runs and results

Following the downtime after an incident in one of the main dipole circuits during the first commissioning in 2008 [8], the operation restarted at lower beam energy to minimize the risk. Therefore, the first proton run (2010-2013) [9] was carried out at 3.5–4 TeV (centre of mass energy 7-8 TeV). Furthermore, a bunch spacing of 50 ns was used instead of the nominal 25 ns. This implied fewer bunches with larger intensity and hence a high peak luminosity (0.8×10^{34} cm^{-2}s^{-1} still smaller than the nominal 10^{34} cm^{-2}s^{-1} luminosity) but

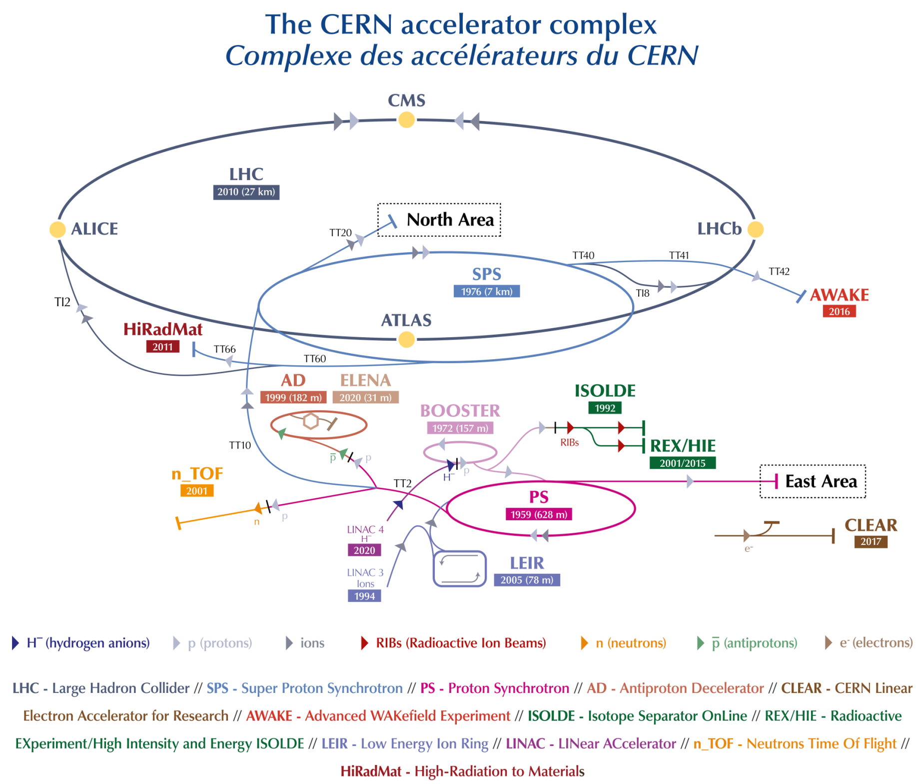


Figure 2: The LHC is the last ring (dark blue line) in a complex chain of particle accelerators. The smaller machines are used in a chain to help boost the particles to their final energies and provide beams to a whole set of smaller experiments.

larger than nominal pileup. Run 1 resulted in about 30 $fb^{-1}$of proton data and important physics results, most notably the discovery of the Higgs boson [10], [11]. Run 1 was followed by a long shutdown (LS1, 2013–2014) with a large number of consolidation and upgrade activities [12]. The bus-bar splices between the superconducting magnets were improved, in order to make sure that the LHC could operate at higher energy without risk of repeating the 2008 incident.

Run 2 (2016-2018) was carried out at 6.5 TeV (center of mass energy 13 TeV) [13]. As shown in Figure 3, out of the 156 $fb^{-1}$of data LHC has delivered, the ATLAS detector has recorded 147 $fb^{-1}$and 139 $\mathrm{fb}^{-1}$of data is certified to be good quality data. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity is slightly smaller than the delivered luminosity, due to the inefficiency of the Data Acquisition (DAQ) and the so called “warm start”: when the stable beam flag is raised, the tracking detectors undergo a ramp of the high-voltage and, for the pixel system, turning on the pre-amplifiers. More details of the ATLAS detector can be

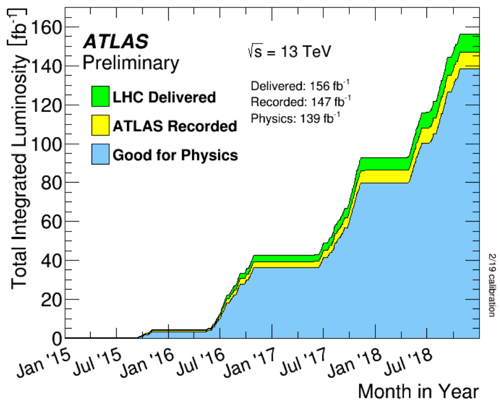


Figure 3: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in Run 2.

found in the following sections. The recorded data is checked carefully to exclude possible hardware or software issues. This is achieved by monitoring detector-level quantities and reconstructed collision event characteristics at key stages of the data processing chain. This procedure led to high efficiency of good quality data: 95.6% [14].

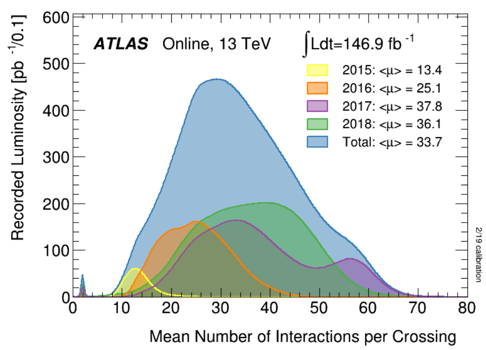


Figure 4: Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the Run 2 data. All data recorded by ATLAS during stable beams is shown, and the integrated luminosity and the mean mu value are given in the figure.

In this thesis the 139 fb^{-1}data recorded by the ATLAS detector of Run 2 is used. The nominal bunch spacing of 25 ns was used, with slightly less bunches (2500) each beam. The LHC experts have continually improved the running scenario to increase the luminosity, and during Run 2 the luminosity surpassed the designed luminosity by a factor of 2. As well as improving the instantaneous luminosity, the availability of the machine was dramatically improved during Run 2 which is an important factor enabling the high efficiency of good quality data as mentioned above. During Run 2, the machine was providing physics collisions during 50% of the allocated physics time, which is very impressive for a super conducting collider. An important parameter for the LHC experiments is the pileup, which is determined by the luminosity per bunch, and is a measure of the number of inelastic pp interactions that occur per bunch crossing.

Higher pileup gives more luminosity (for a fixed number of bunches) but makes physics analysis more difficult due to the signals in the detector from the additional interactions. The distribution of the recorded luminosity over the pileup is shown in Figure 4. It's a challenging task for the trigger and reconstruction algorithms to achieve robustness under such high pileup condition.

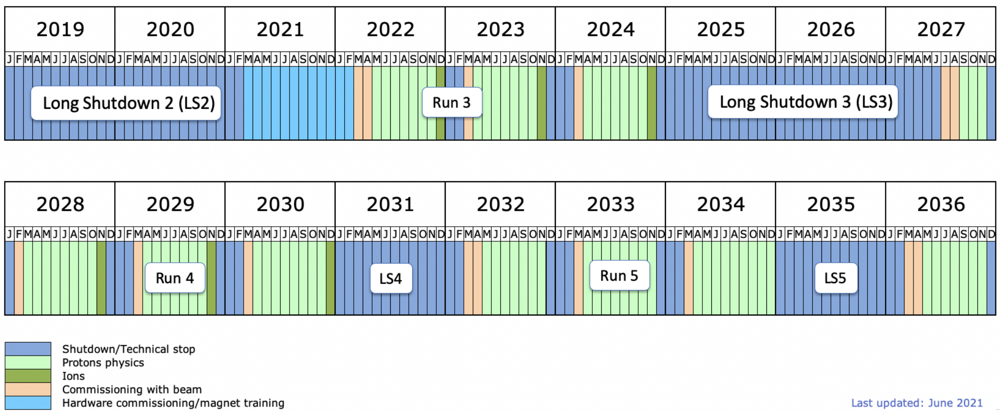


Figure 5: Longer term LHC operation schedule.

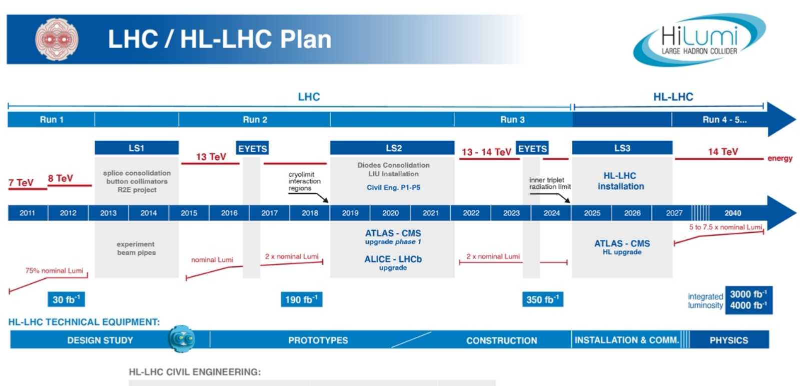


Figure 6: LHC operation schedule and luminosity targets.

The LHC operation and shutdown schedule is shown in Figure 5 and Figure 6. The operation of CERN's accelerators is subject to scheduled shutdowns to allow important repair and upgrade work to take place. The present shutdown, LS2, is devoted to preparations for Run 3 of the LHC, which will have an integrated luminosity equal to the two previous runs combined, and for the High-Luminosity LHC (HL-LHC), the successor to the LHC, which will begin operation at the end of 2027, and eventually generate 10 times the integrated luminosity to all Run 1, 2, 3 combined!

The LS2 schedule has had to be modified due to the COVID-19 pandemic, which the new schedule anticipates that the first test beams will circulate in the LHC at the end of September 2021, four months later than the date planned before the COVID-19 crisis, to give the LHC's main experiments – ATLAS, CMS, ALICE and LHCb – time to prepare their own upgrade. Run 3 of the LHC will begin at the start of March 2022. The third long shutdown (LS3) will begin at the start of 2025 and end in mid-2027. This is when the equipment for the HL-LHC and its experiments will be installed.

3.2 The ATLAS Detector

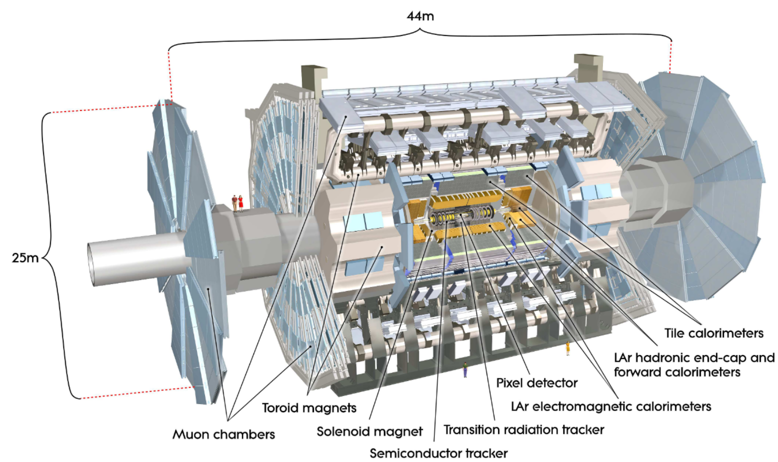


Figure 7: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes.

ATLAS is one of the two general purpose detectors built for probing proton-proton collision. This detector represents the work of a large collaboration of several thousand physicists, engineers, technicians, and students over a period of fifteen years of dedicated design, development, fabrication, and installation. A overall layout of the detector is shown in Figure 7.

The high interaction rates, radiation doses, particle multiplicities and energies, as well as the requirements for precision measurements have set strigent standards on the design of the ATLAS detector. Therefore the ATLAS detector is designed to fufiled the following requirements:

- The detector requires fast, radiation-resistant electronics and sensor elements and high detector granularity to meet the requirements due to the high frequency of collision, high particle fluxes and high radiation environment of the detector.

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \bigoplus 1\%$	± 2.5	None
EM calorimetry	$\sigma_E/E = 10\%/E \bigoplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap forward	$\sigma_E/E = 50\%/E \bigoplus 3\%$ $\sigma_E/E = 100\%/E \bigoplus 10\%$	± 3.2 $3.1 < \lvert\eta\rvert < 4.9$	± 3.2 $3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_{\mathrm{T}}=$ 1 TeV	± 2.7	± 2.4

Table 1: General performance goals of the ATLAS detector. Note that, for high-p_{T} muons, the muon-spectrometer performance is independent of the inner-detector system. The units for E and p_{T} are in GeV [2].

- Due to the geometry of the detector, to collect most of the collision data, large acceptance in pseudorapidity (which will be defined in the next subsection: 3.2.1) with almost full azimuthal angle coverage is required.
- Good energy and momentum resolution are required to enable accurate physical objects reconstruction. The high resolution of energy can be obtained with very good electromagnetic (EM) calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missingtransverse energy measurements.
- For offline tagging of τ-leptons and b-jets, vertex detectors close to the interaction region are required to observe secondary vertices.
- Good muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high p_{T} muons are fundamental requirements.
- Triggering with high efficiency on p_{T} objects with sufficient background rejection is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest.

The main performance goals of the detector are listed in Table 1.

3.2.1 Coordinate system

The ATLAS coordinate system is a right-handed Cartesian system with the nominal interaction point defined as the origin of the coordinate system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards, as shown in Figure 8.

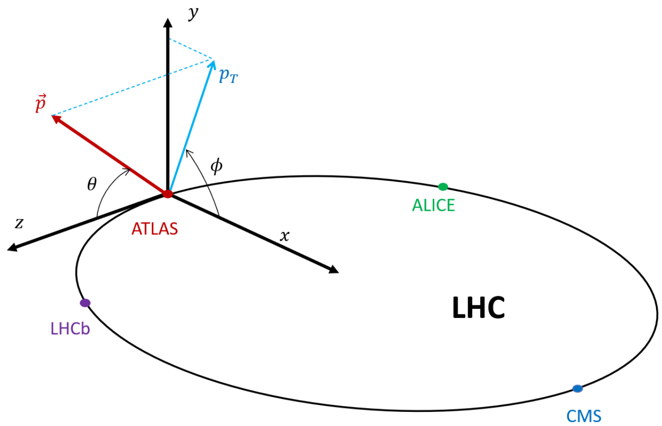


Figure 8: Illustration of the coordinate systems used at the ATLAS experiment in the geographical context of the LHC.

The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. In high energy physics, it's more common to use the pseudorapidity instead of the polar angle θ, defined as:

$$\eta = -ln\ tan(\theta/2).$$

In the case of highly relativistic particles (which is the comman case in high energy physics), the pseudorapidity approaches the rapidity,

$$y = 1/2ln[(E + p_z)/(E - p_z)],$$

with E the energy of the particle, m its mass and p_z the momentum along the z axis.

There are two main reasons for using pseudorapidity. The first reason is that the rapidity is invariant under Lorentz transformation, while capturing the characteristic of the particle direction of travel: $y \to \pm\infty$ when the particle is travelling close to the beam pipe (positive for along the beam pipe, negative for the opposite direction) and $y \to 0$ when p_z is small. While the second reason is that due to the limited angle coverage of the detector, it's usually hard to determine the total energy and the momentum along the z axis, especially when the direction of the particles are close to the beam pipe. While the pseudorapidity is determined only by the polar angle, which is much easier and faster to compute.

The transverse momentum p_{T}, the transverse energy E_T, and the missing transverse energy E^T_{miss} (MET) are defined in the x-y plane unless stated otherwise. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}.$$

3.2.2 Magnets

ATLAS has a unique hybrid system of four large superconducting magnets [15]. This magnetic system is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ. Figure 7 shows the general layout, the four main layers of detetors and the four superconducting magnets which provide the magnetic field over a volume of approximately 12000 m^3.

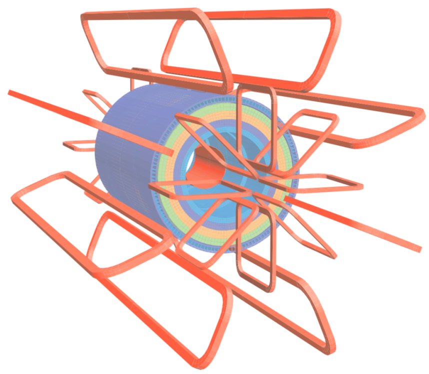


Figure 9: Geometry of magnet windings and tile calorimeter steel.

The spatial arrangement of the coil windings is shown in Figure 9. The ATLAS magnet system consists of two parts:

- a solenoid, which is aligned on the beam axis and provides a 2 T axial magnetic field in the z-direction. For the inner detector, while minimising the radiative thickness in front of the barrel electromagnetic calorimeter;
- a barrel toroid and two end-cap toroids, which produce a toroidal magnetic field of approximately 0.5 T and 1 T for the muon detectors in the central and end-cap regions, respectively. The barrel toroid generates the magnetic field in the central zone of the muon spectrometer, along the tangential direction of the circumferences centered on the z-axis (ϕ direction). The end-cap toroids are two smaller toroids designed to provide the magnetic field in the forward areas of the muon spectrometer.

Therefore in the ATLAS magnetic system, tracks in the inner detector is bended in the ϕ direction, while in the muon spectrometer in the θ direction.

3.2.3 Inner detector

The innder detector [16] consists of three parts: the pixel detector [17] and the insertable B-Layer (IBL) [18] (as one part), the silicon micro strip tracker (SCT) []ATLAS-TDR-04, the transition radiation tracker (TRT) [16]. Figure 10 shows a charge track traverses the sensors and structural elements. The track traverses successively the beryllium beam-pipe,

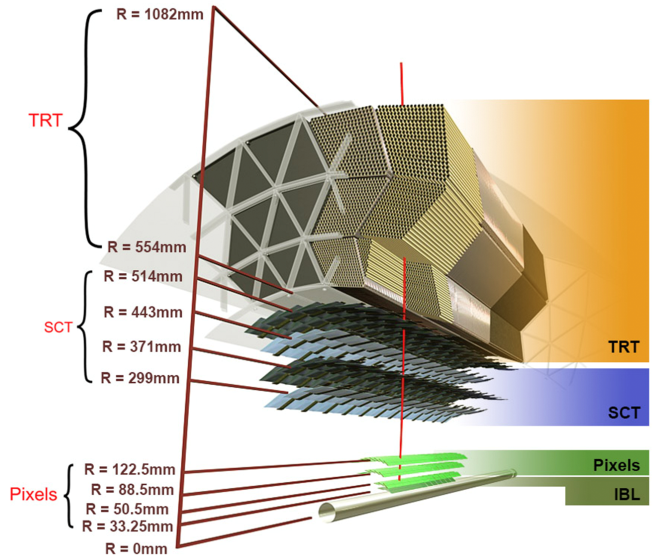


Figure 10: Cut-away view of the inner detector.

the four cylindrical silicon-pixel layers with individual sensor elements of $50 \times 400\ \mu m^2$, the four cylindrical double layers of barrel SCT of pitch 80 μm, and approximately 36 axial straws of 4 mm diameter contained in the barrel TRT modules within their support structure.

- **Pixel detector and IBL**
  The silicon pixel detector is the closest ATLAS component to the collision. It is composed of layers of silicon pixels and designed to have a very high granularity for reconstructing primary and secondary interaction vertices. The detector layers are formed of silicon sensor modules and in total there are approximately 92 million pixels (consequently, 92 million readout channels) in the system. It consists of three cylindrical layers in the barrel region, positioned at the radial distances of 50.5, 88.5 and 122.5 mm and on disks perpendicular to the beams in the end-caps at the longitudinal distances of 49.5, 58.0 and 65.0 mm. In 2014, during the first LHC long shutdown, a fourth pixel layer was installed inside the existing detector, the insertable B-Layer (IBL) at a radius of 33 mm from the beam axis. The new pixel layer provides an additional space point very close to the interaction point, which significantly improves the identification of jets coming from b-quark hadronisation (b jets).

- **SCT**
  The next constituent of the inner detector is the SCT, which consists of four layers of strips located axially on the beam direction in the barrel region and placed along the R-direction in the end-cap region. Each layer of strip is glued back to back with an angle of 40 mrad. The spatial resolution of the sensor is $\sigma_\phi = 17\ \mu m$ in the bending

direction ($R-\phi$) and $\sigma_\phi = 580\ \mu m$ in the z (barrel) and R (end-cap) direction.

- **TRT**
  The most external part of the inner detector is the TRT. It is a straw drift tube tracker, with additional electron identification capabilities from transition radiation. It consists of modules of 4 mm diameter straws, filled with a mixture of gas of 70% Xe, 27% CO2 and 3% O2 immersed in a polypropylene radiator. Up to 73 straw layers of straws are interleaved with fibres in the barrel and 160 starw planes are interleaved with foils in the end-cap. All charged tracks with $p_{\mathrm{T}} > 0.5$ GeV and $|\eta| < 2.0$ will traverse at least 36 straws, except in the barrel-end-cap transition region ($0.8 < |\eta| < 1.0$), where this number decreases to a minimum of 22 crossed straw. The TRT measures the track position only in the bending direction ($R-\phi$), with a spatial resolution of $\sigma_\phi = 130\ \mu m$. The TRT contributes significantly to the pattern recognition and momentum reconstruction, depsite the low resolution compared to the silicon tracker and the lack of a measurement along the z axis. This is the result of the large number of measurements and longer measured track length. In addition, the TRT provides extra ability to identify elecctrons due to the polypropylene fibres (foils) in the barrel (end-cap) emit photons when a charged particle traverses the boundaries of the material. These photons are then absorbed by the Xenon gas mixture, which the intensity depends on gamma factor of the traversing particle. This information can then be exploited for eletron/pion discrimination.

