Doctorado en Fsica

Departamento de Fsica,
Universitat Autnoma de Barcelona
y
Institut de Física d'Altes Energies

Measurements of the charge asymmetry in top quark pair production at the LHC with the ATLAS detector

PhD dissertation

September 2014

Doctorando
Francesco Rubbo

Director
Prof. Aurelio Juste Rozas
Institució Catalana de Recerca i Estudis Avançats
Institut de Física dAltes Energies

Tutor
Prof. José María Crespo
Universitat Autònoma de Barcelona
Institut de Física dAltes Energies

Contents

Chapter 1

Introduction

Asymmetry is the lack of equality or equivalence between parts or aspects of something [1]. Symmetries are powerful tools in particle physics, because it has become evident that practically all laws of nature originates in symmetry. Therefore, violations of symmetry present theoretical puzzles leading to a deeper understanding of nature. From the experimental point of view, the measurement of asymmetries is virtually free from systematic uncertainties, and allow to distinguish small asymmetric signals from large symmetric backgrounds. In the recent history of particle physics, asymmetries have been exploited to investigate the nature of the weak interaction. Due to the parity violation typical of the weak interaction, processes show differences between particles and antiparticles, or between left-handed or right-handed particles. Various asymmetries have been studied at collider experiments in the last thirty years:

- A *left-right asymmetry* is defined as $A_{LR} = (N_L - N_R)\ \ (N_L + N_R)$, where N_L is the number of events with a left-polarized particle, and N_R the number of right-polarized events. The left-right asymmetry in Z boson production was measured in 1992 at the Stanford Linear Collider of polarized and unpolarized electron beams [2].
- A *charge asymmetry* is defined as $A_{+-} = (N_+ - N_-)\ \ (N_+ + N_-)$, where N_+ is the number of events with a positively charged particle, and N_- the number of events with a negative particle. Charge asymmetries are used to measure CP violation in B meson decays.
- A *forward-backward asymmetry* is defined as $A_{FB} = (N_F - N_B)\ \ (N_F + N_B)$, where N_F is the number of events with a particle moving *forward* with respect to a chosen reference direction, and N_B the number of *backward* events.

The latter is the type of asymmetry more relevant for this dissertation. A famous example of forward-backward asymmetry is the measurement in $\mu^+\mu^-$ production in e^+e^- annihilation by the TASSO collaboration [3]. The measurement, performed at center-of-mass energies between 35 and 46 8 GeV, showed how the asymmetry provides sensitivity to probe for physics at higher energy, beyond reach for direct observation. In the specific case, Fig. 1.1 shows that, while the inclusive cross section measurements agree with the symmetric QED description, a forward backward asymmetry is induced by the interference with a virtual Z boson discovered at a mass of ~ 91 GeV in proton antiproton collisions by the UA1 and UA2 collaborations [4, 5].

An analogous asymmetry in top quark pair ($t\bar{t}$) production might provide a probe for new physics processes at scales beyond the current reach of direct searches. The CDF and Dfi collaborations reported an unexpectedly large forward backward asymmetry in $t\bar{t}$ production