3.2.4 Calorimeter system

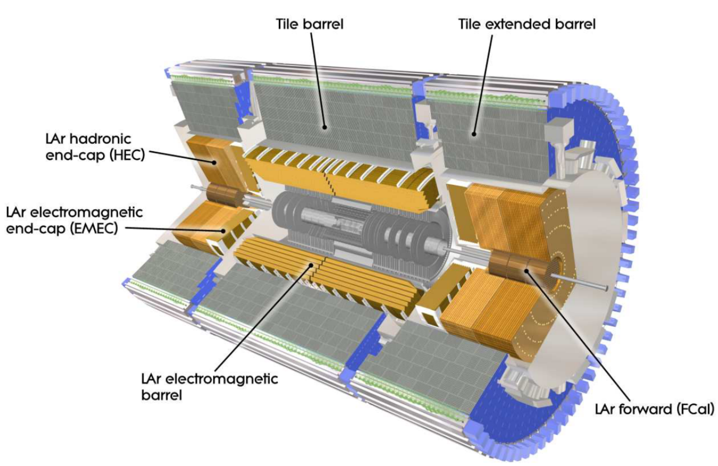


Figure 11: Cut-away view of the ATLAS calorimeter system.

The ATLAS calorimeters [19] are sampling calorimeters, meaning different material are used for the absorber and the active part. A cut-away view of the calorimeters is shown in Figure 11. These calorimeters cover the range $|\eta|< 4.9$, using different techniques suited to the widely varying requirements of the physics processes of interest and of the radiation environment over this large η-range. Over the η region matched to the inner detector, the fine granularity of the EM calorimeter is ideal for precision measurements of electrons and photons. The coarser granularity of the rest of the calorimeter is sufficient for the jet reconstruction and E^T_{miss} measurements.

The electromagnetic and hadronic showers must be contained in the calorimeter to avoid punch-through into the muon system. The calorimeter depth is hence an important design consideration. The total thickness of the EM calorimeter is greater than 22 radiation lengths (X_0) in the barrel and greater than 24 X_0 in the end-caps. The approximate 9.7 interaction lengths (λ) of active calorimeter in the barrel (10 λ in the end-caps) are adequate to provide good resolution for high-energy jets. The total thickness, including 1.3 λ from the outer support, is 11 λ at $\eta = 0$ and has been shown both by measurements and simulations to be sufficient to reduce punch-through well below the irreducible level of prompt or decay muons. Together with the large η-coverage, this thickness will also ensure a good E^T_{miss} measurement.

- **LAr electromagnetic calorimeter**
  The EM calorimeter is a lead-Liquid argon (lead-LAr) detector [20] with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete ϕ symmetry without azimuthal cracks. The lead thickness in the absorber plates has been optimised as a function of η in terms of EM calorimeter performance in energy resolution. The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in their own cryostat. The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at z = 0. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$.

  The calorimeter has three layers along the transverse direction: a pre-sampler with very high granularity in η, in order to reconstruct the neutral pions decaying to two photons and particles which already starts showering in the inner detector. The pre-sampler is followed by longer towers of relatively high granularity, which is the major part of detecting EM showers, and reponsible for measuring the η and ϕ coordinates of the particles. The last layer detects showers generated from particles other than electrons or photons that start showering inside the EM calorimeter before leaving it.

- **Tile calorimeter**
  The tile calorimeter [21] is placed directly outside the EM calorimeter envelope. Its barrel covers the region $|\eta| < 1.0$, and its two extended barrels the range $0.8 < |\eta| <$

1.7. It is a sampling calorimeter using steel as the absorber and scintillating tiles as the active material. The barrel and extended barrels are divided azimuthally into 64 modules. It is segmented in depth in three layers, approximately 1.5, 4.1 and 1.8 λ thick for the barrel and 1.5, 2.6, and 3.3 λ for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ at $\eta = 0$.

- **LAr hadronic end-cap calorimeter**
  The Hadronic End-cap Calorimeter (HEC) consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter and sharing the same LAr cryostats. The HEC covers the range of $1.5 < |\eta| < 3.2$, slightly overlapping with the forward calorimeter which will be described in the following paragraph (around $|\eta|= 3.1$) and the tile calorimeter ($|\eta| < 1.7$). This overlap is to reduce the drop in material density at the transition between the different calorimeters.

- **LAr forward calorimeter**
  The Forward Calorimeter (FCal) covers $3.1 < |\eta| < 4.9$ and is approximately 10 interaction lengths deep. It consists of three modules in each end-cap: the first, made of copper, is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions. Due to high particle fluxes and energies in the forward region, the calorimeter must contain relatively long showers in the small volume allowed by design constraints, and thus must be very dense.

3.2.5 Muon Spectrometer

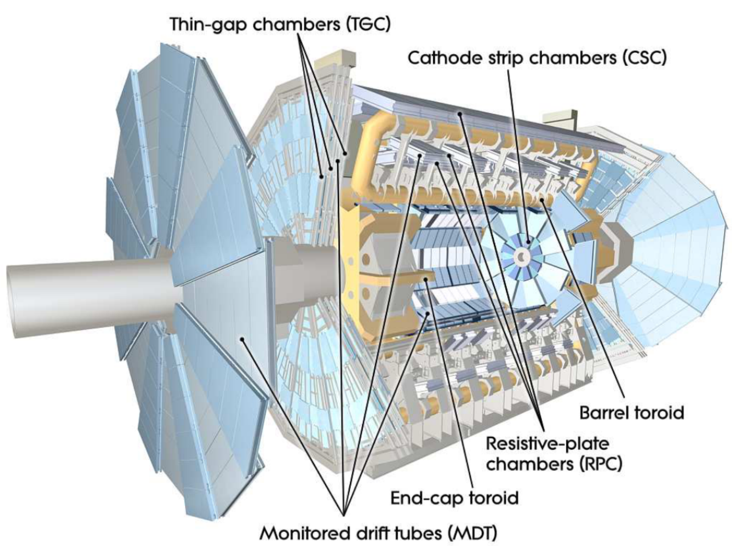


Figure 12: Cut-away view of the ATLAS muon spectrometer system.

The muon spectrometer (MS) [17] is the outermost and largest detector of ATLAS. The cut-away view of the MS is illustrated in Figure 12. It fully covers the calorimeter system and occupies a large part of the ATLAS cavern. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. The configuration of magnets provides a field mostly orthogonal to the muon trajectories, hence minimising the degradation of resolution due to multiple scattering.

The MS measures momentum of muons tracks depends on the momentum of the tracks. The resolution is typically 2–3% over most of the kinematic range apart from very high momenta, where it increases to about 10% at $p_{\mathrm{T}} = 1$ TeV. The MS consists of four subsystems which rely on four different gas detector technologies. Two of them, the resistive plate chambers (RPC) in the barrel region and the thin gap chambers (TGC) in the end-cap region, provide trigger signals, while the other two, the monitored drift tubes (MDT) and the cathode strip chambers (CSC) provide the momentum measurement. The MDT chambers provide high precision measurements in the bending direction over most of the detector acceptance while the CSC are used in the forward region where the particle flux is too high for the MDT chambers. The muon chambers are arranged in the barrel ($|\eta| < 1.05$) in three cylindrical layers around the beam axis, while in the end-cap regions ($1.05 < |\eta| < 2.7$) they are placed in three wheels.

3.2.6 Trigger system

As mentioned in section 3.1.2, the spacing of each bunch is 25 ns. This translates to a frequency of 40 MHz of bunch crossing. It will be difficult and not meaningful to keep all these events, which includes a lot of "uninteresting" physics events. Figure 13 shows a summary of the most dominant background processes production cross-sections. Many of these processes produce high multiplicity of jets and are not of experimental interest. Therefore, it's necessary to adopts a trigger system that make fast decisions whether to save the event or not. The ATLAS trigger system consists of two consecutive parts: the Level 1 trigger [22], followed by the Level 2 and the event filter, which together form the High Level Trigger (HLT) [23].

The L1 trigger searches for signatures from high-p_{T} muons, electrons/photons, jets, and τ-leptons decaying into hadrons. It also selects events with large MET and large total transverse energy. The L1 trigger uses reduced-granularity information from a subset of detectors: the RPC and TGC for high-p_{T} muons, and all the calorimeter sub-systems for

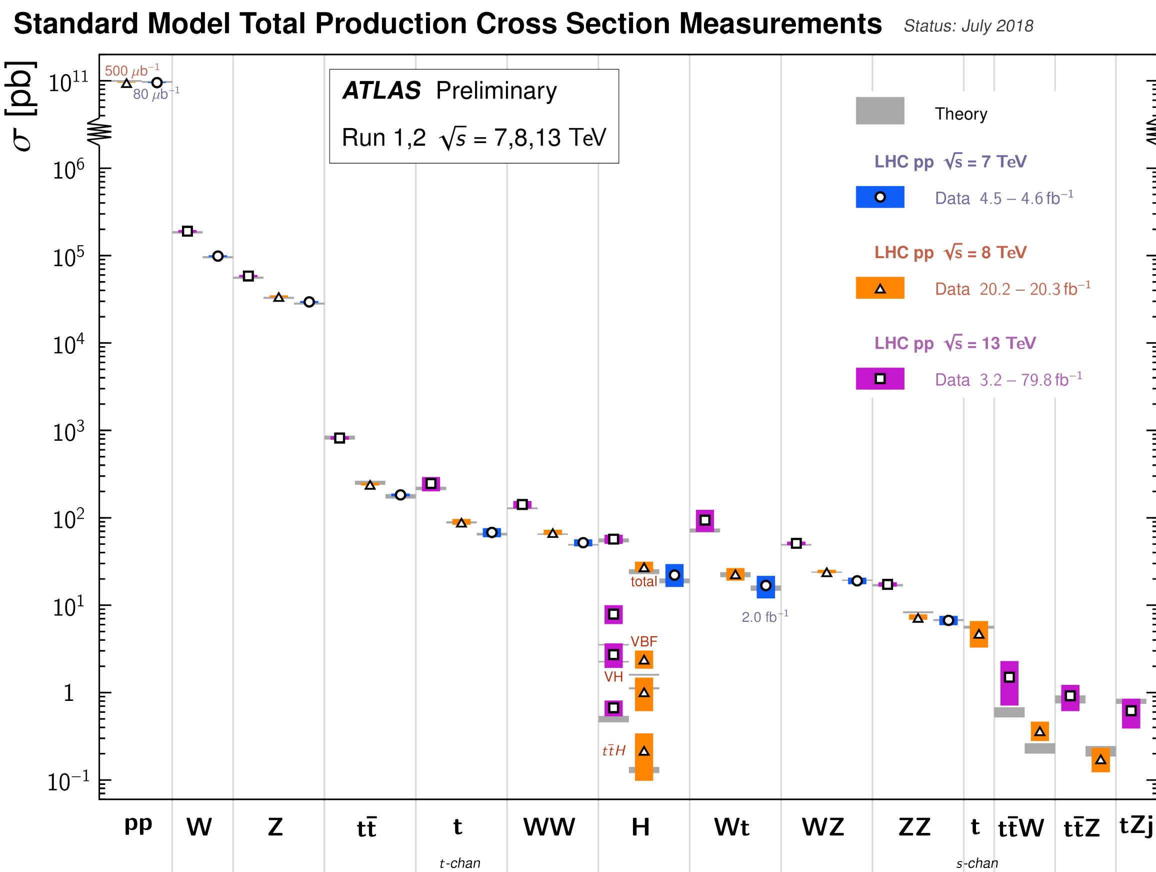


Figure 13: Plot showing the production cross-sections of the most dominant background processes at the LHC.

electromagnetic clusters, jets, τ-leptons, E^T_{miss} , and large total transverse energy. As a result, the L1 trigger reduces the event rate from 40 MHz to a maximum of 100 kHz. The decision is made by Central Trigger Processor (CTP), which operates on signals from dedicated hardware in the calorimeter and muon detector systems. The decision time, at under 2.5 μs, is faster than the ID can process events so ID information is omitted. For each data-taking period, the L1 trigger is loaded with a trigger menu, a list of up to 256 criteria used to determine whether an event is accepted. The trigger menus are designed to accomodate a broad physics programme, with high acceptance for both BSM searches and SM precision measurements.

The L1 trigger also uses detector information with reduced granularity to identify Regions of Interest (RoI) [24] in ϕ and η. The L2 selection is seeded by the RoI information provided by the L1 trigger over a dedicated data path. L2 selections use, at full granularity and precision, all the available detector data within the RoI's. The L2 menus are designed to reduce the trigger rate to approximately 3.5 kHz, with an event processing time of about 40 ms, averaged over all events. The final stage of the event selection is carried out by the event filter, which reduces the event rate to roughly 200 Hz. The selections are implemented using offline analysis procedures within an average event processing time of the order of four seconds.

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 [25, 26, 27]. It also plays an important role in the $HH \to bb\tau\tau$ searches presenting in Chapter 7.

The ATLAS Collaboration uses various algorithms to identify b jets [28], 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, as illustrated in Figure 14.

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 jetsvia 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 [28], have been improved and retuned for Run 2 [29]. Secondly, in order to maximise

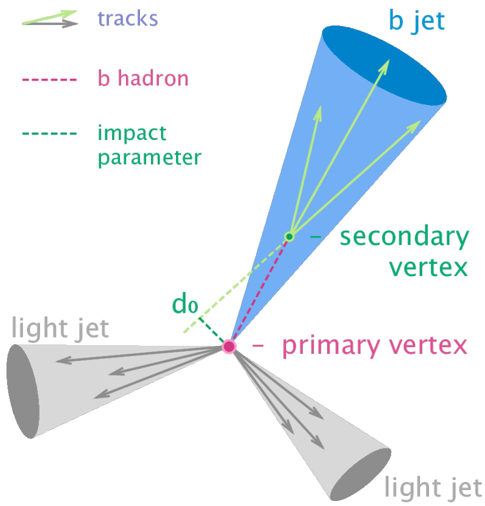


Figure 14: A diagram showning the b hadron decay initiated jets.

the b-tagging performance, the results of the low-level b-tagging algorithms are combined into high-level algorithms via multivariate classifiers.

The most performant algorithms presently in use in physics analyses at ATLAS are based on multivariate combinations of the available information (MV2) or additionally using a deep feed-forward neural network (DL1) [30, 31], as shown in Figure 15, where the performance is characterised by the probability of tagging a b jet (b jet tagging efficiency, ϵ_b) and the probability of mistakenly identifying a c jetor a light-flavour jet as a b jet labelled $\epsilon_c(\epsilon_l)$. In addition, the distribution of the output discriminant of the MV2 and DL1 tagger for b jet, c jet, and light-flavour jets in the $t\bar{t}$ simulated events are shown in Figure 16. 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 [32] and the DL1r tagger uses a Recurrent Neural Network Impact Parameter tagger (RNNIP) [31]. 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 $1/\epsilon_l$.

5.6 Missing transverse energy

5.7 Hadronically decaying τ lepton

6 Charm jet mis-tagging calibration

MC simulations are not able to model exactly 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 [29]. The

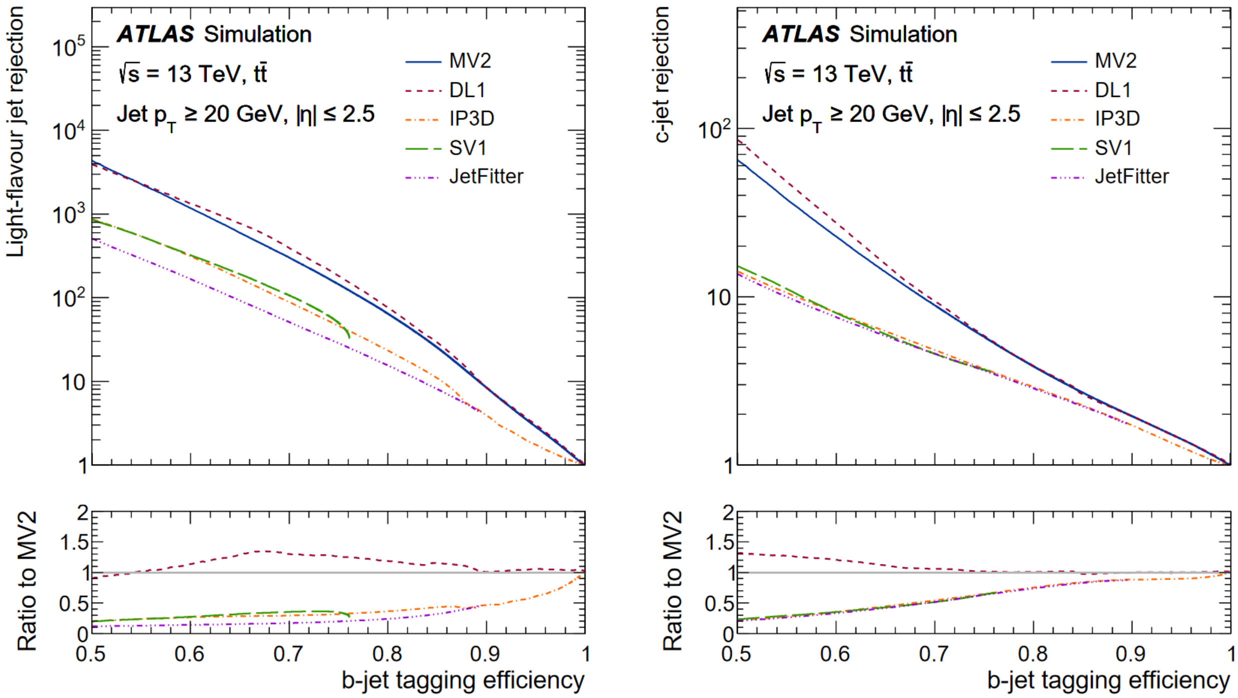


Figure 15: The light-flavour jet (left) and c jet (right) rejections versus the b jet tagging efficiency for the IP3D, SV1, JetFitter, MV2 and DL1 b-tagging algorithms evaluated on the $t\bar{t}$ events [29].

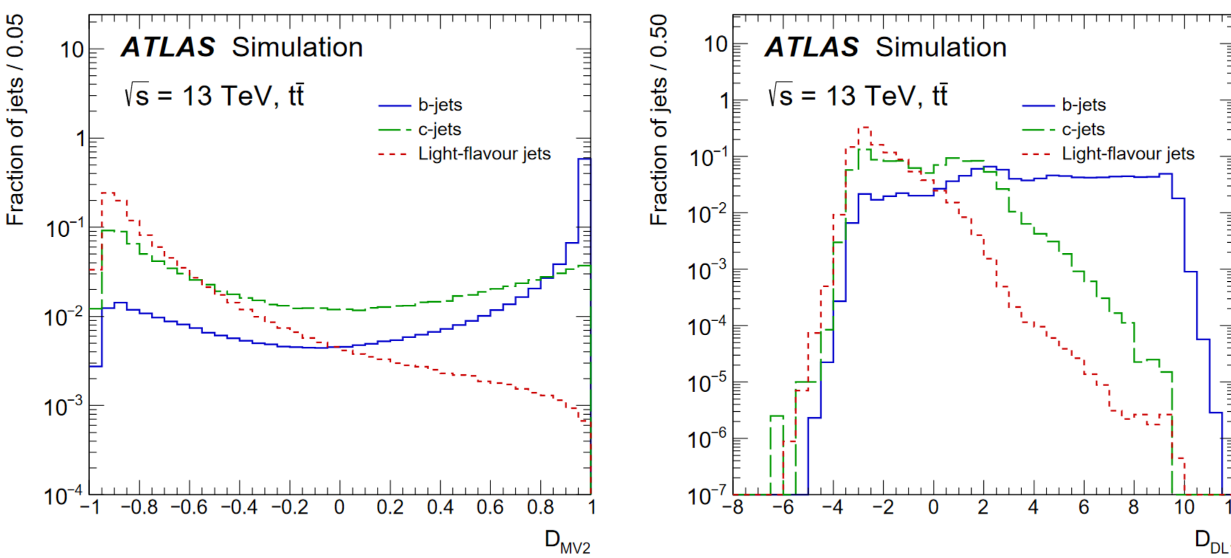


Figure 16: The fraction of light-flavour jets and c jets versus the b jets in the MV2 (left) and DL1 (right) b-tagging algorithms output distribution evaluated on the $t\bar{t}$ events [29].

calibration is performed for all supported jet collections(TODO: refer back to the object definition chapter) and working points, which are cuts in the b-tagging algorithm output identifying the different tagging efficiencies and corresponding light jet and c jetrejection rate. 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.

6.1 Calibration methods for b jet and light jet

In general, the efficiency is calculated with data and simulations, and scale factors are then calculated to recover the efficiency extracted from simulations to the data. The production of $t\bar{t}$ pairs at the LHC provides an abundant source of b jetsby 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-sign 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 emulate the performance of the algorithms to the data [29].

For the light jet mis-tagging calibration, two methods are used to measure the mis-tagging rate from the data [33]. 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 im 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 systematic uncertainties.

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) in 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. The calibration is performed on the PFlow jets and the VR-Track jets (TODO: refer back to definition in Object reconstruction section).

As determined by the CKM matrix [34, 35], the W boson decays dominantly to a pair of light quarks (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 a u quark and d quark pair or a s quark and c quark pair is 33.1%, and to pairs containing a b quark is only 0.057% [36]. Therefore, b-tagged jets from the W decay are most likely to be mis-tagged c jetsor light jets.

Furthermore, given the ratio between the DL1 light jet rejection and the corresponding charm jet rejection ranges from 10 to 40 (Figure 15), the c jet is much more likely to be mis-tagged than the light jet. This allows for a source of mis-tagged c jetsto be obtained in the $t\bar{t}$ events, requiring that one W boson decays 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 mis-tagged c jets. 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 17, 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. The following notation will be used: the jets that are the decay products of the W boson are referred to as W jets and the remaining two jets are referred to as top-jets.

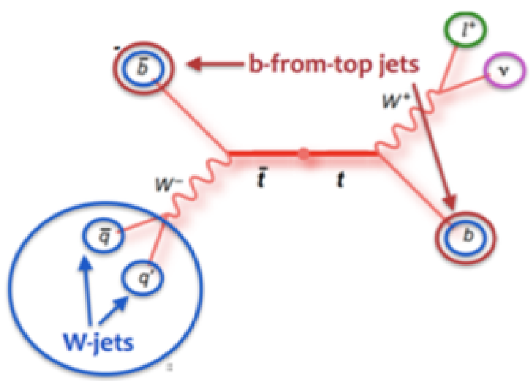


Figure 17: Feynman diagram of the semi-leptonic $t\bar{t}$ events.

A kinematic likelihood technique, referred to as KLFitter [37], is used to assign jets to the proper $t\bar{t}$ decay product (more details in Section 6.4). 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.

The charm jet efficiency is defined as the ratio of events with either of the W jet is tagged. The efficiency is evaluated in four p_{T} intervals, with boundaries of 20, 40, 65, 140 and 250 GeV for the PFlow jets and 15, 20, 40, 140 GeV for the VR-Track jets; and for four tagging intervals with the boundaries of 85%, 77%, 70% and 60%.

The choice of the bin boundaries ensures enough statistics for each bin and hence relatively flat statistical uncertainty, given the underlying charm-jet p_{T} spectrum as shown in Figure 23. The boundaries for the VR-Track jets are lower than for PFlow jets, since the track jets miss the neutral particles the reconstructed energy is significantly below the true jet energy.

The main method described in the chapter is for the "fixed-cut" calibration where the efficiency is defined as the fraction of b jetspassing the tagger. Jets are said to be tagged (untagged) at particular working point if they have DL1r scores greater (less) than the DL1r score of that working point. The events with both W jet are discarded to simplfiy the fit described in the following.

To extract the scale factors of the charm jet mis-tagging, a fit is performed by minimising the χ^2 defined as:

$$\chi^2 = \sum_{t=1}^{4}\sum_{i=1}^{4}\sum_{j=1}^{4}(N^t_{\mathrm{data}}(i,j) - p(i,j)[c^t(i)N^t_C(i,j) + N^t_J(i,j) + \sum_k c^4(k)N^t_X(i,j,k)])^2/N^t_{\mathrm{data}}(i,j)$$

$$+\sum_{i=1}^{4}\sum_{j=i}^{4}[N^{\mathrm{untag}}_{\mathrm{data}}(i,j) - p(i,j)N^{\mathrm{untag}}_{\mathrm{MC}}(i,j)]^2/N^{\mathrm{untag}}_{\mathrm{data}}(i,j). \quad (1)$$

The $c^t(i)$ is the main floating parameter in the fit, which is the charm jet mis-tagging scale factor at working point t of p_{T} bin labelled i. The other main floating parameter is $p(i,j)$ that is the normalisation factor scaling the MC to data. The $N^t_{\mathrm{data}}(i,j)$ is the number of data events with a tagged W jet in the p_{T} bin labelled i and the other (untagged) W jet in the p_{T} bin labelled j. Similarly the $N^t_C(i,j)$ is the number of MC events with a tagged W jet while the tagged W jet is indeed a c jet which can be seen as "signal". In contrast the $N^t_J(i,j)$ is the number of events with neither the tagged W jet nor the top jets are c jets and the $N^t_X(i,j,k)$ is the number of events with one of the top jets is a c jet These two types of events can be seen as "background". The later case is slightly more complicated, as the c jet lies in a different p_{T} bin to the tagged jet, denoted as k, and it only depends on the $c^4(k)$ (which is the scale factor of the 4th working point i.e. 60%) as the top jets are tagged at 60% working point. The calibration is then given as the scale factors of the four working points in bins of p_{T} defined in the above text.

6.3 Data and Monte Carlo samples

TODO: remove the overlap between this section and the Data MC chapter in the thesis Dedicated MC are used to model SM processes. The data analysed in this study correspond to 139 fb^{-1} [38, 39, 40, 41], 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 [42] and single-electron triggers [43]. 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.

All samples were produced using the ATLAS simulation infrastructure [44] and GEANT4 [45]. A subset of samples use a faster simulation based on a parameterisation of the calorimeter response and GEANT4 for the other detector systems [44]. 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 [46] using the NNPDF2.3LO set of PDFs [47] and a set of tuned parameters called the A3 tune [48].

The events that are used in this study originate mostly due to $t\bar{t}$ production. This process is modelled using the POWHEGBOX v2 [49, 50, 51, 52] generator at NLO with the `NNPDF3.0nlo` parton distribution function (PDF) set and the h_{damp} parameter[1] set to 1.5 m_{top} [53]. 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 [54]. 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.

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` [55] parton distribution function (PDF) set. The associated production of top quarks with W bosons (tW) is modelled

[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.

using the PowhegBox v2 [56, 50, 51, 52] generator at NLO in QCD using the five-flavour scheme and the NNPDF3.0nlo set of PDFs [55]. The diagram removal scheme [57] 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 [58] 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 [59] and OpenLoops [60, 61, 62] libraries. They are matched with the Sherpa parton shower [63] using the MEPS@NLO prescription [64, 65, 66, 67] using the set of tuned parameters developed by the Sherpa authors. The NNPDF3.0nnlo set of PDFs [55] is used and the samples are normalised to a next-to-next-to-leading order (NNLO) prediction [68].

Samples of diboson final states (VV) are simulated with the Sherpa v2.2.1 or v2.2.2 [58] 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 [59, 63] using the MEPS@NLO prescription [64, 65, 66, 67]. The virtual QCD correction are provided by the OpenLoops library [60, 61, 62]. 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 MadGraph5_aMC@NLO v2.3.3 [69] generator at NLO with the NNPDF3.0nlo [55] parton distribution function (PDF). The events are interfaced to Pythia8.210 [46] using the A14 tune [70] and the NNPDF2.3lo [55] PDF set. The decays of bottom and charm hadrons are simulated using the EvtGen v1.2.0 program [71].

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. [37]. 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 to avoid bias.

6.5 Maximising likelihood

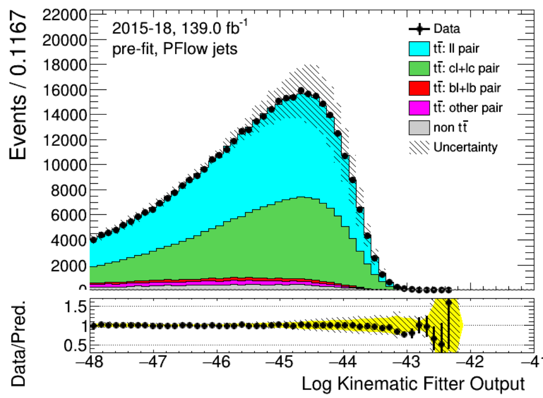


Figure 18: Distribution of the negative logarithm of the likelihood that is used to reconstruct the $t\bar{t}$ decay.

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. 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 18. 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.

	PFlow jets	Track jets
Data	227118	218351
$t\bar{t}$	235670 $\pm$ 200	223770 $\pm$ 180
Non $t\bar{t}$	7610 $\pm$ 120	7280 $\pm$ 100
Data/MC	0.934 $\pm$ 0.002	0.945 $\pm$ 0.002

Table 2: Standard selection: prefit comparison of the number of events in data and in simulation considering the PFlow jets and the VR-Track jets for events with exactly 4 jets.

6.6 Event selection

6.6.1 Standard selection

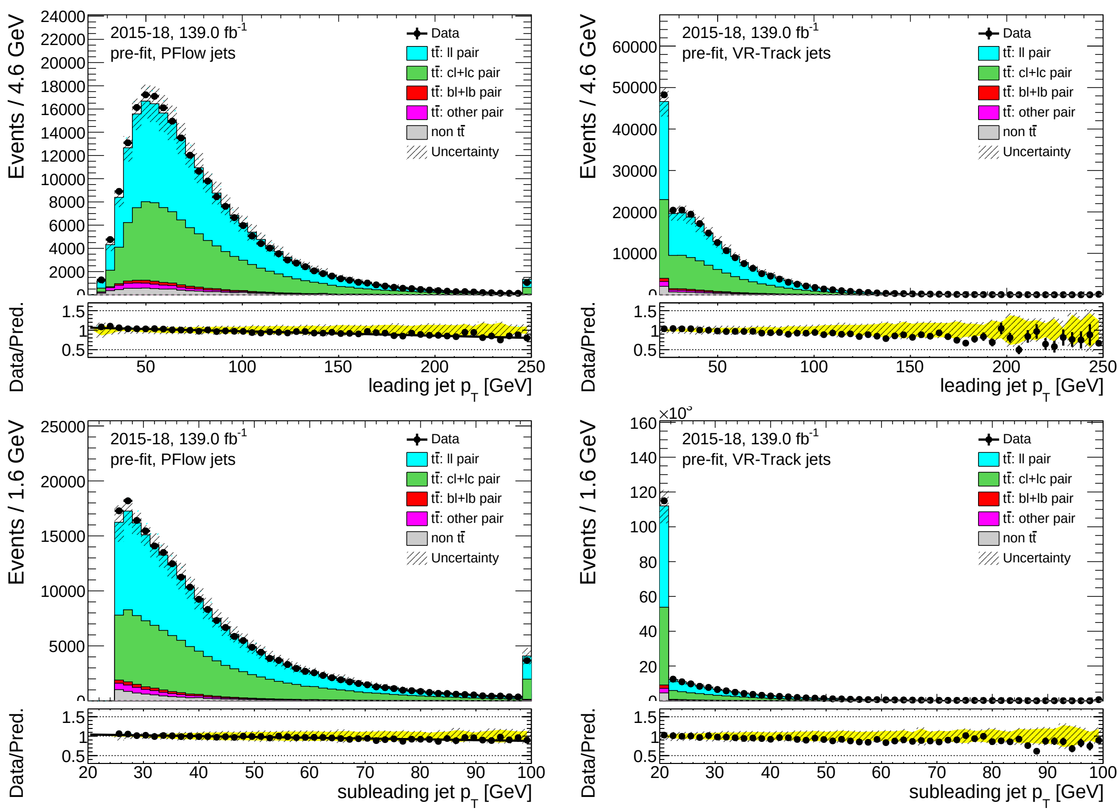


Figure 19: Standard selection: data versus simulation of the leading and sub-leading W jet p_{T} for the PFlow jets in the left column and for VR-Track jets in the right column. The leading jet and sub-leading jet refer to the highest p_{T} W jet and the second highest p_{T} jet, respectively. The 'non $t\bar{t}$' background indicates background comes from non-$t\bar{t}$ processes like W or Z production in association with jets or single-top production. The error in the table (and the following yields tables for different selection) is stats-only.

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 $E_{\mathrm{T}}^{\mathrm{miss}}$. The yields of the data and the MC are given in Table 2. An example of the p_{T} distributions before any tagging or fitting and after the standard selection is shown in Figure 19. More plots can be found in Appendix A.1.1. The yellow band in the lower pad shows the overall systematic 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 Low-p_{T} selection

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 PFlow jets with $20 < p_{\mathrm{T}} < 25$ GeV. The p_{T} threshold of the VR-Track jets is 10 GeV therefore the low-p_{T} selection is not needed. 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 requirements for the selection are the same. This additional cut provides candidates for the PFlow W jet that is used for calibration in the $20 - 25$ GeV region. The inclusive yields of the low-p_{T} selection of the data and the MC are given in Table 3, and the p_{T} distributions of the W jets are shown in Figure 20. More plots of the kinematic distributions are shown in Appendix A.1.2. 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.

	PFlow jets
Data	59987
$t\bar{t}$	56530 ± 90
Non $t\bar{t}$	3340 ± 60
Data/MC	1.002 ± 0.004

Table 3: Low-p_{T} selection: prefit comparison of the number of events in data and MC for the PFlow W jets. Events are required to have exactly 3 jets with $p_{\mathrm{T}} > 25$ GeV and one jet with $20 < p_{\mathrm{T}} < 25$ GeV.

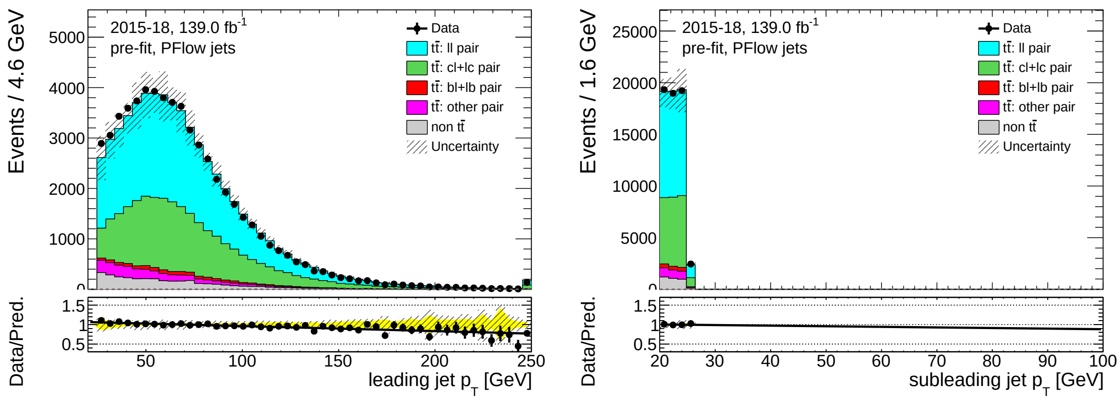


Figure 20: Low-$p_{\rm T}$ selection: data versus simulation of the PFlow W jets $p_{\rm T}$.

6.6.3 High-p_T selection

It has been observed that in the previous calibrations that the statistics are relatively low for the high-$p_{\rm T}$ region (e.g. jet $p_{\rm T} > 100$ 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 $p_{\rm T} > 70$ GeV. Other than that, all requirements for the selection remain the same.

The choice of cut value at 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:

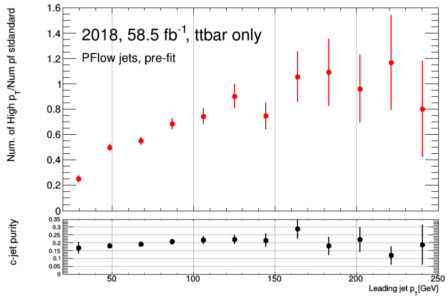
$$c \text{ jet purity} = \frac{\mathrm{N}_{\text{true } c \text{ jet}}}{\mathrm{N}_{\text{all}}}, \tag{2}$$

where $N_{\text{true } c \text{ jet}}$ 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-$p_{\rm T}$ selection will maximally increase the statistics while mininally decreasing the c jet purity, therefore a figure of merit $P^{\rm Cut}$ is defined as:

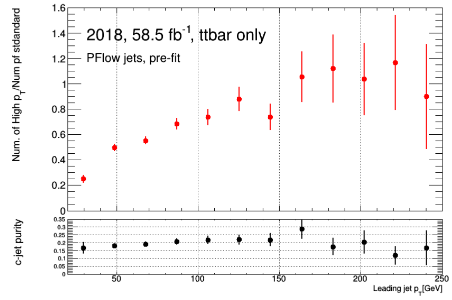
$$P^{\rm Cut} = \frac{\sum_i \text{Gain in stats}_{\rm i}^2}{\sum_i c \text{ jet purity}_{\rm i}^2},$$

where i stands for the number of bins. The “Gain in stats” stands for increase in statistics and it’s summed over all bins in Figure 21. The c jet purity and the statistical gain are calculated for 4 different cut values as shown in Figure 21, comparing with the cut value of 0. The value of 70 GeV is chosen as it gives the highest value of $P^{\rm Cut}$.

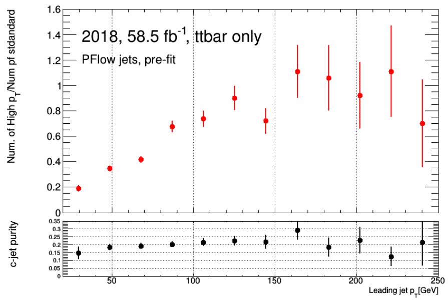
The yields of the data and the MC are given in Table 4. An example of the $p_{\rm T}$ distributions before any tagging or fitting, applying the high-$p_{\rm T}$ selection is shown in Figure 22. In general the event statistics improve about 80% in region with $p_{\rm T} > 70$ GeV as desired. More plots



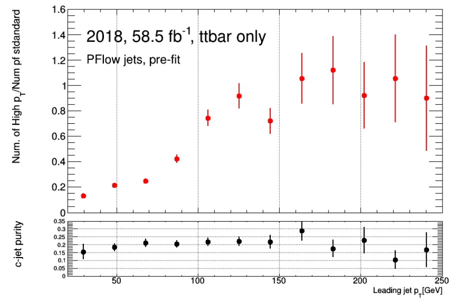
(a) No cut



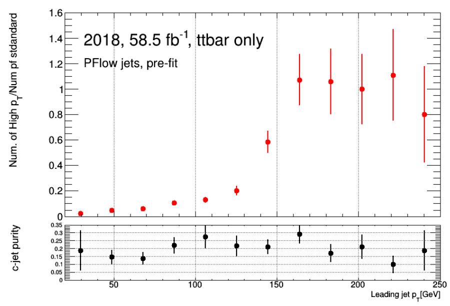
(b) Cut value: 40 GeV



(c) Cut value: 70 GeV



(d) Cut value: 90 GeV



(e) Cut value: 140 GeV

Figure 21: Comparison of different cut values in terms of gain in stats and c jet purity.

can be found in Appendix A.2.

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 and the MC are given in Table 5, an example of the p_{T} distributions before any tagging or fitting and after the combined selection is shown in Figure 23. More plots can be found in Appendix A.3.

	PFlow jets	Track jets
Data	98273	83957
$t\bar{t}$	99430 ± 120	87476 ± 110
Non $t\bar{t}$	1842 ± 21	1570 ± 20
Data/MC	0.97 ± 0.003	0.94 ± 0.003

Table 4: High-p_{T} selection: prefit comparison of the number of events in data and in simulation considering the PFlow W jets and the VR-Track jets.

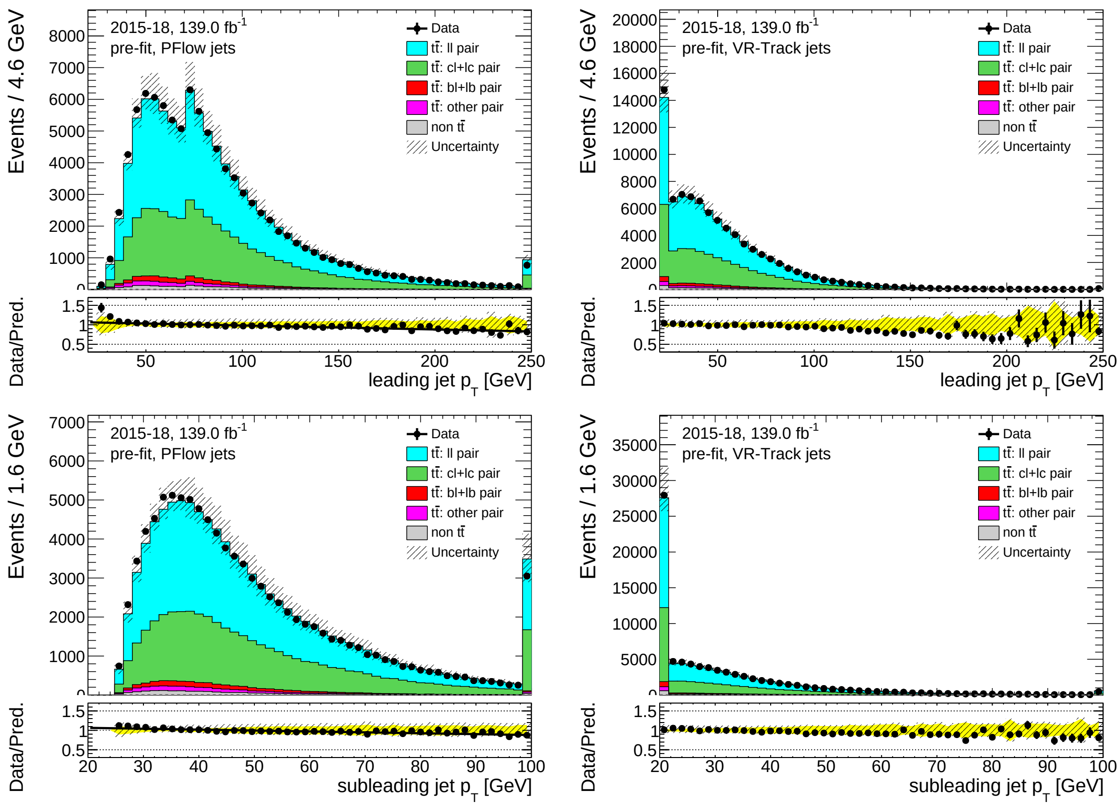


Figure 22: High-p_{T} selection: data versus simulation of W jets p_{T} for PFlow jets in the left column and for VR-Track jets in the right column.