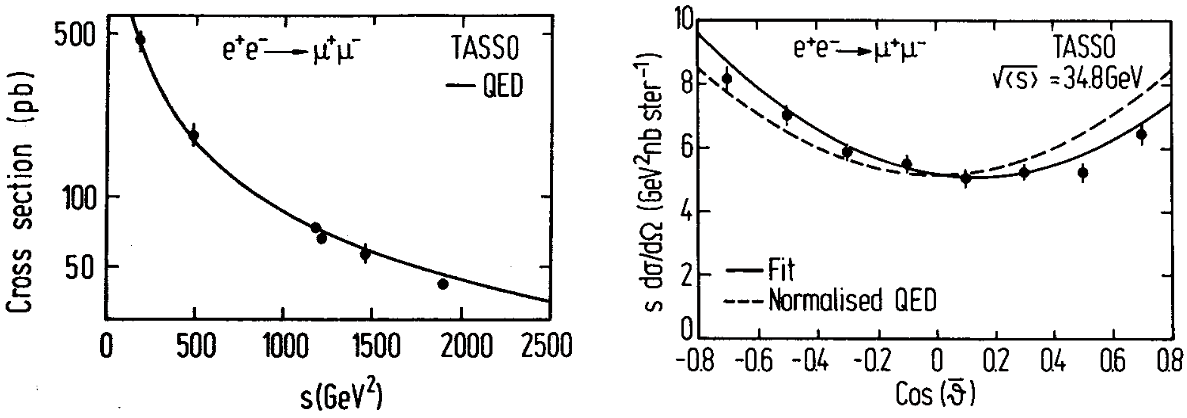


Figure 1.1: Inclusive cross section as a function of center-of-mass energy (left) and angular distribution of the muons (right) for $\mu^+\mu^-$ production at e^+e^- collisions [3].

using proton–antiproton collisions at $\sqrt{s} = 1.96$ TeV at Fermilab's Tevatron collider [6, 7]. As shown in Fig. 1.2, the asymmetry has a dependence on the $t\bar{t}$ invariant mass (a proxy for the partonic center–of–mass energy) with a interesting hint of discrepancy with respect to the Standard Model prediction. With the end of Tevatron operations in 2011, no additional

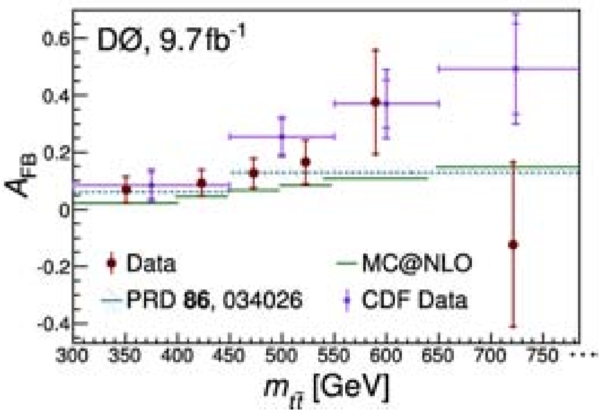


Figure 1.2: Dependence of the forward–backward asymmetry measurements by the CDF and D0 Collaborations on the invariant mass of the $t\bar{t}$ system. The horizontal error bars indicate the binning used in each experiment. Also shown is the Standard Model prediction.

data will be collected. Therefore little room is left for improving the precision of the forward–backward asymmetry measurement at $p\bar{p}$ collisions, with the goal of clarifying the experimental status. However, complementary probes of the asymmetry in $t\bar{t}$ production can be studied at pp collisions, provided by the Large Hadron Collider at CERN.

This thesis work focus on the study of a central–forward asymmetry in $t\bar{t}$ production at pp collisions, also referred to as *charge* asymmetry. This document is organized as follows: Chapter 2 gives an introduction to the Standard Model of particle physics, with a focus on the top quark phenomenology and the production asymmetry. Chapter 3 describes the experimental setup of the ATLAS detector within the Large Hadron Collider accelerator facility. The simulation of the physics processed used in the analysis is described in Chapter 4, while the definition of the physics objects follows in Chapter 5. Chapter 6 illustrates the analysis strategy followed by the description of the event selection and the background determination. The reconstruction of the $t\bar{t}$ system kinematics is described in Chapter 7, while Chapter 8 describes the unfolding procedure used to account for acceptance and resolution effects. The results of the inclusive and differential measurements of the $t\bar{t}$ charge asymmetry using the full LHC Run 1 dataset collected with the ATLAS detector are summarized in Chapter 9, along with a comparison to other existing measurements at the LHC. Finally the summary of this thesis and an outlook is given in Chapter 10.

Chapter 2

The top charge asymmetry in theory and experimental results

2.1 The Standard Model

The Standard Model (SM) is a quantum field theory describing the interactions between the elementary constituents of matter through the fundamental forces. The elementary particles three *leptons*, three *neutrinos* and six *quarks* are fermions categorized in three families. For each particle there exist an antiparticle with same mass and opposite quantum numbers for a total of 24 particles. All particles carry a weak charge and interact with the weak force carriers W and Z. All of the particles, except for neutrinos, are electrically charged and interact with the electromagnetic force, mediated by γ. Only the quarks have a color charge and undergo strong force interactions with *gluons*. Table 2.1 shows the elementary particles and the fundamental force carriers of the Standard Model.