6.7 Systematic uncertainties

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

6.7.1 Experimental uncertainties

TODO: refer to the analysis Part Experimental uncertainties are related to the detector and estimated using data-driven methods or MC simulations. The lepton energy scale and resolution are corrected to provide better agreement between MC predictions and data,

	PFlow 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 5: Combined selection: prefit comparison of the number of events in data and in simulation considering the PFlow jets and the VR-Track jets for an inclusive selection.

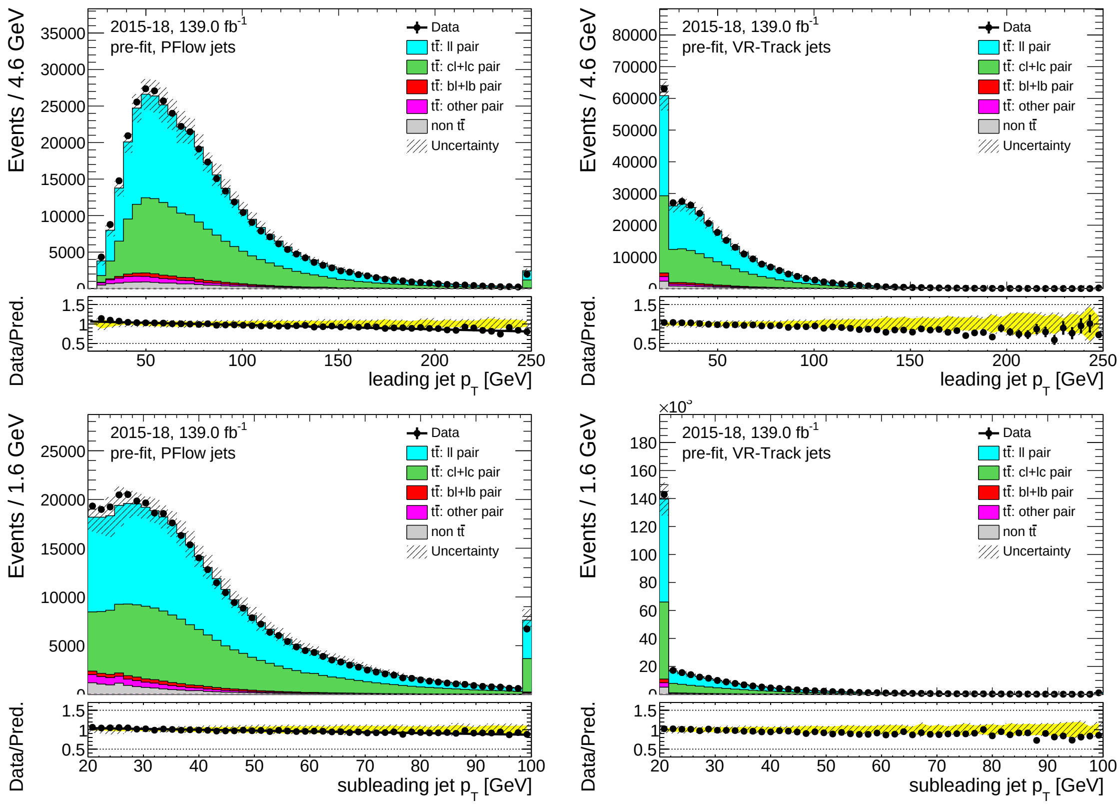


Figure 23: Combined selection: data versus simulation of W jetsp_{T} for PFlow jets in the left column and for VR-Track jets in the right column.

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.

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) are 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 (mis-tagging) probabilities for b (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.9.

6.7.2 Modelling uncertainties

The uncertaity due to different choices of the parton shower models is estimated by comparing the MC samples generated with nominal parton shower model and with the alternative parton shower model. More specifically, it is derived by comparing the prediction from POWHEG interfaced either to PYTHIA8 or Herwig++. The uncertainty due to additional radiation in the initial state and the final state is estimated by comparing the nominal MC samples with the MC samples with alternative scale of renormalisation and factorisation. 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 renormalisation and factorisation scales decreased by a factor of two and the *hdamp* parameter doubled (which controls the p_{T} of the first additional emission), while the sample with a decrease in radiation has the scales increased by a factor of two. In all cases, MC-to-MC SFs are taken into account. In addition, the uncertainty due to the variations samples being produced by fast simulation while the nominal samples being produced full simulation is also considered. The comparisons of the nominal $t\bar{t}$ sample and the samples with each systematic uncertainty are shown in Table 6.

	PFlow 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 6: Comparison of the number of events in data and in simulation considering the PFlow jets and the VR-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 splits into a pair a b quarks and one of them is assigned as a W jet. The first source can be excluded by requiring no c jetsin the W jets, meaning 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 7 and Figure 24, where an extra cut requiring at least one W jet with DL1r > 8 is added to the combined selection to reject most of the true c jetsand true light jet. A more thorough study is done in Ref. [72], where the mis-modelling factor is measured to be 1.25 ± 0.25, which is also consistent with the 30% mismodelling observed in the previous study. Therefore, 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 jets events), are scaled by 1.25 ± 0.25. 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.

	PFlow jets	VR-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 7: Yields of the 2018 data and 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.

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, the calibration includes the PFlow jet and the 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.

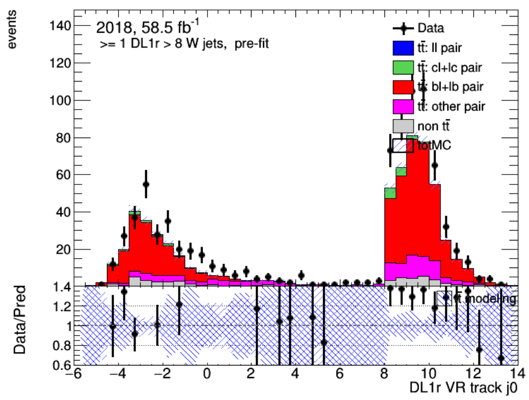
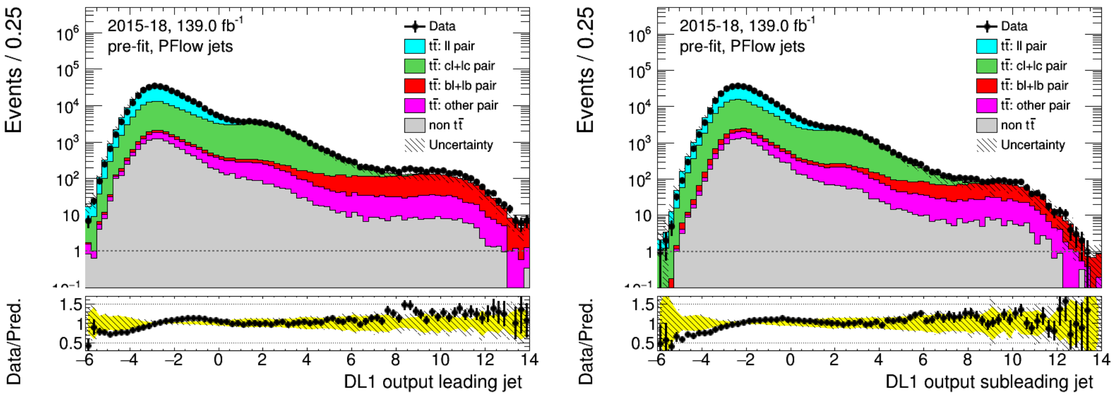


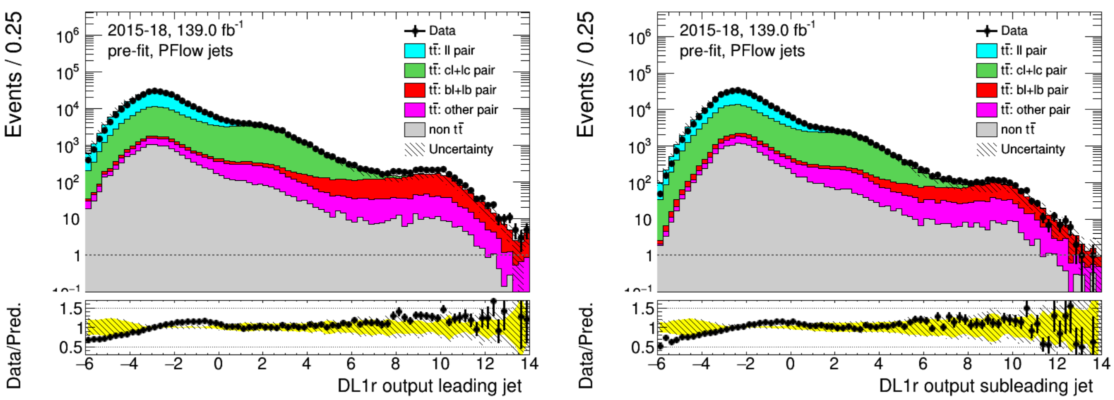
Figure 24: The DL1r score distribution of the leading VR-Track jet, requiring at least 1 VR-Track jets have DL1r $>$ 8 to reject most of the light and the c jets, with $t\bar{t}$ modelling and statistical uncertainties.

6.9.2 b-tagging algorithms output distribution

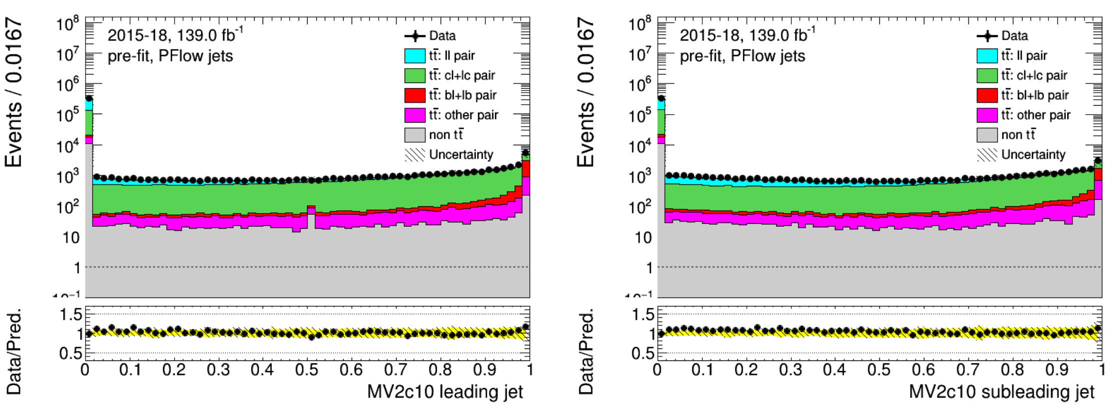
The distributions of the b-tagging algorithm' output of the MC and the data of the latest calibration (December 2020) are shown in Figure 25 for the PFlow jets and Figure 26 for the VR-Track jets, combining the standard selection, low p_{T} and the high-p_T selection. In these figure, the data events 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 systematic uncertainties and the black band represents the $t\bar{t}$ modelling systematic uncertainty, which dominates at low b-tagging discriminant (DL1 or DL1r $<$ 4). The experimental systematic uncertainty is in general very small. At high b-tagging discriminant (DL1 or DL1r $>$ 4), the uncertainty due to the 1.25 ± 0.25 scale factor becomes more important.



(a) DL1 tagger output

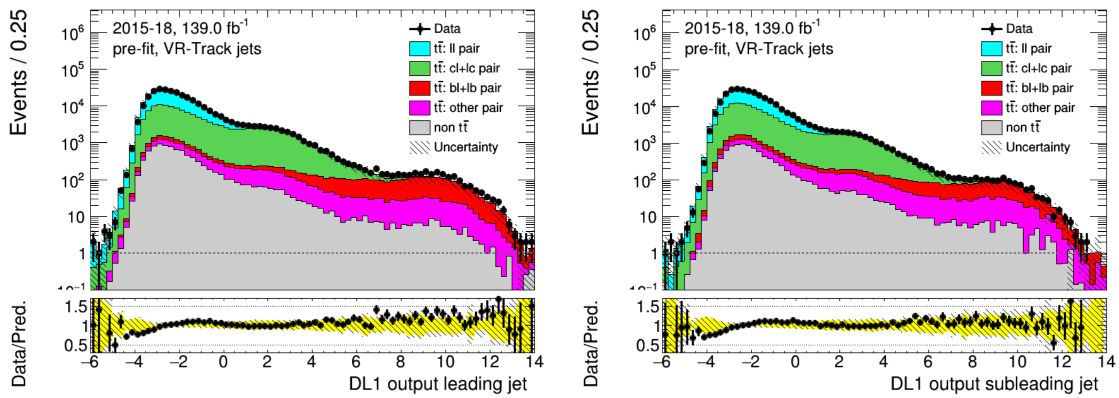


(b) DL1r tagger output

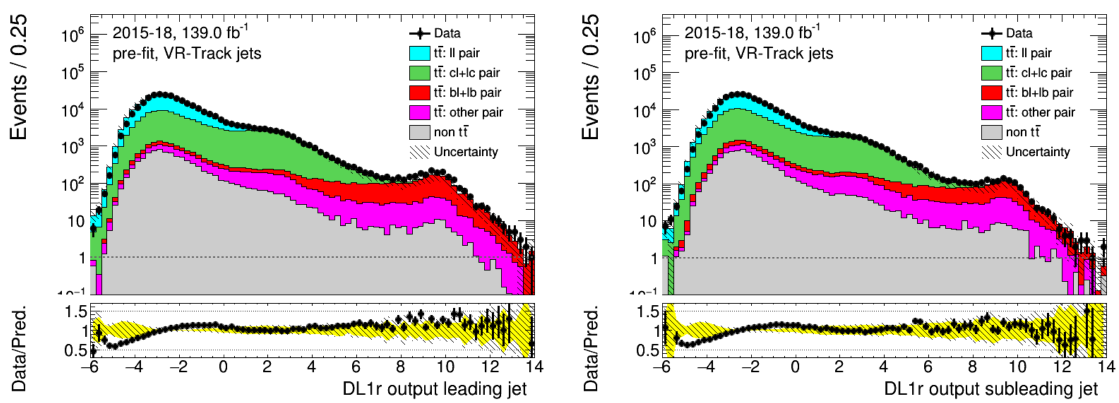


(c) MV2c10 tagger output

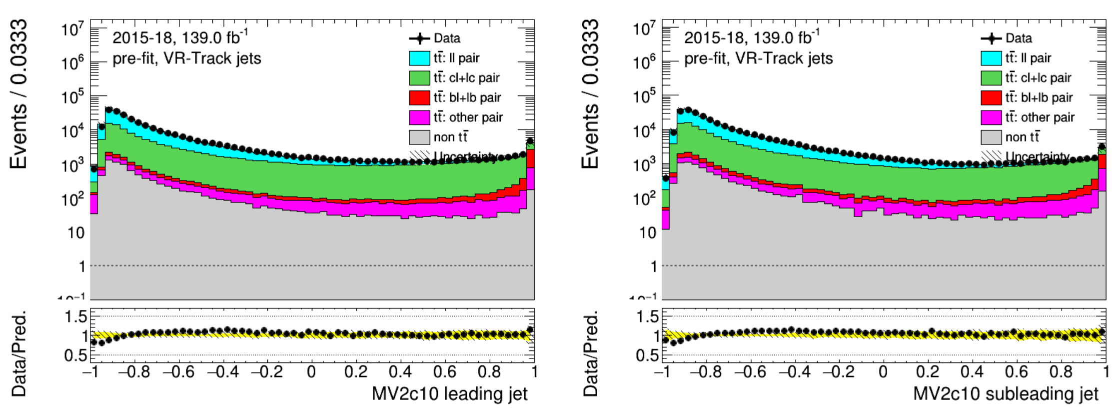
Figure 25: PFlow jets: distributions of the DL1, DL1r and MV2c10 tagger outputs of the combined selection, leading jet in the left column and sub-leading jet in the right column, before fitting or tagging with full uncertainties.



(a) DL1 tagger output



(b) DL1r tagger output



(c) MV2c10 tagger output

Figure 26: VR-Track jets: distributions of the DL1, DL1r and MV2c10 tagger outputs of the combined selection, leading jet in the left column and sub-leading jet in the right column, before fitting or tagging with full uncertainties.

6.9.3 Efficiencies and Scale Factors

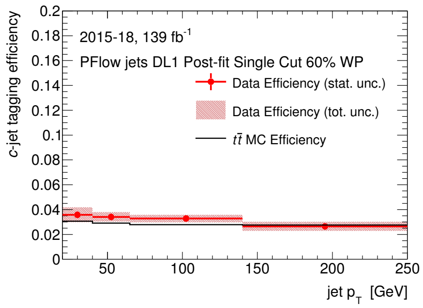
The Dl1 and DL1r c jetefficiencies and scale factors with systematics uncertainties are calculated with four fixed cut working points for the PFlow and VR VR-Track jets collection in the latest derivation in December 2020.

The c jet mis-tagging efficiencies are shown in Figure 27-30 for the PFlow jet collections and the VR-Track jets with the DL1 and the DL1r tagger. For PFlow jets, these results combine the standard selection, lowp_{T} selection and the high-p_T selection and for the VR-Track jets they combine the standard selection and the high-p_T selection.

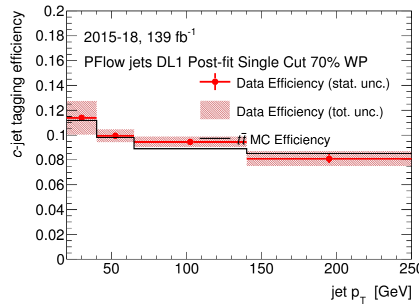
The 1.25$\pm$0.25 scale factor is applied on events with 3 true b jets. The overall uncertainties are shown in the red band. The scale factors are shown in Figure 31-34 for the PFlow jets and the VR-Track jets with the DL1 and DL1r tagger. The tighter working points (60%, 70%) show larger uncertainties and bigger deviation from 1, while the looser working points (77%, 85%) have much smaller uncertainty and the simulation is able to recover the data well due to more abundant events statistics. For the PFlow jets, in most of the working points the systematic uncertainties dominate in the low-p_{T} bins ($p_{\mathrm{T}} < 150$) and the statistical error, represented by the error bars on the markers, become more important in the last bin. For the VR-Track jets the statistical uncertainty is relatively constant for all bins while the systematic uncertainty increases as the p_{T} increases. To demonstrate the effect on statistics with the high-p_{T} selection, the fractional statistical uncertainties of 60% working point scale factor are shown in Table 8 for the standard and the combined selection. In some bins the statistical uncertainty can decrease up to 30%, suggesting that the high-p_{T} selection is successful at increasing events statistics.

	PFlow jets			VR-Track jets		
	Standard selection	High-p_{T} selection	Fractional decrease	Standard selection	High-p_{T} selection	Fractional decrease
Bin No.1	3.3%	3.3%	0.0%	5.6%	5.3%	5.7%
Bin No.2	3.1%	2.8%	10.7%	4.2%	3.7%	13.5%
Bin No.3	3.4%	2.6%	30.8%	5.8%	4.9%	18.4%
Bin No.4	12.1%	9.3%	30.1%	7.2%	5.6%	28.6%

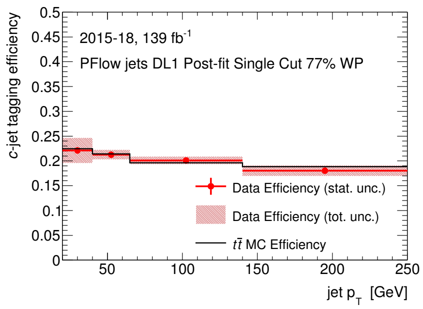
Table 8: Comparison of the fractional statistical uncertainty in the DL1r 60% working point scale factor. The p_{T} range of each bin can be found in section 6.2.



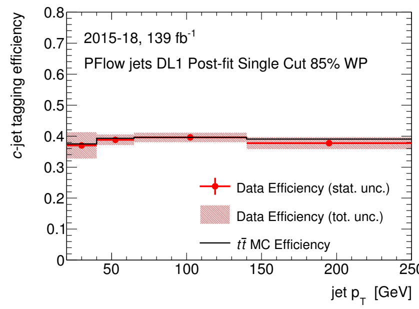
(a) 60% working point



(b) 70% working point

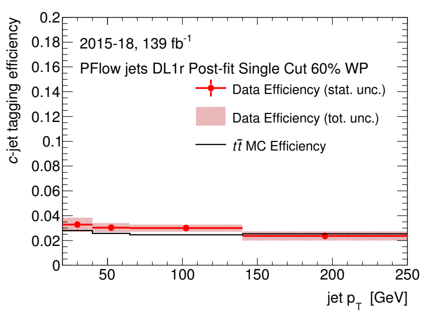


(c) 77% working point

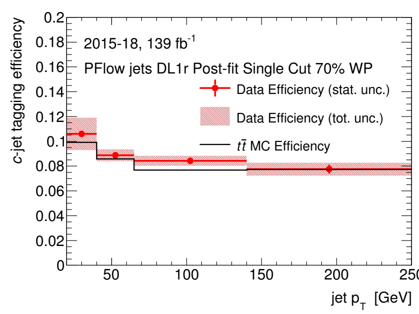


(d) 85% working point

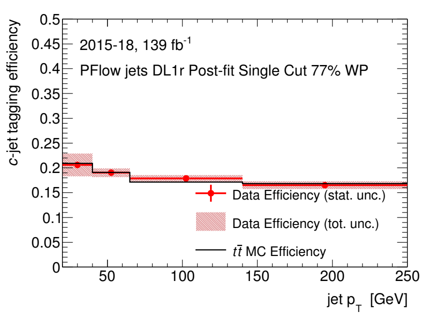
Figure 27: Charm-jet efficiencies for the PFlow jets collection with the DL1 tagger.



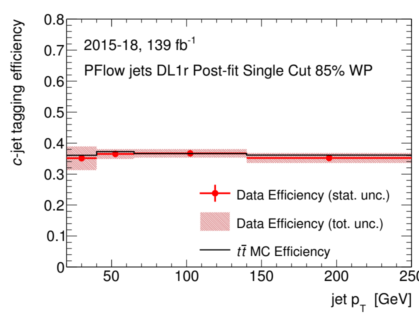
(a) 60% working point



(b) 70% working point

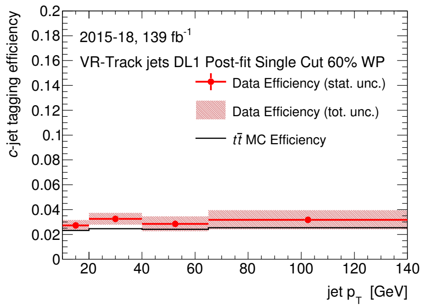


(c) 77% working point

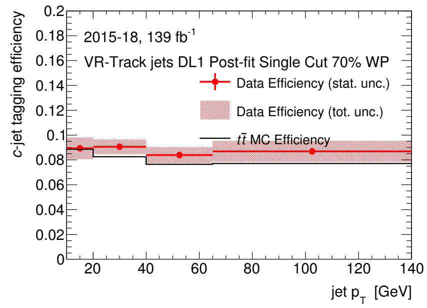


(d) 85% working point

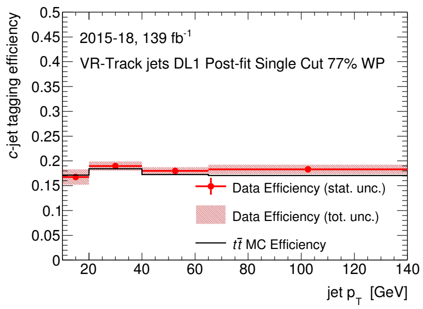
Figure 28: Charm-jet efficiencies for the PFlow jets collection with the DL1r tagger.



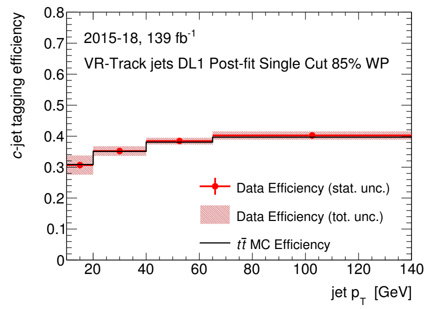
(a) 60% working point



(b) 70% working point

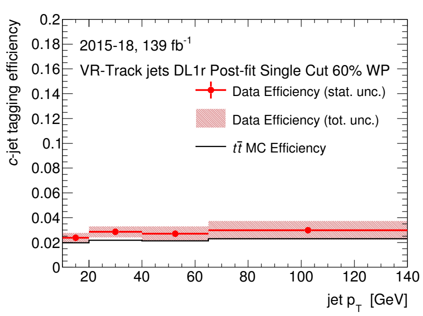


(c) 77% working point

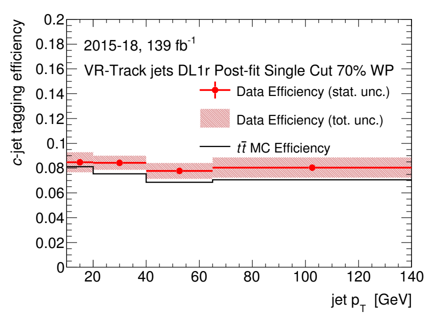


(d) 85% working point

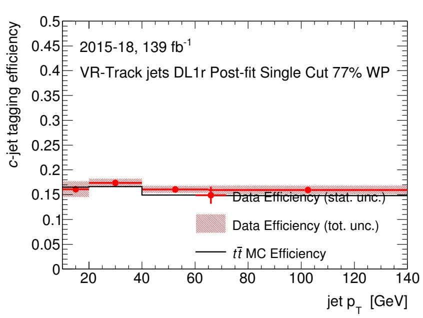
Figure 29: Charm-jet efficiencies for the VR-Track jets collection with the DL1 tagger.



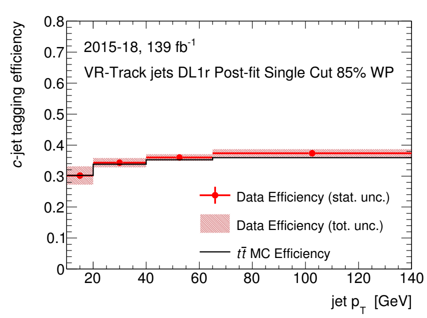
(a) 60% working point



(b) 70% working point

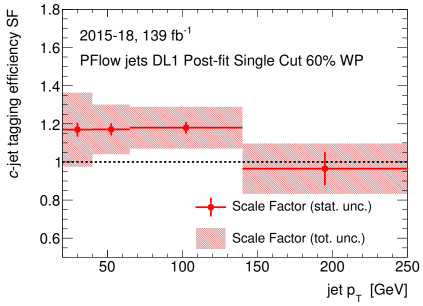


(c) 77% working point

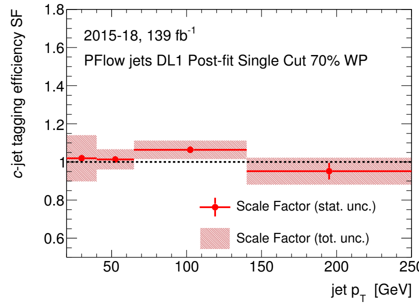


(d) 85% working point

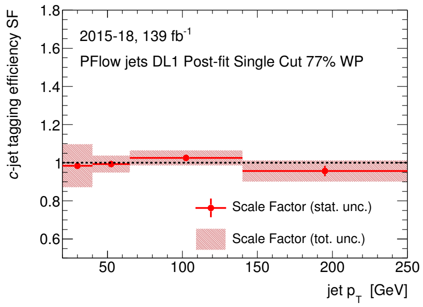
Figure 30: Charm-jet efficiencies for the VR-Track jets collection with the DL1r tagger.



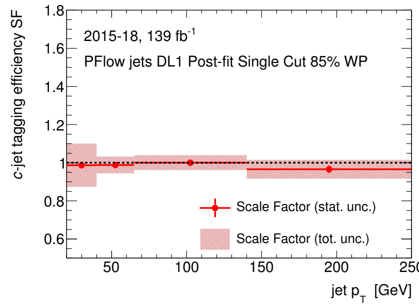
(a) 60% working point



(b) 70% working point

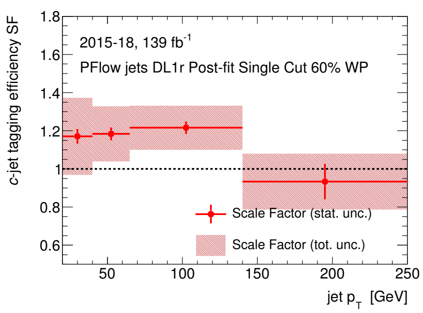


(c) 77% working point

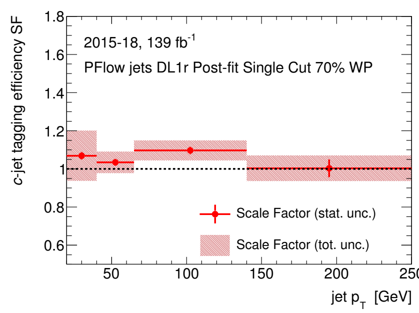


(d) 85% working point

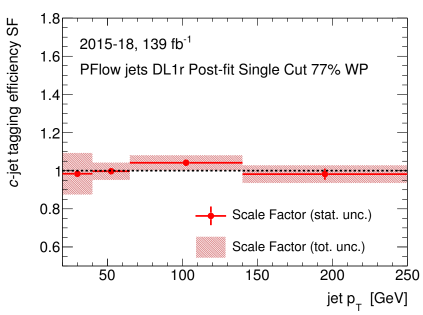
Figure 31: Charm-jet scale factors for the PFlow jets collection with the DL1 tagger.



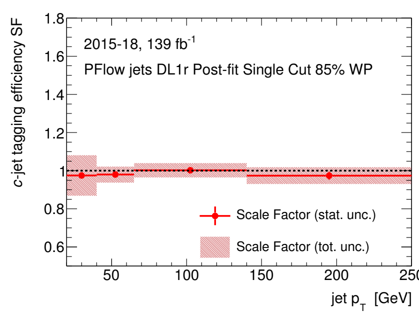
(a) 60% working point



(b) 70% working point

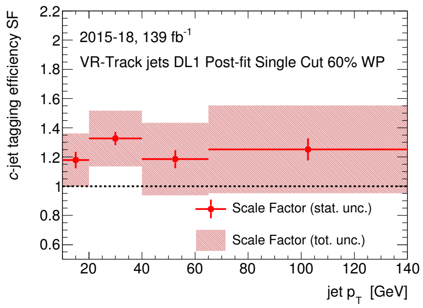


(c) 77% working point

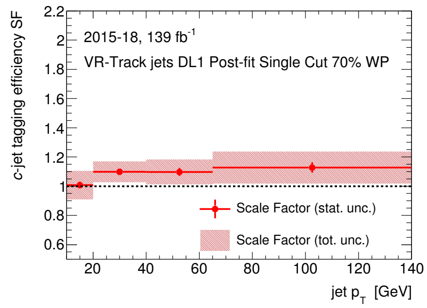


(d) 85% working point

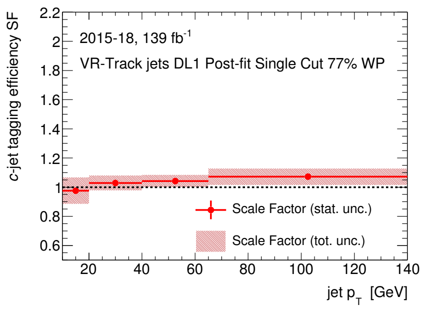
Figure 32: Charm-jet scale factors for the PFlow jets collection with the DL1r tagger.



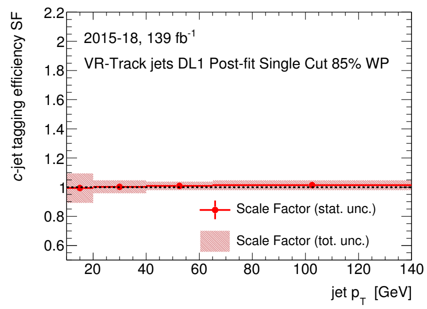
(a) 60% working point



(b) 70% working point

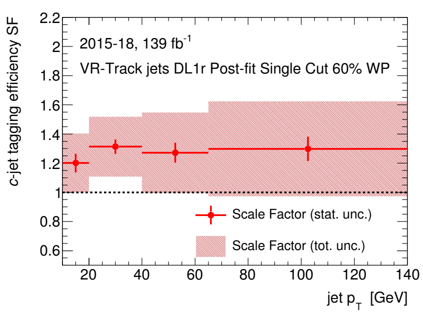


(c) 77% working point

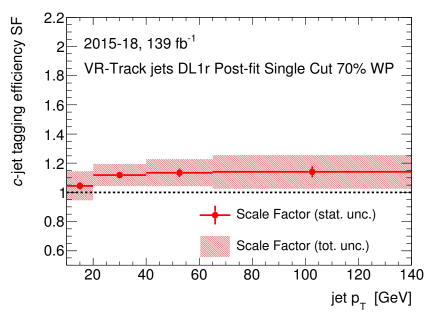


(d) 85% working point

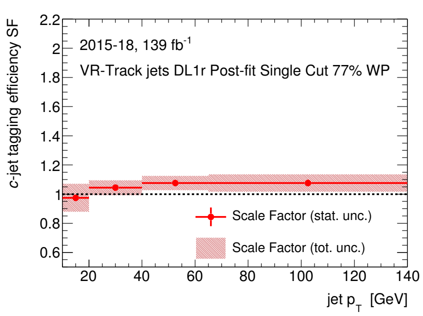
Figure 33: Charm-jet scale factors for the VR-Track jets collection with the DL1 tagger.



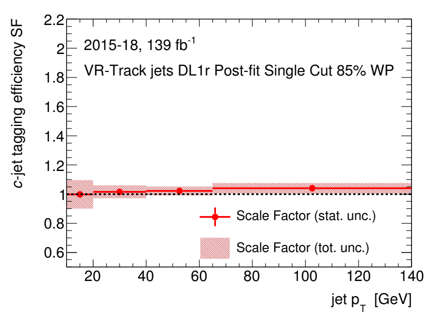
(a) 60% working point



(b) 70% working point



(c) 77% working point



(d) 85% working point

Figure 34: Charm-jet scale factors for the VR-Track jets collection with the DL1r tagger.

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] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[2] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[3] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[4] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[5] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[6] L Rossi. "The LHC superconducting magnets". In: 1 (2003), pp. 141–145.

[7] J. P. Blewett. "200-GeV Intersecting Storage Accelerators". In: *eConf* C710920 (1971), p. 501.

[8] Ph Lebrun et al. "Report of the Task Force on the Incident of 19 September 2008 at the LHC". In: *CERN LHC Project Rep* 1168 (2009).

[9] R Alemany-Fernandez et al. "Operation and Configuration of the LHC in Run 1". In: (2013).

[10] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[11] ATLAS Collaboration. "A Particle Consistent with the Higgs Boson Observed with the ATLAS Detector at the Large Hadron Collider". In: *Science* 338 (2012), p. 1576. DOI: 10.1126/science.1232005.

[12] F Bordry et al. "The first long shutdown (LS1) for the LHC". In: (2013).

[13] Jorg Wenninger. "Operation and Configuration of the LHC in Run 2". In: (Mar. 2019). URL: https://cds.cern.ch/record/2668326.

[14] Georges Aad et al. "ATLAS data quality operations and performance for 2015–2018 data-taking". In: *JINST* 15.04 (2020), P04003.

[15] ATLAS Collaboration. *ATLAS Magnet System: Magnet Project Technical Design Report, Volume 1*. ATLAS-TDR-6; CERN-LHCC-97-018. 1997. URL: https://cds.cern.ch/record/338080.

[16] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 1*. ATLAS-TDR-4; CERN-LHCC-97-016. 1997. URL: https://cds.cern.ch/record/331063.

[17] ATLAS Collaboration. *ATLAS Pixel Detector: Technical Design Report*. ATLAS-TDR-11; CERN-LHCC-98-013. 1998. URL: https://cds.cern.ch/record/381263.

[18] ATLAS Collaboration. *ATLAS Insertable B-Layer: Technical Design Report*. ATLAS-TDR-19; CERN-LHCC-2010-013. 2010. URL: https://cds.cern.ch/record/1291633.

[19] ATLAS Collaboration. *ATLAS Calorimeter Performance: Technical Design Report*. ATLAS-TDR-1; CERN-LHCC-96-040. 1996. URL: https://cds.cern.ch/record/331059.

[20] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter: Technical Design Report*. ATLAS-TDR-2; CERN-LHCC-96-041. 1996. URL: https://cds.cern.ch/record/331061.

[21] ATLAS Collaboration. *ATLAS Tile Calorimeter: Technical Design Report*. ATLAS-TDR-3; CERN-LHCC-96-042. 1996. URL: https://cds.cern.ch/record/331062.

[22] ATLAS Collaboration. *ATLAS Level-1 Trigger: Technical Design Report*. ATLAS-TDR-12; CERN-LHCC-98-014. 1998. URL: https://cds.cern.ch/record/381429.

[23] ATLAS Collaboration. *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*. ATLAS-TDR-16; CERN-LHCC-2003-022. 2003. URL: https://cds.cern.ch/record/616089.

[24] Robert Blair et al. "The ATLAS High Level Trigger Region of Interest Builder". In: *JINST* 3 (2008), P04001. DOI: 10.1088/1748-0221/3/04/P04001. arXiv: 0711.3217 [physics.ins-det].

[25] 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].

[26] 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.

[27] 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.

[28] "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.

[29] 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].

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

[31] 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.

[32] "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run". In: ATL-PHYS-PUB-2016-012 (June 2016). URL: https://cds.cern.ch/record/2160731.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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].

[39] 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].

[40] 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].

[41] 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.

[42] 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].

[43] 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].

[44] 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].

[45] 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.

[46] 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].

[47] 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].

[48] 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.

[49] 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].

[50] 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.

[51] 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].

[52] 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].

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

[54] 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.

[55] 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].

[56] 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].

[57] 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].

[58] 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].

[59] 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].

[60] 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].

[61] 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].

[62] 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].

[63] 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].

[64] 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].

[65] 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].

[66] 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.

[67] 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].

[68] 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.

[69] 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].

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

[71] 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.

[72] 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 $\sqrt{s} = 13\,\text{TeV}$ 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 c jet calibration

A.1 Additional plots for kinematic variables

A.1.1 Standard selection

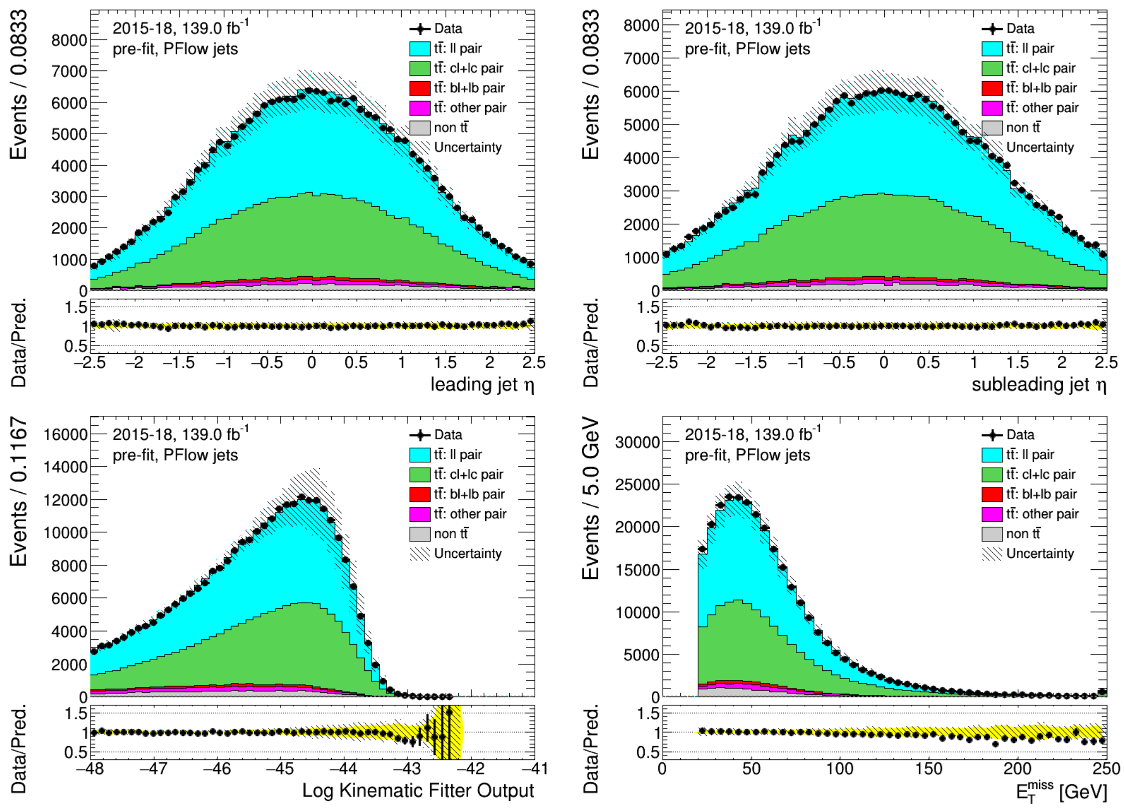


Figure 35: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the standard selection, before fitting or tagging with full uncertainties.