Matter					Force carriers
	Electric Charge	Family 1	Family 2	Family 3	
quarks	$+2\ 3$	u	c	t	γ
	$-1\ 3$	d	s	b	$W^{\pm}$
leptons	-1	e	μ	τ	Z
	0	ν_e	ν_μ	ν_τ	g

Table 2.1: Table of particles and forces in the Standard Model

The theory describing the particle interactions is structured according to the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The three terms represent the fundamental symmetries corresponding to the forces describing the interaction of particles: $SU(3)_C$ is the *color* (C) symmetry corresponding to the strong force; $SU(2)_L \times U(1)_Y$ is the spontaneously broken symmetry with respect to the *isospin* (L) and *hypercharge* (Y) gauge groups. The SM is then the combination of two theories: the *Quantum Chromodynamics* (QCD) lagrangian, which describes the

Spin-1 gauge bosons	mass	Spin-0 scalar boson	mass
γ	0	H	~ 125 GeV
g	0		
$W^{\pm}$	$80\ 385 \pm 0\ 015$ GeV		
Z	$91\ 188 \pm 0\ 002$ GeV		

Spin-$\frac{1}{2}$ fermions	I generation		II generation		III generation	
leptons	ν_e	~ 0	ν_μ	~ 0	ν_τ	~ 0
	e	0 511 MeV	μ	105 7 MeV	τ	1 777 GeV
quarks	u	1 7 − 3 1 MeV	c	$1\ 29^{+0.05}_{-0.11}$ GeV	t	$173\ 3 \pm 0\ 8$ GeV
	d	4 1 − 5 7 MeV	s	100^{+30}_{-20} MeV	b	$4\ 19^{+0.18}_{-0.06}$ GeV

Table 2.2: Mass values for the elementary particles of the Standard Model, as measured at experiments.

strong interaction arising from the $SU(3)$ color symmetry, and the *Electroweak* lagrangian, which accounts for the electroweak interactions corresponding to the $SU(2) \times U(1)$ isospin and hypercharge symmetries.

$$\begin{aligned}\mathcal{L}_{SM} &= \mathcal{L}_{SU(3)} + \mathcal{L}_{SU(2)\times U(1)} \\ &= \mathcal{L}^{Gauge}_{SU(3)} + \mathcal{L}^{Matter}_{SU(3)} + \mathcal{L}^{Gauge}_{SU(2)\times U(1)} + \mathcal{L}^{Matter}_{SU(2)\times U(1)} + \mathcal{L}^{Higgs}_{SU(2)\times U(1)} + \mathcal{L}^{Yukawa}_{SU(2)\times U(1)}\end{aligned} \tag{2.1}$$

The $\mathcal{L}^{Gauge}$ terms describe the dynamics of the gauge fields: the gluons in QCD, and the W, Z and γ boson in the electroweak theory. The $\mathcal{L}^{Matter}$ terms describe the interaction of particles with the gauge fields. The $\mathcal{L}^{Higgs}$ and $\mathcal{L}^{Yukawa}$ terms arise from the spontaneous symmetry breaking of the $SU(2) \times U(1)$ gauge theory via the Higgs mechanism, and they are responsible, respectively, for the interaction of the Higgs field with the other particles and force carriers, and for generating their masses. The Higgs field needs to be introduced in the $SU(2) \times U(1)$ theory in order to account for the non-vanishing masses of W and Z bosons and of lepton and quarks. The Higgs boson does not interact with the gluon and the photon; therefore they are the only two massless particles of the SM. Table 2.2 summarize the masses of the SM particles.

2.2 The top quark

The top quark was discovered in 1995 at the Tevatron by the CDF [8] and DØ [9] collaborations. With a mass of ≈ 170 GeV it is the heaviest elementary particle observed so far. Its proximity to the electroweak scale suggests that the top quark might play a relevant role in new physics scenarios. Therefore its production, decay modes and properties have been studied in details at the two hadron colliders powerful enough to produce it: the Tevatron, where protons and antiprotons were collided ($p\bar{p}$), and the Large Hadron Collider (LHC), colliding protons.