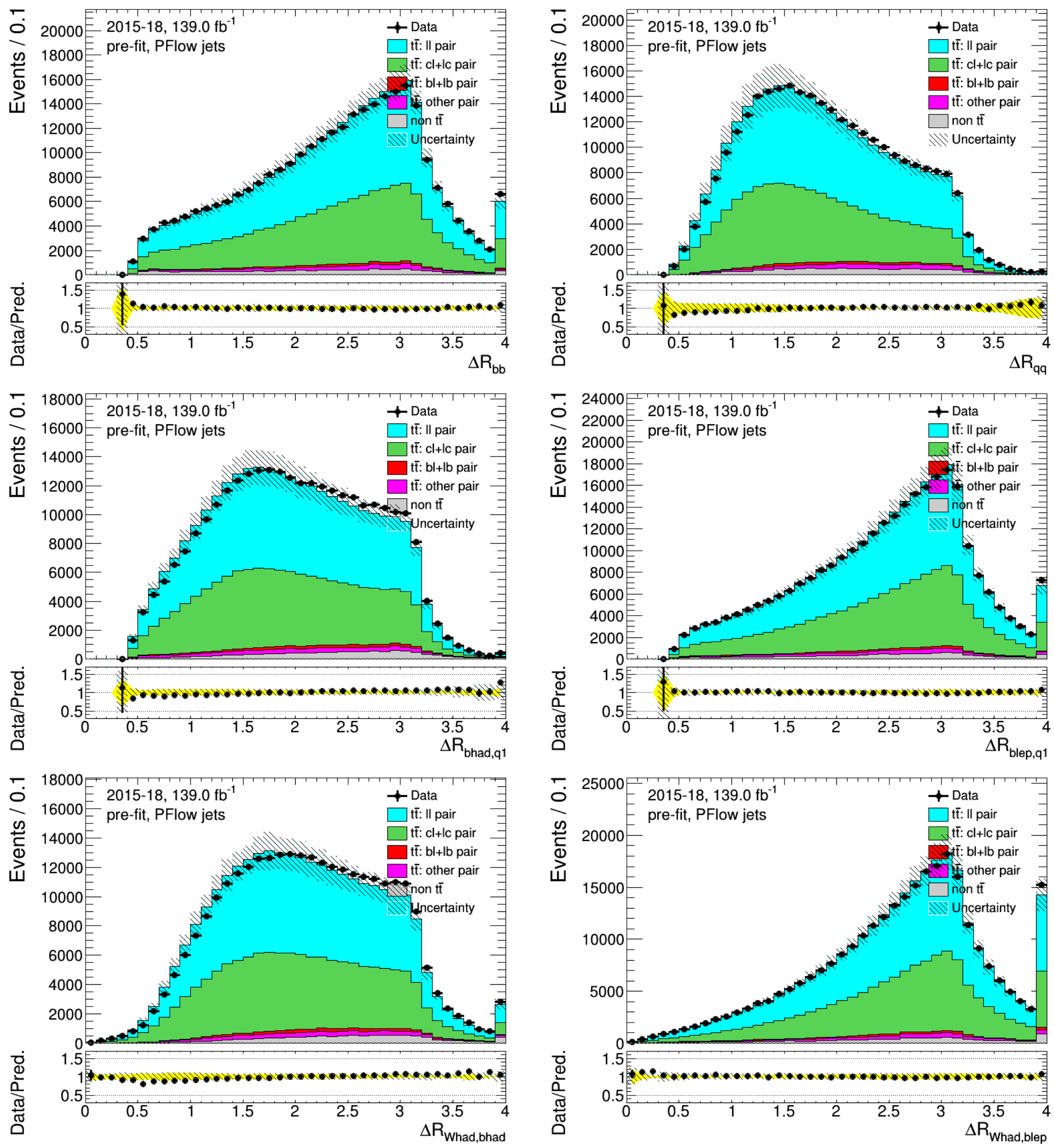


Figure 36: PFlow jets: distributions of angle related variables of the combination of the standard selection, before fitting or tagging with full uncertainties.

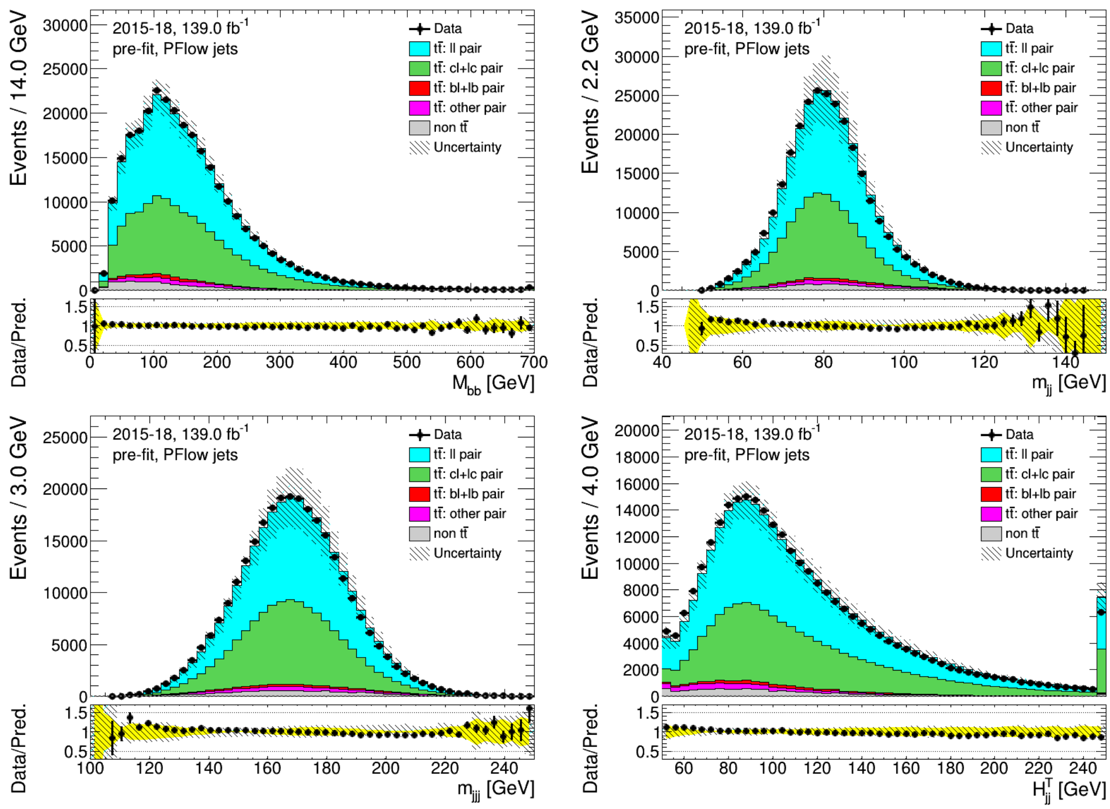


Figure 37: PFlow jets: distributions of mass related variables of the standard selection, before fitting or tagging with stat-only uncertainties.

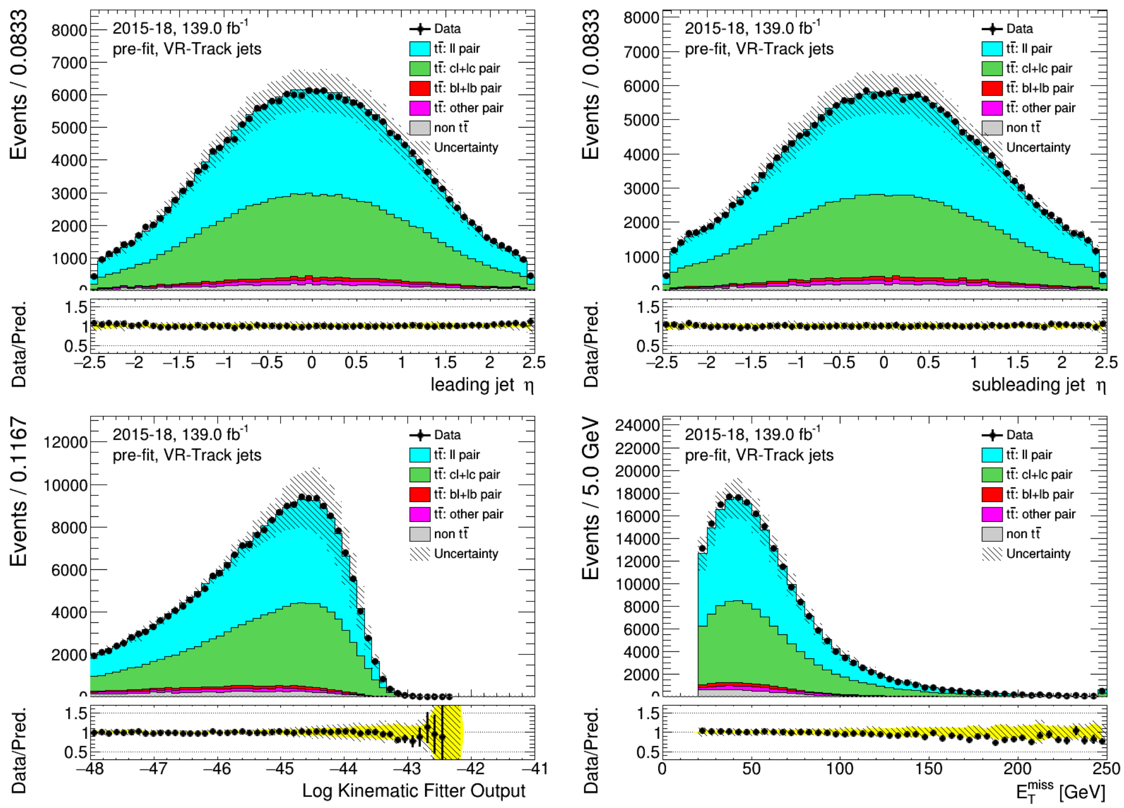


Figure 38: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the standard selection, before fitting or tagging with full uncertainties.

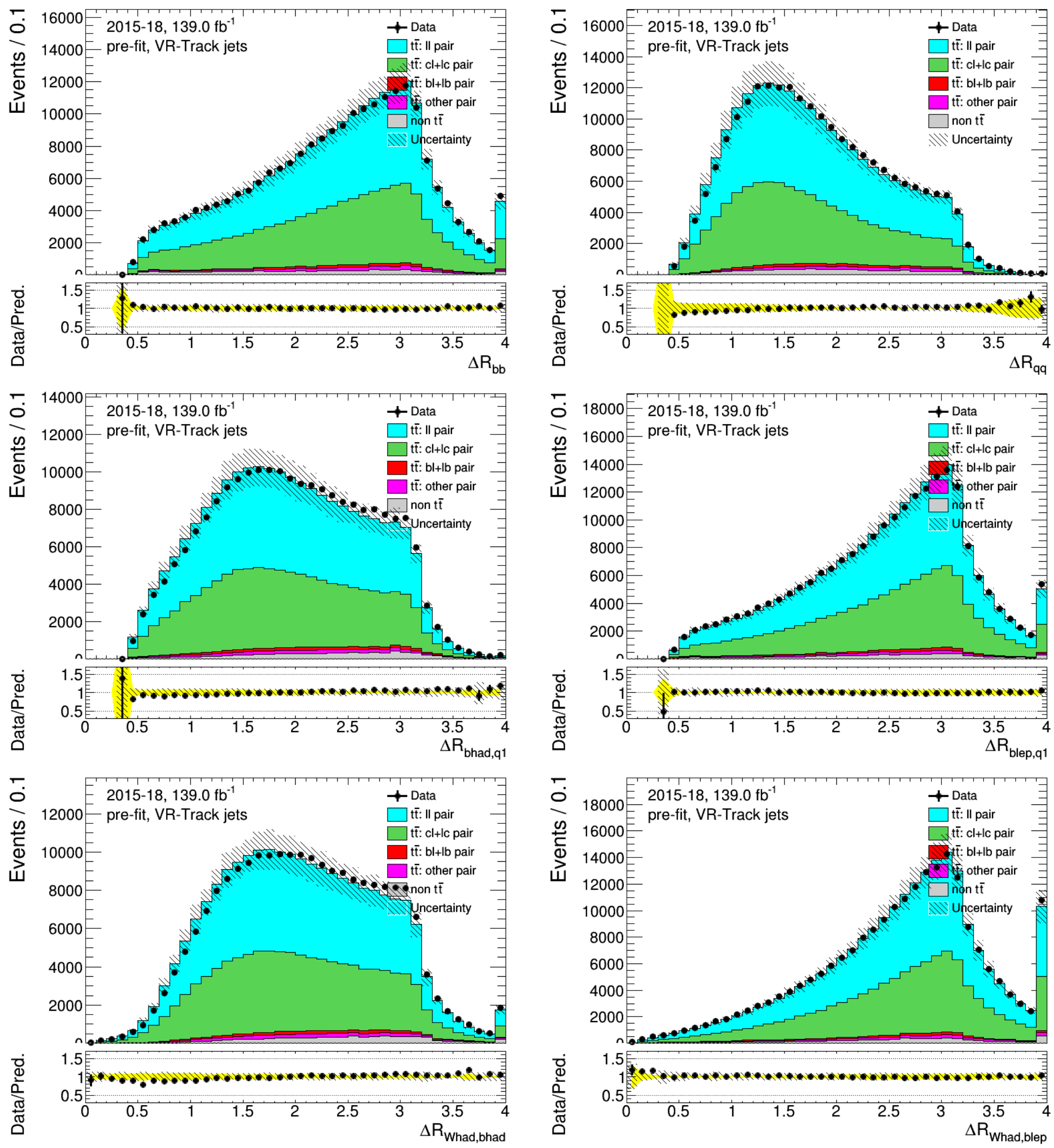


Figure 39: VR-Track jets: distributions of angle related variables of the combination of the standard selection, before fitting or tagging with full uncertainties.

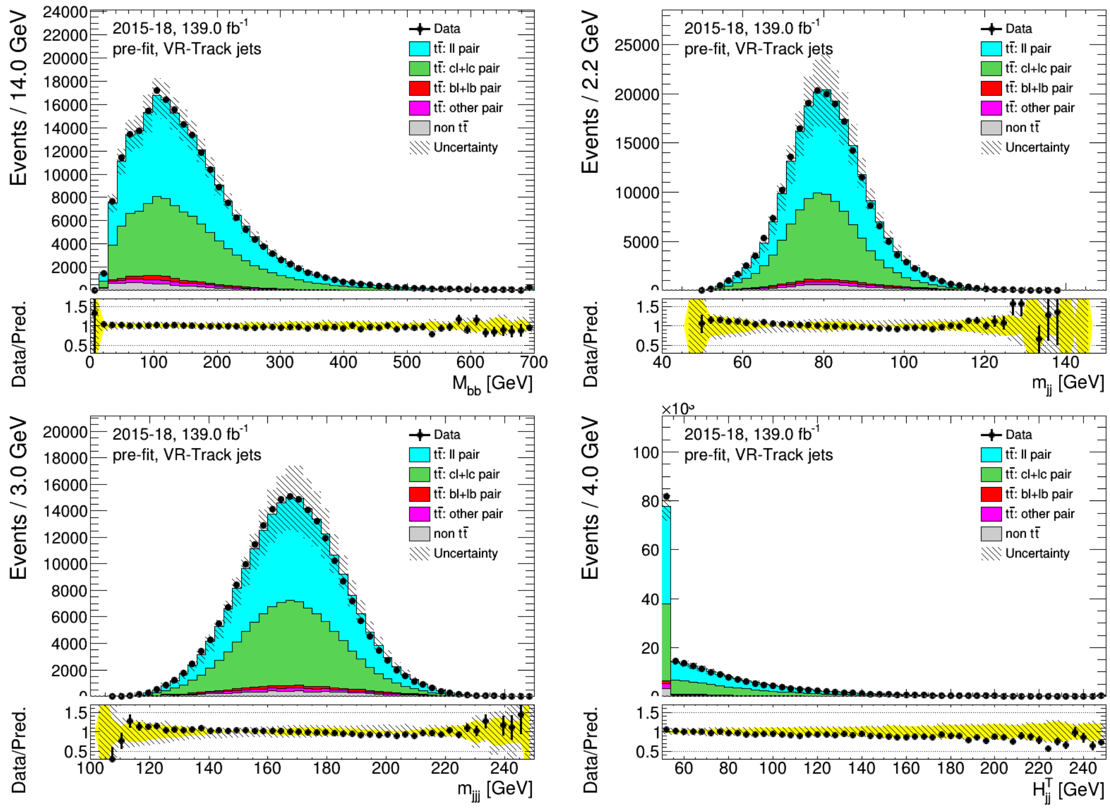


Figure 40: VR-Track jets: distributions of mass related variables of the standard selection, before fitting or tagging with stat-only uncertainties.

A.1.2 Low-p_T selection

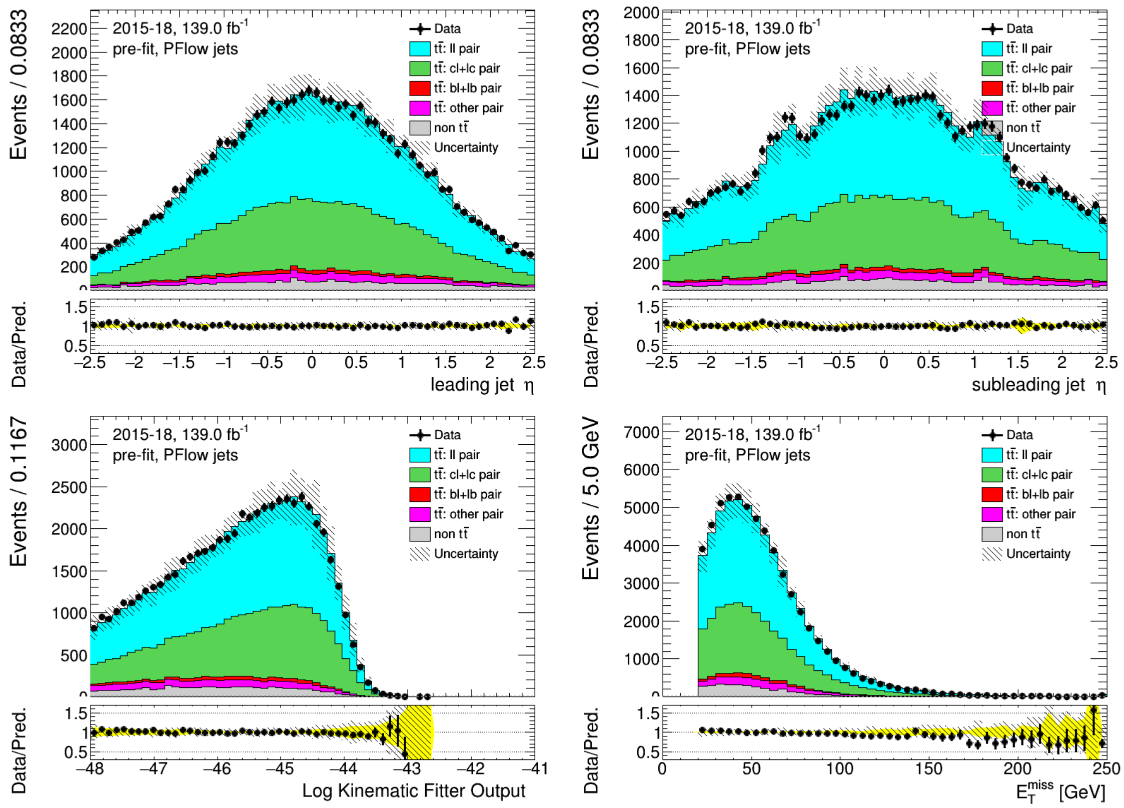


Figure 41: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the low-p_T selection, before fitting or tagging with full uncertainties.

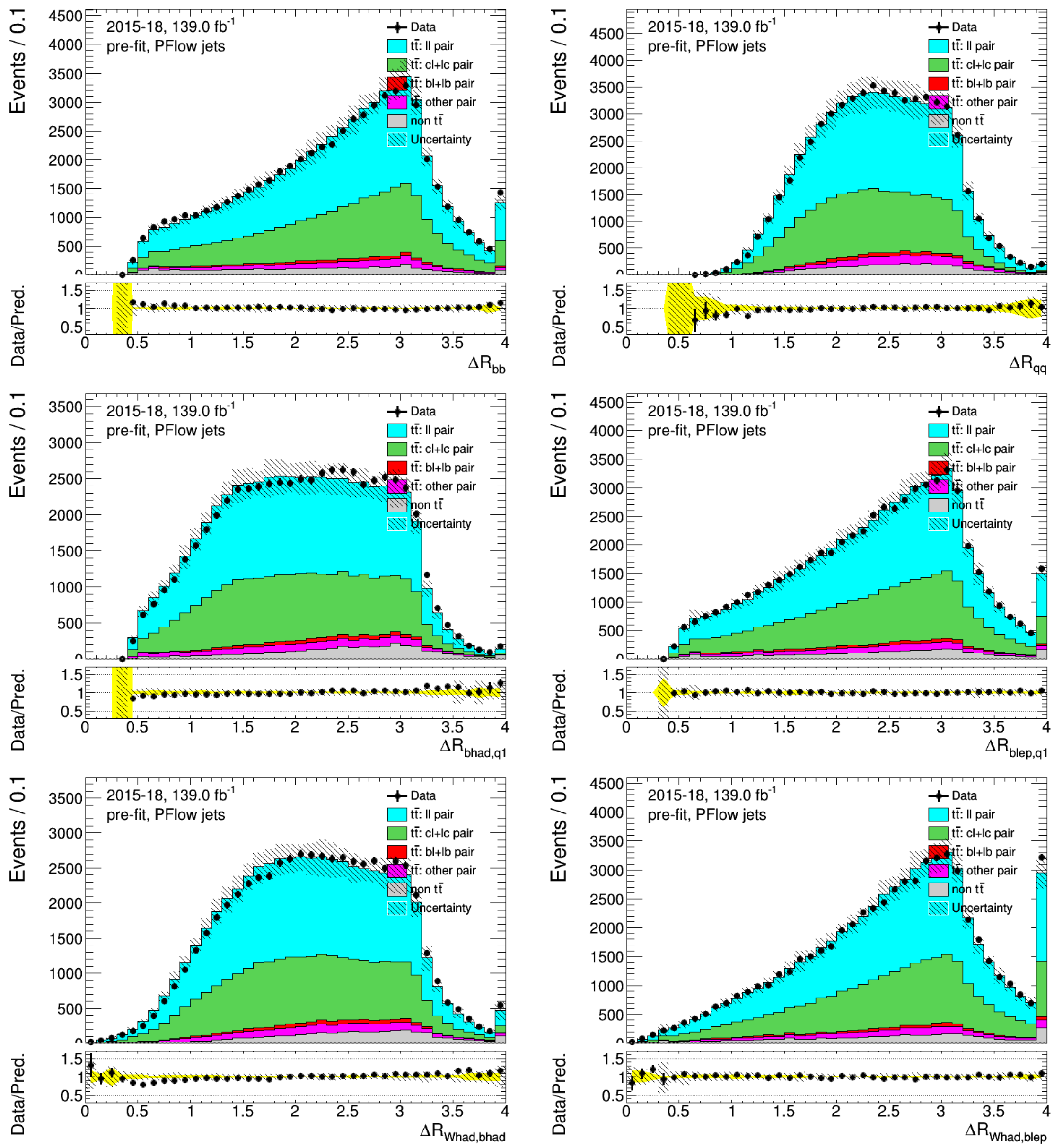


Figure 42: PFlow jets: distributions of angle related variables of the combination of the low-p_{T} selection, before fitting or tagging with full uncertainties.

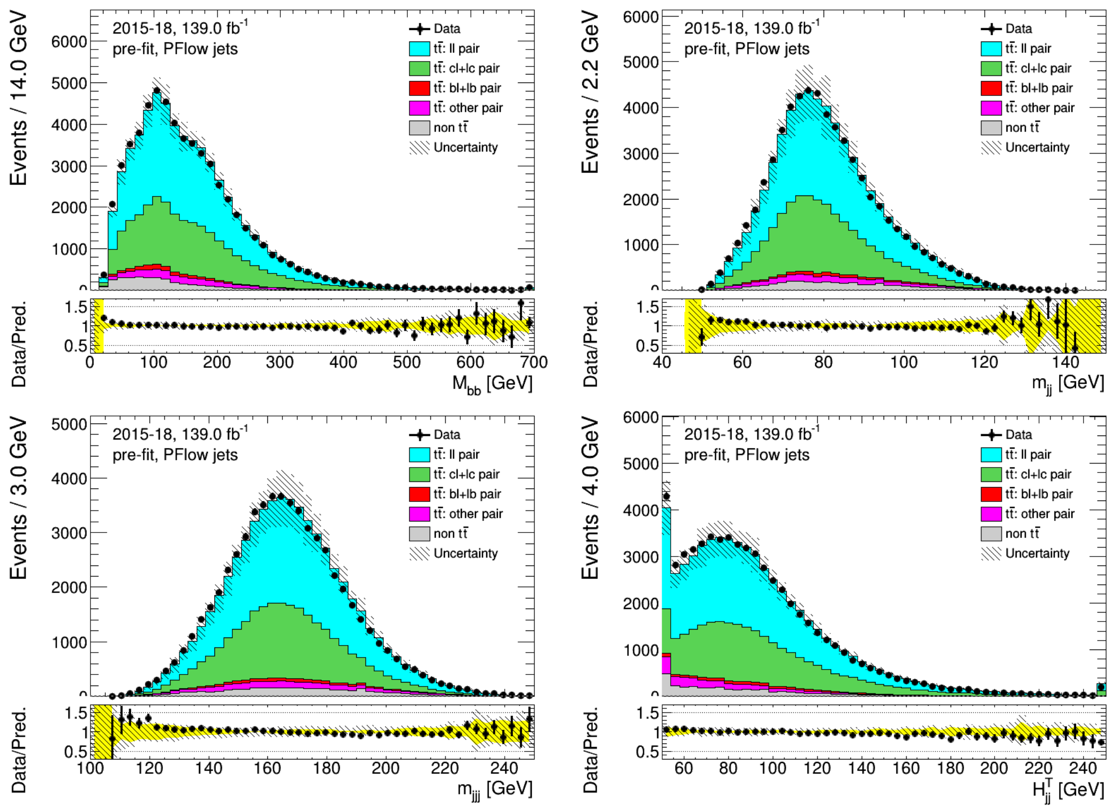


Figure 43: PFlow jets: distributions of mass related variables of the low-p_{T} selection, before fitting or tagging with stat-only uncertainties.

A.2 High-p_{T} selection

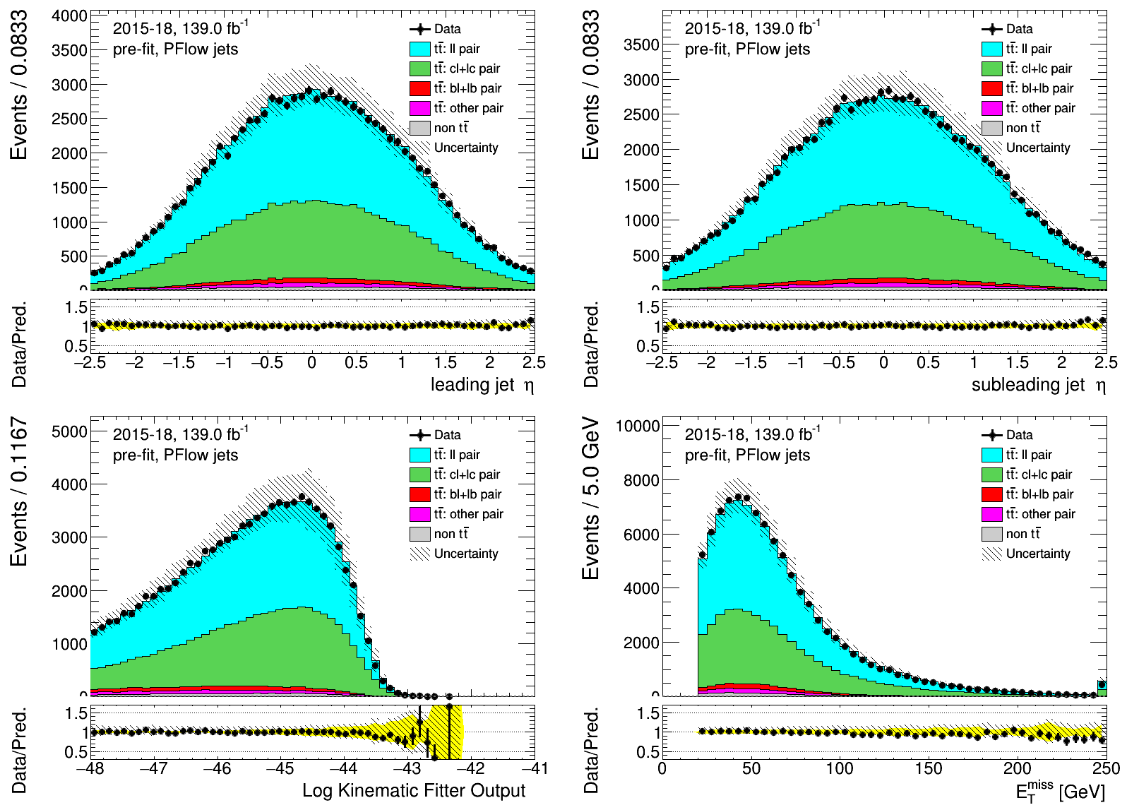


Figure 44: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the high-p_{T} selection, before fitting or tagging with full uncertainties.

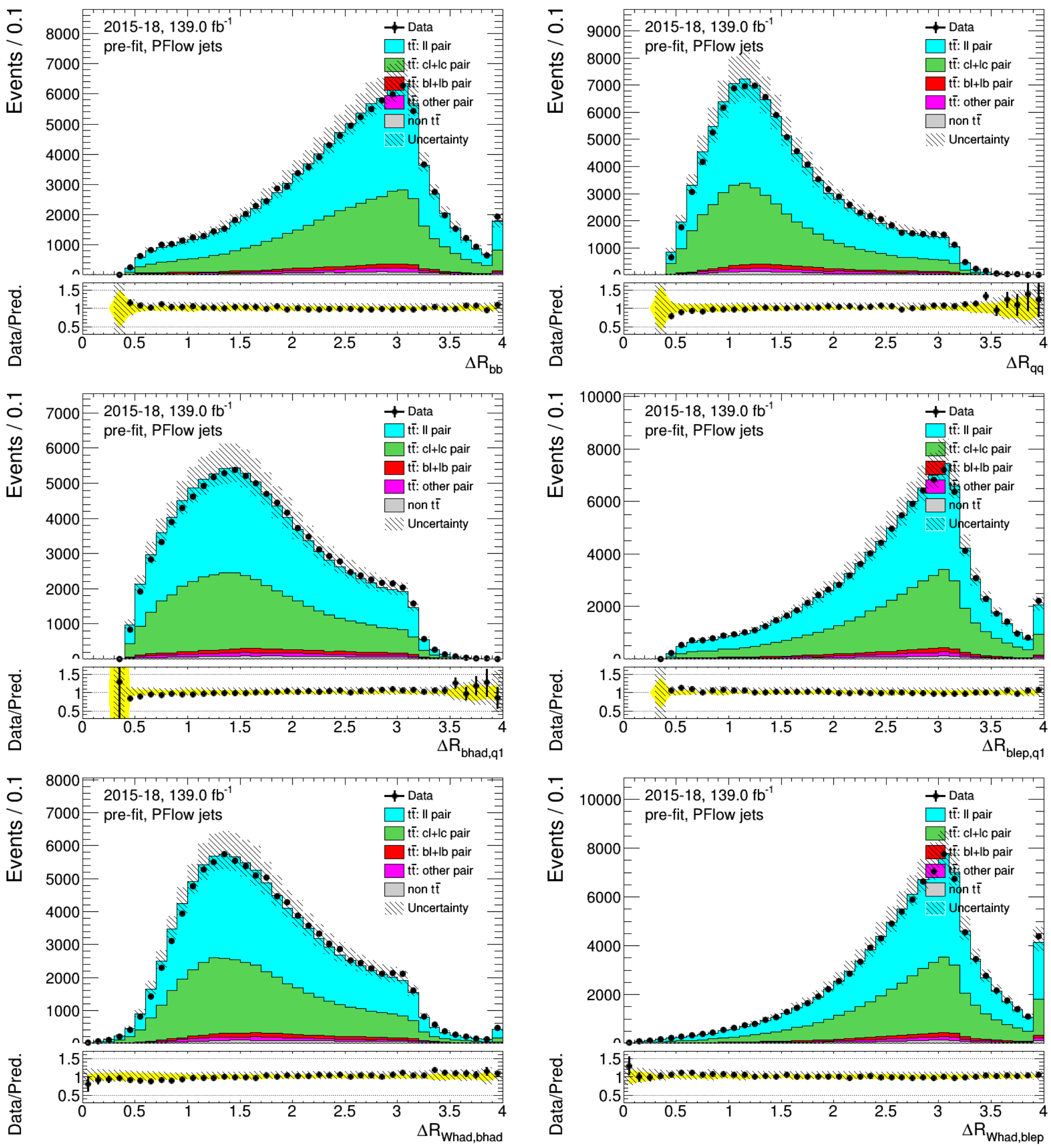


Figure 45: PFlow jets: distributions of angle related variables of the combination of the high-p_{T} selection, before fitting or tagging with full uncertainties.

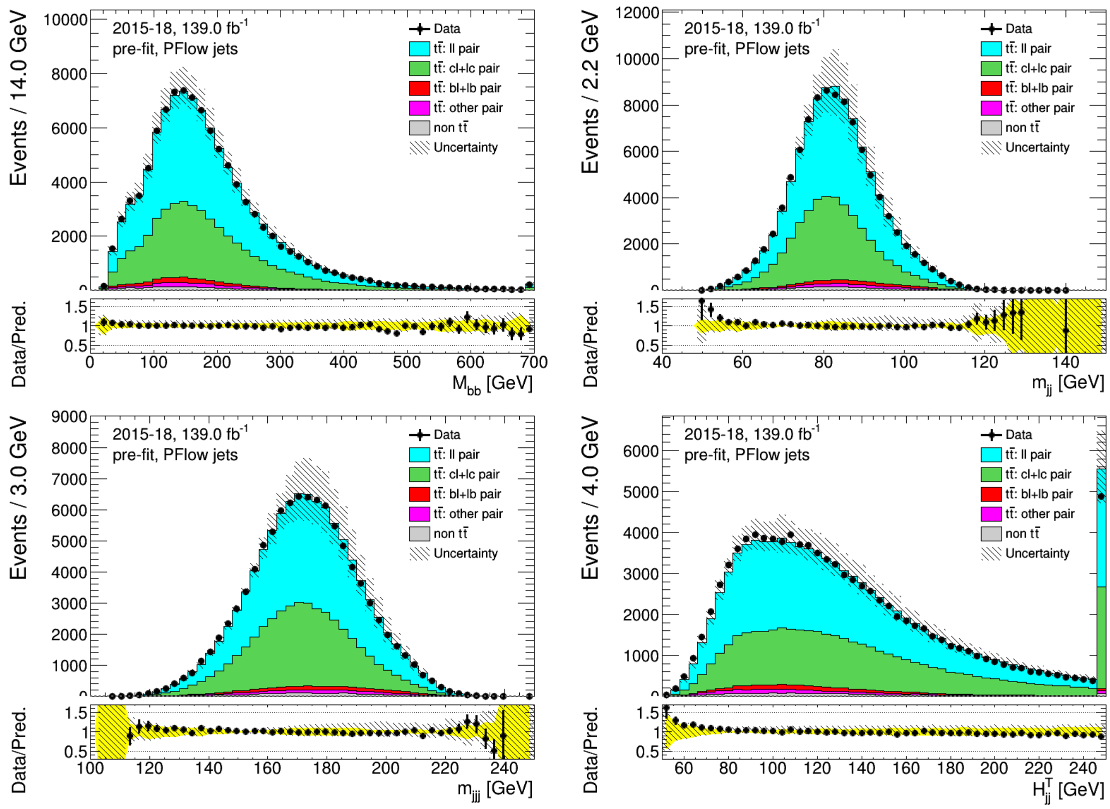


Figure 46: PFlow jets: distributions of mass related variables of the high-p_{T} selection, before fitting or tagging with stat-only uncertainties.

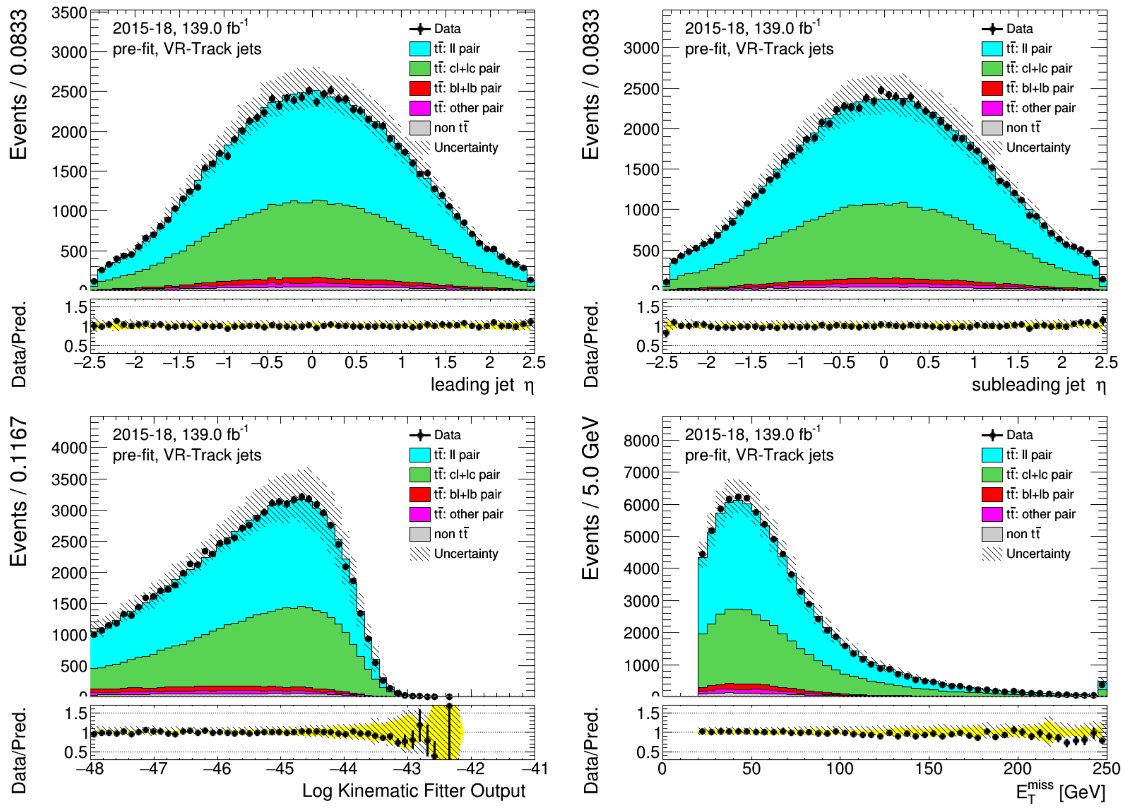


Figure 47: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the high-p_{T} selection, before fitting or tagging with full uncertainties.

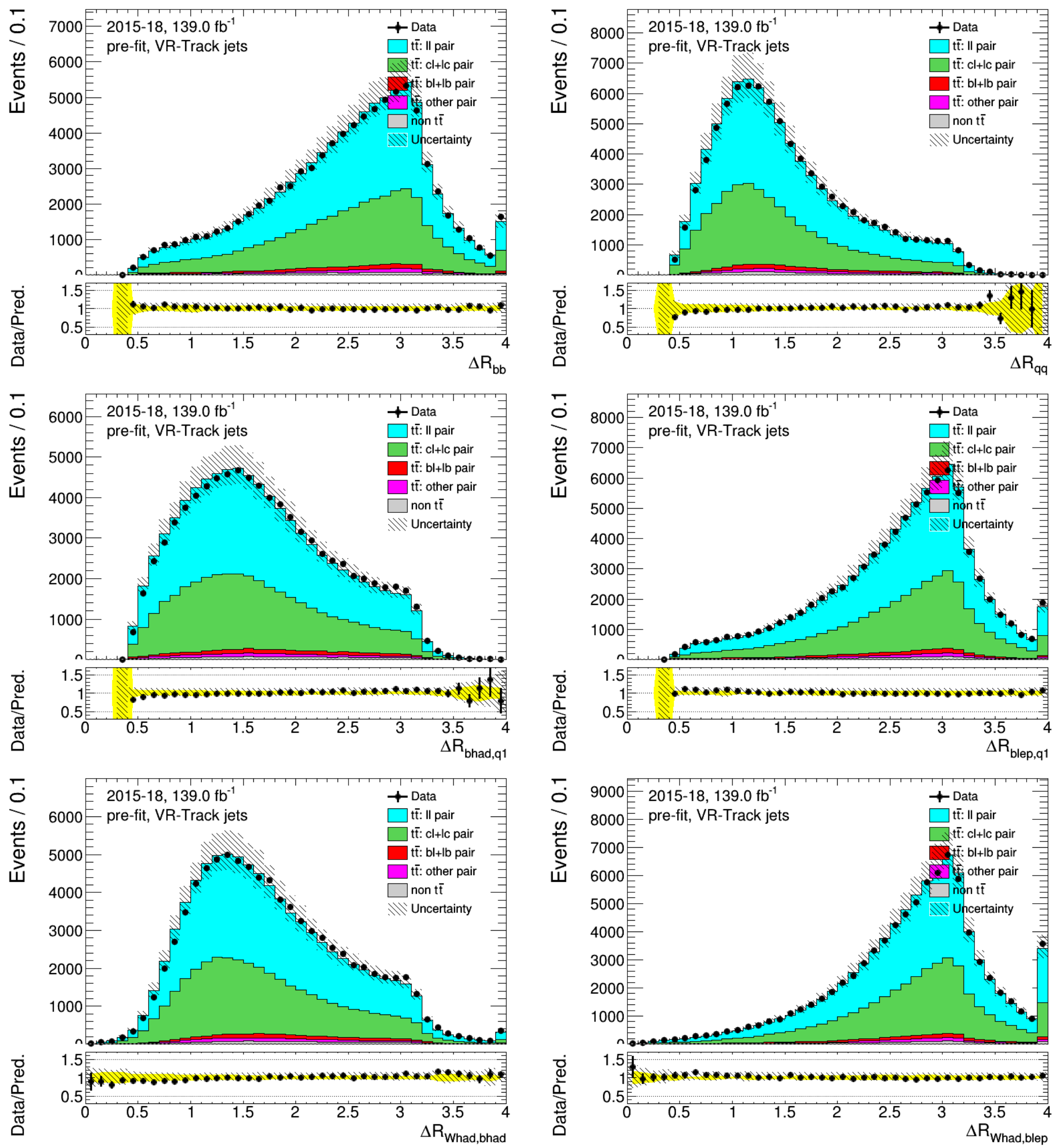


Figure 48: VR-Track jets: distributions of angle related variables of the combination of the high-p_{T} selection, before fitting or tagging with full uncertainties.

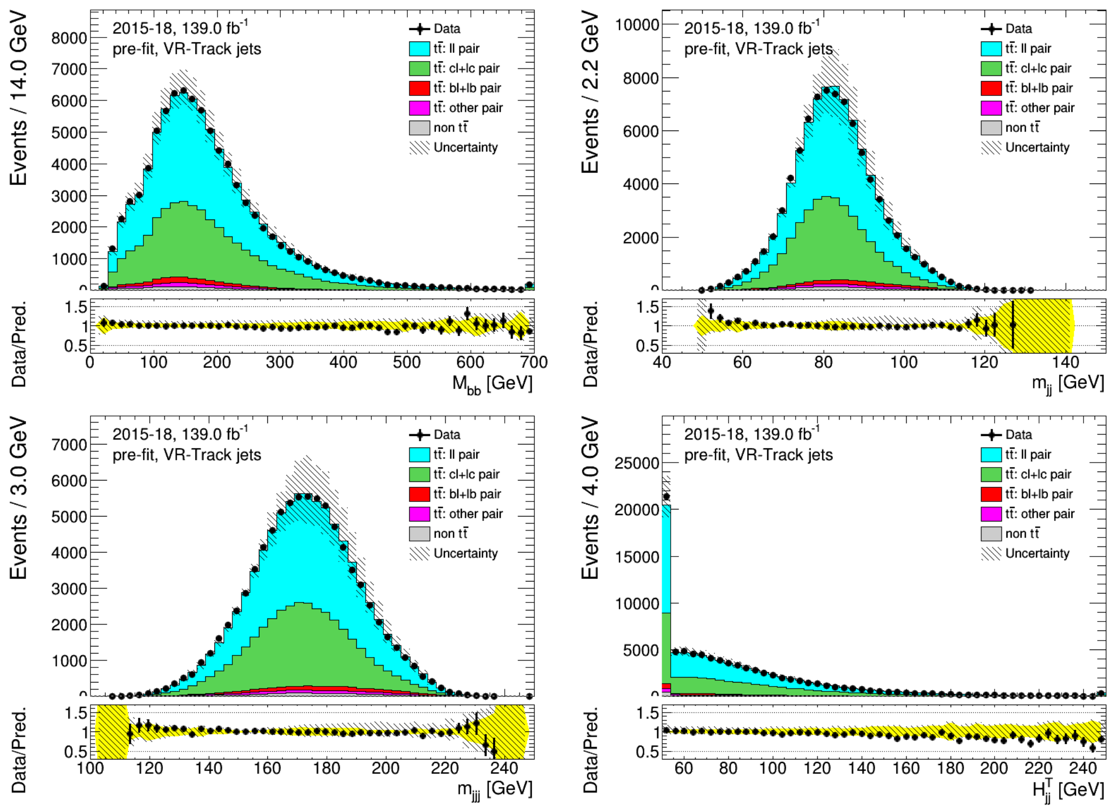


Figure 49: VR-Track jets: distributions of mass related variables of the high-p_{T} selection, before fitting or tagging with stat-only uncertainties.

A.3 Combined selection

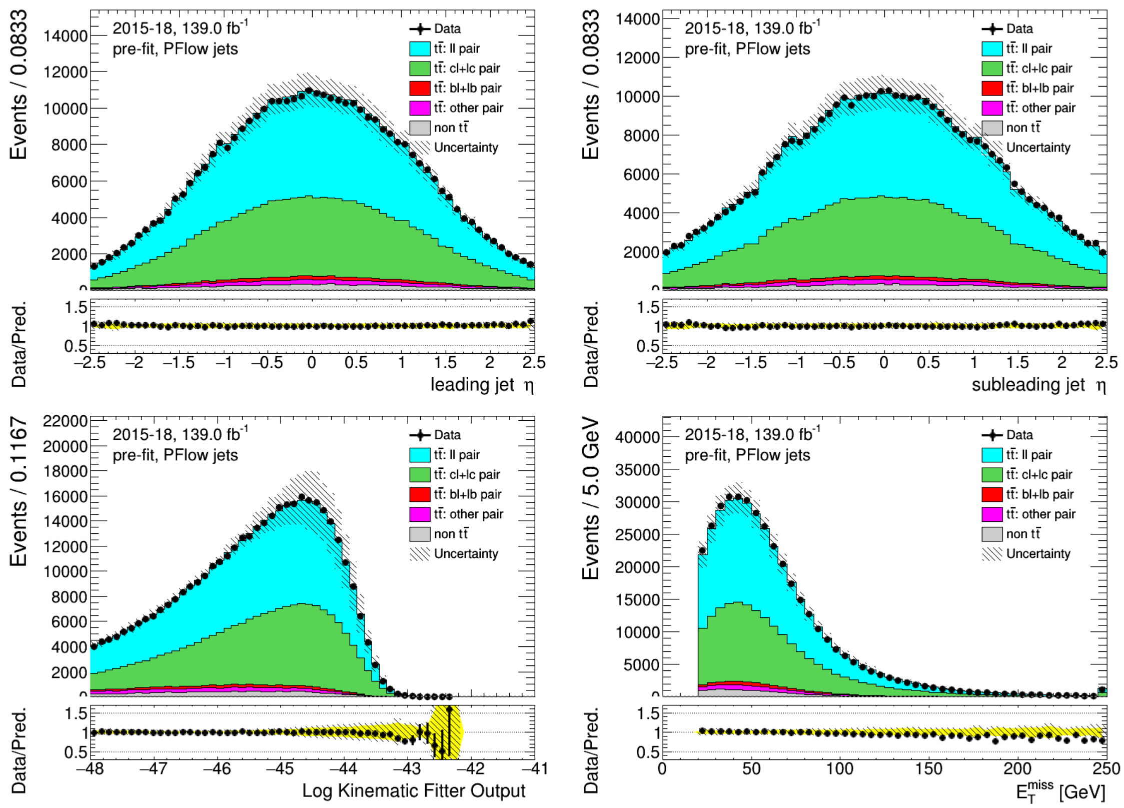


Figure 50: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combined selection, before fitting or tagging with full uncertainties.

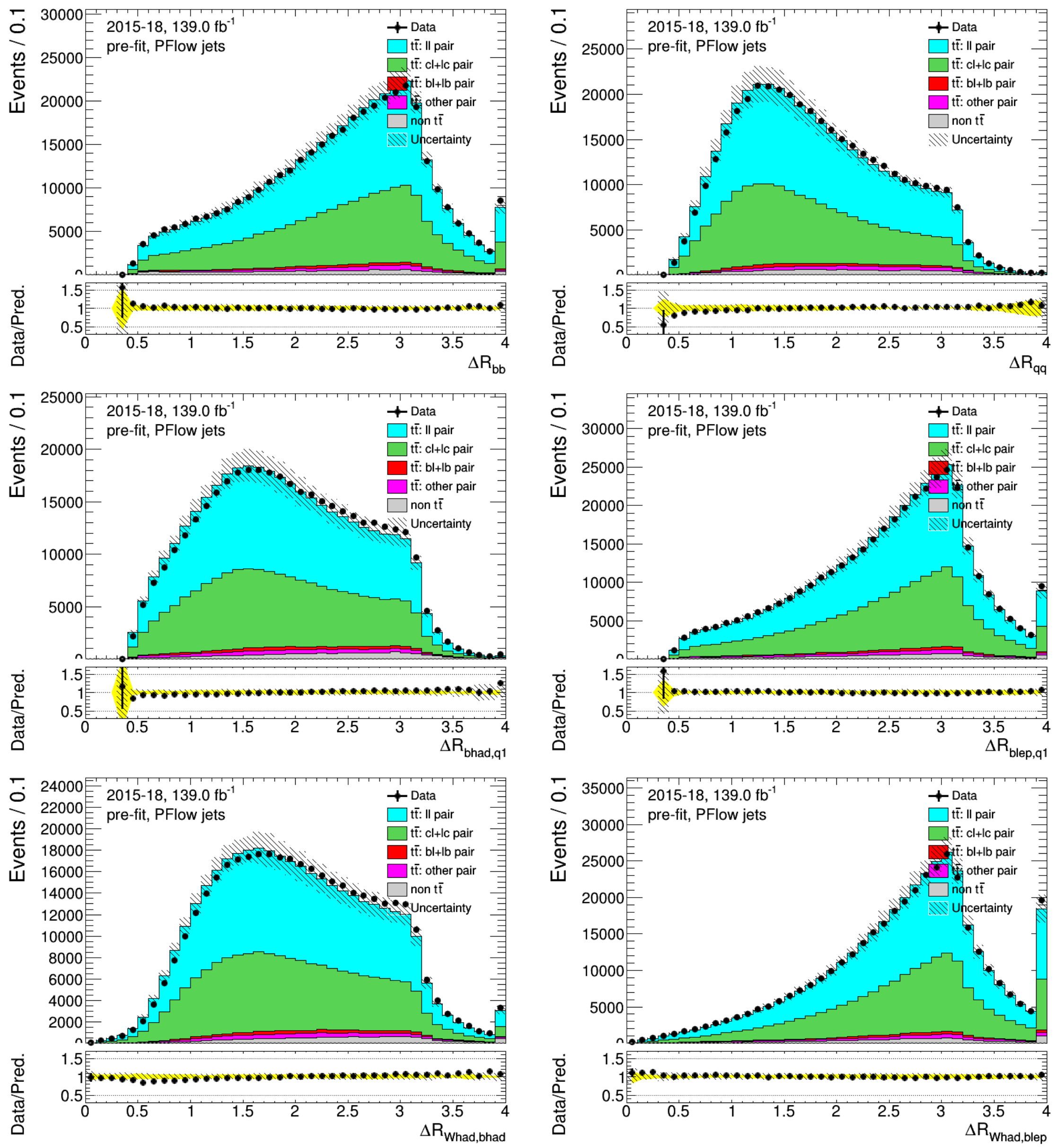


Figure 51: PFlow jets: distributions of angle related variables of the combination of the combined selection, before fitting or tagging with full uncertainties.

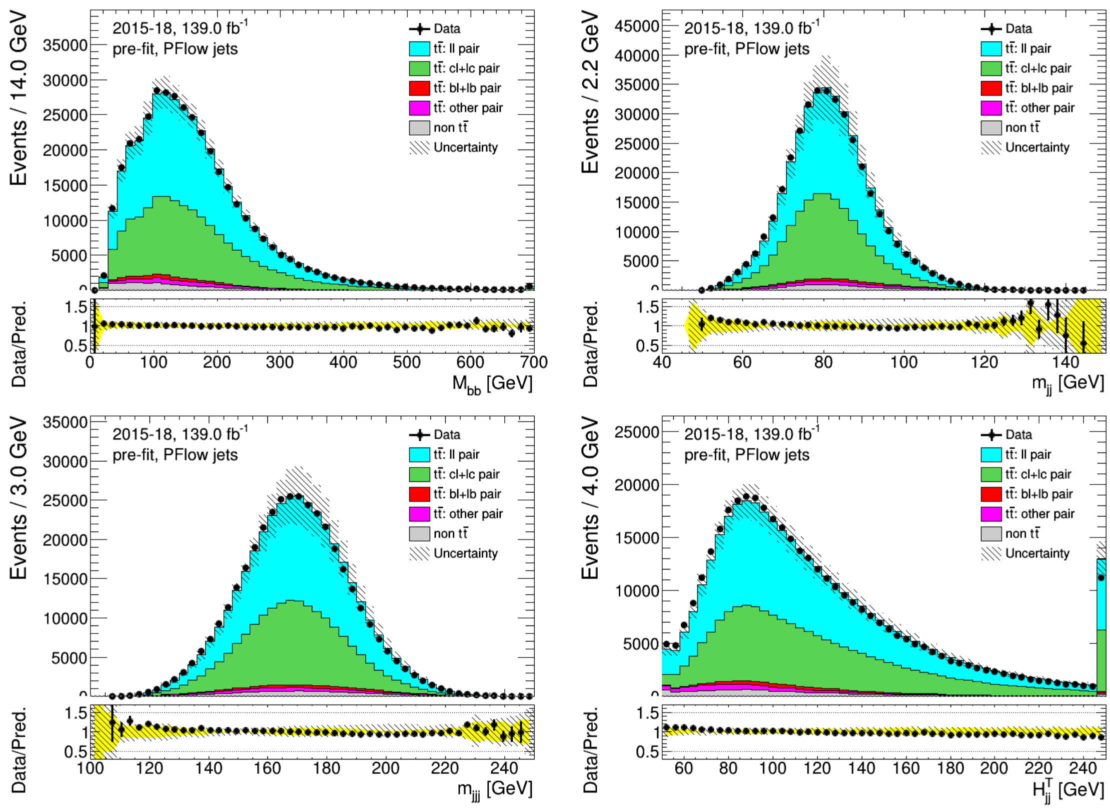


Figure 52: PFlow jets: distributions of mass related variables of the combined selection, before fitting or tagging with stat-only uncertainties.

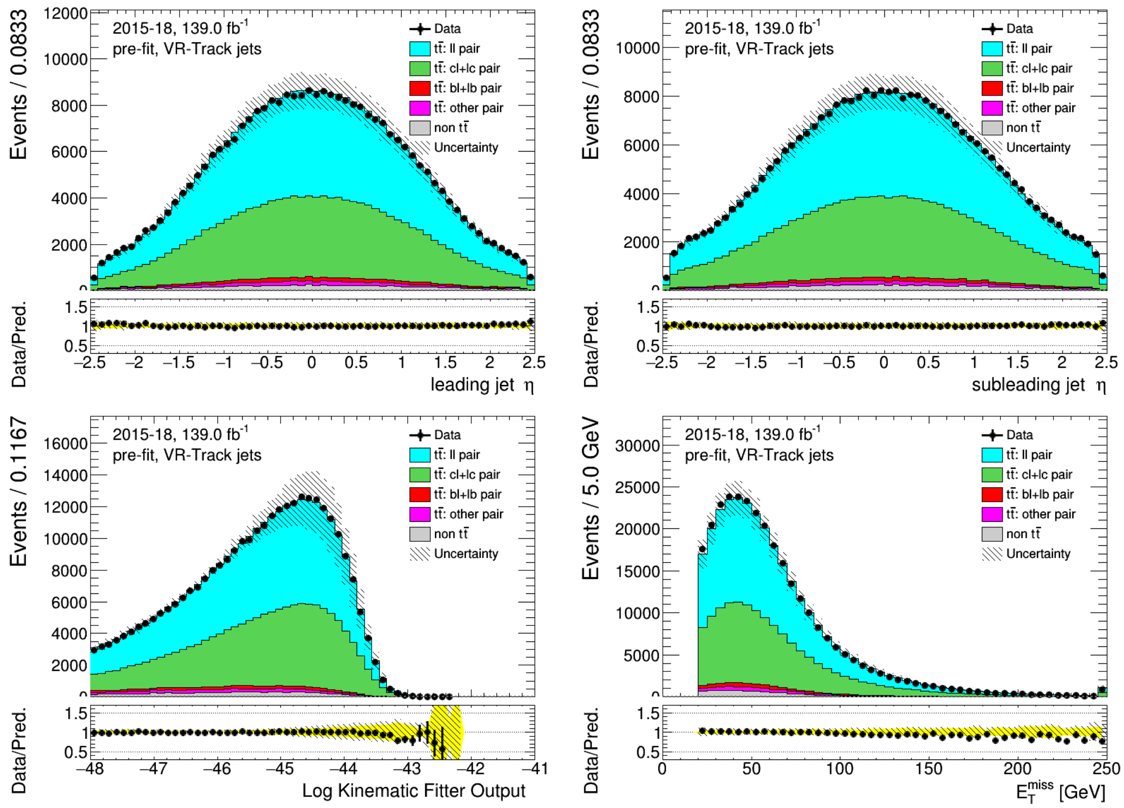


Figure 53: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combined selection, before fitting or tagging with full uncertainties.

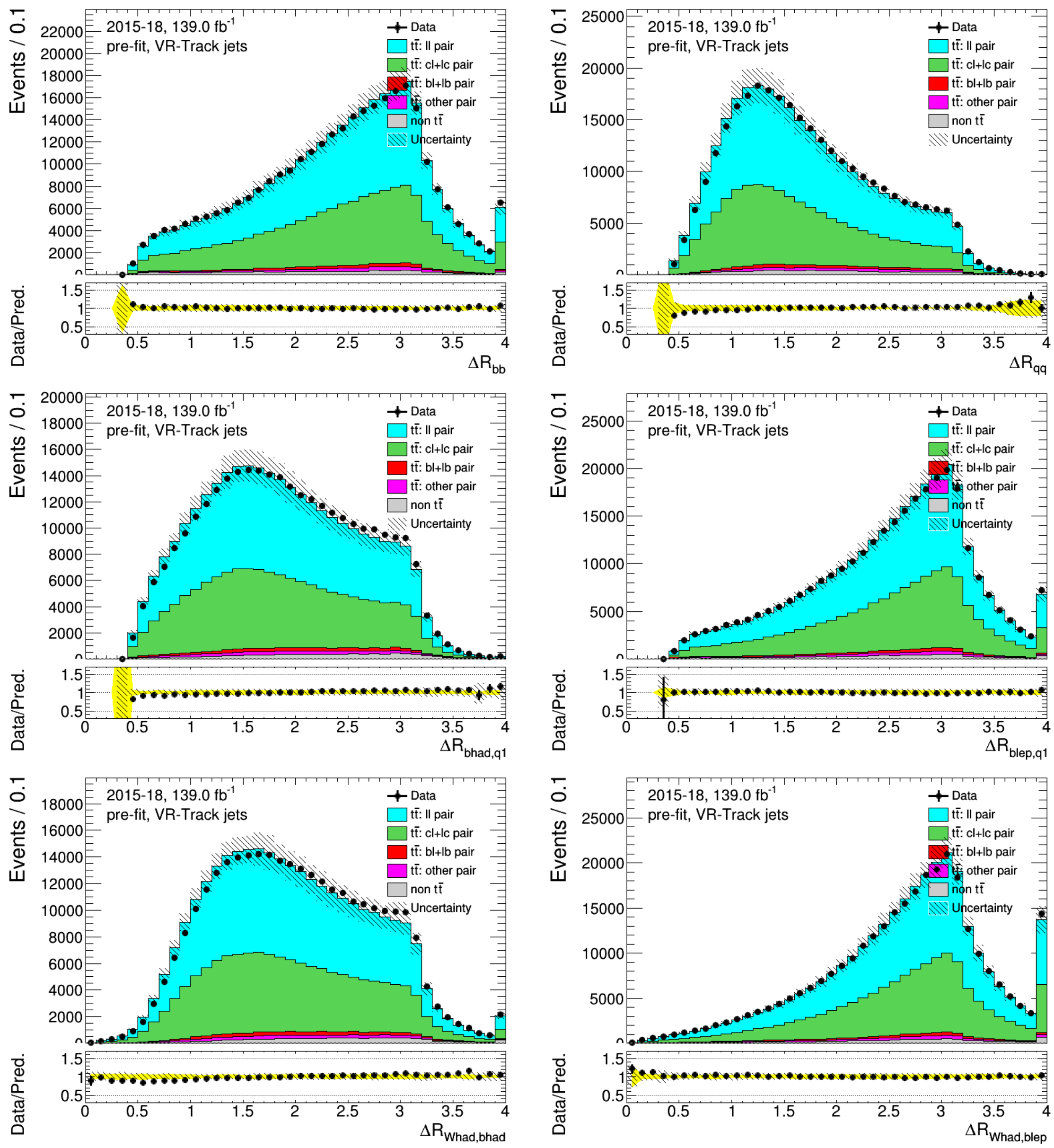


Figure 54: VR-Track jets: distributions of angle related variables of the combination of the combined selection, before fitting or tagging with full uncertainties.

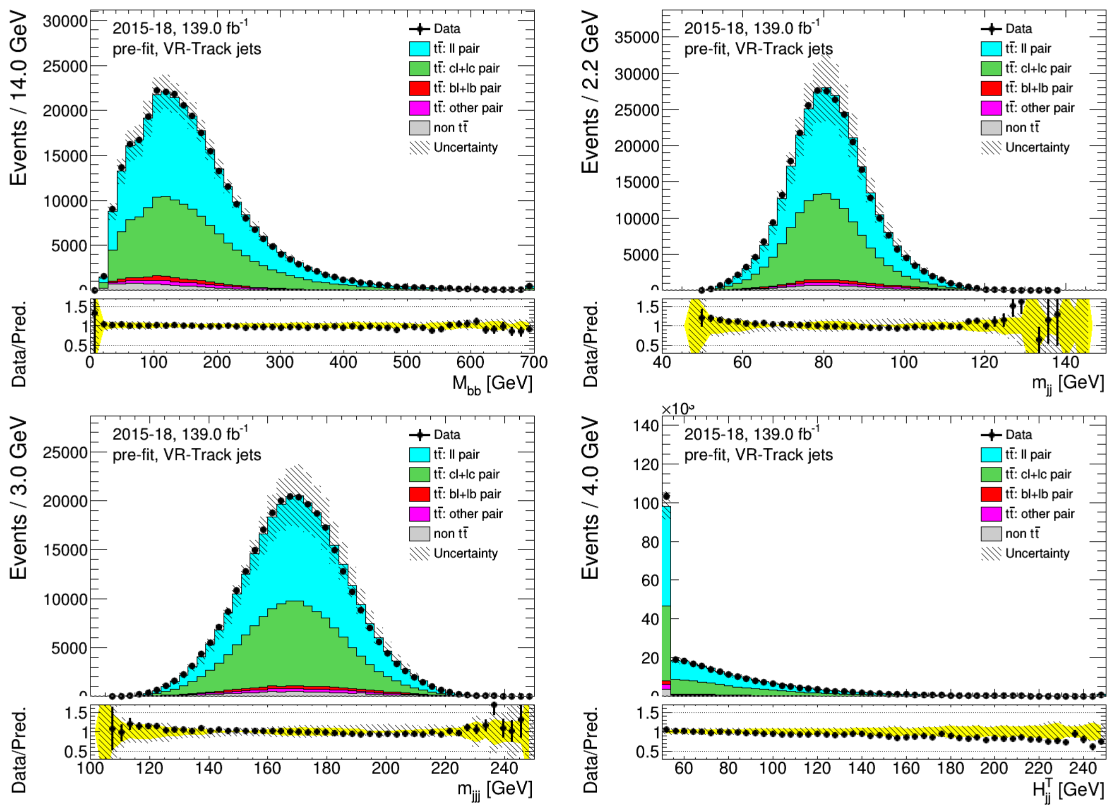


Figure 55: VR-Track jets: distributions of mass related variables of the combined selection, before fitting or tagging with stat-only uncertainties.

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 9: List of experimental systematics.