2.2.1 Pair production at hadron colliders

At hadron colliders, top quarks are mainly produced in pairs through QCD processes. The partons constituents of the colliding hadrons (protons or antiprotons) participate in a hard scattering process and produce a top quark and an antitop quark. The leading production modes are: *gluon–gluon fusion* and *quark–antiquark annihilation.* The leading order Feynman diagrams for these two processes are shown in Fig. 2.1. The following discussion focuses on the top quark pair production in proton collisions; however similar considerations apply to the case of proton-antiproton collisions.

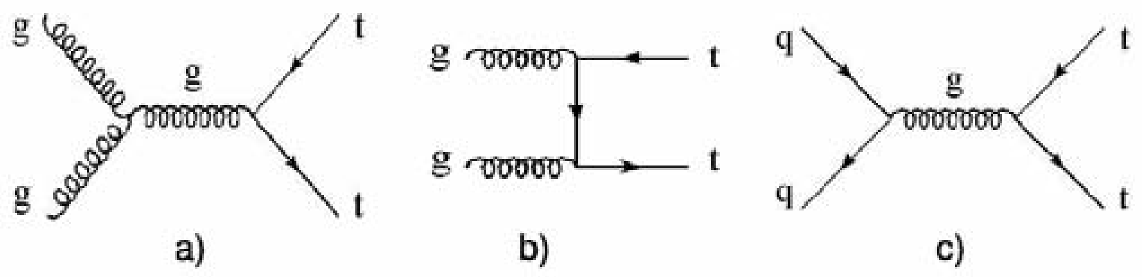


Figure 2.1: Leading order diagrams for QCD top quark pair production. Gluon fusion, `a)` and `b)`, is the dominant process at LHC energies, while quark antiquark annihilation, `c)`, is the dominant one at Tevatron energies [10].

Due to the composite nature of the proton, most of the collisions involve only soft (i.e. long distance) interactions of the constituent quarks and gluons. Such interactions cannot be described with perturbative QCD because the expansion parameter α_s is large for small momentum exchange. In some collisions, however, in addition to the long distance parton interactions, two quarks or two gluons undergo a hard (i.e. short distance) scattering where a top quark pair is produced. The *factorization principle* [11] states that the perturbative QCD description of this hard process is possible by factorizing the long distance effects into functions f describing the proton structure. Thus the cross section for top quark pair production in a collision of two protons with momenta P_1 and P_2 is given by:

$$\sigma(p(P_1)+p(P_2) \quad t\bar{t}+\mathrm{X}) = \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \sum_f f_f(x_1) f_f(x_2) \cdot \sigma(q_f(x_1 P_1)+\bar{q}_f(x_2 P_2) \quad t\bar{t}+\mathrm{X}) \tag{2.2}$$

where the sum runs over all type of partons, and $x_{1,2}$ are the fractions of proton momentum carried by the constituents participating in the hard interaction.

The *parton distribution functions* (PDFs) $f_f(x \ \mu_F^2)$ represent the probability density for a parton of a type i to carry a fraction of proton momentum x. The *factorization scale* μ_F^2 is set arbitrarily and defines the distinction between short and long distance interactions based on the transferred momentum Q^2. An additional renormalization scale μ_R^2 accounts for higher order corrections. For calculations and simulations, both scales are set to the typical transferred momentum of the process studied. In the case of $t\bar{t}$ production the scale is chosen to be equal to the top mass m_t.

The PDFs cannot be computed in perturbation theory; instead, they are measured in deep inelastic scattering experiments and at hadron colliders. For the studies presented in this thesis, the PDF estimations provided by the CTEQ [12], MRST [13], MSTW [14], CT10 [15] and NNPDF [16] collaborations are used. Fig. 2.2 shows the PDFs of valence quarks, gluons and sea quarks for two values of transferred momentum Q^2 at which the proton is probed.

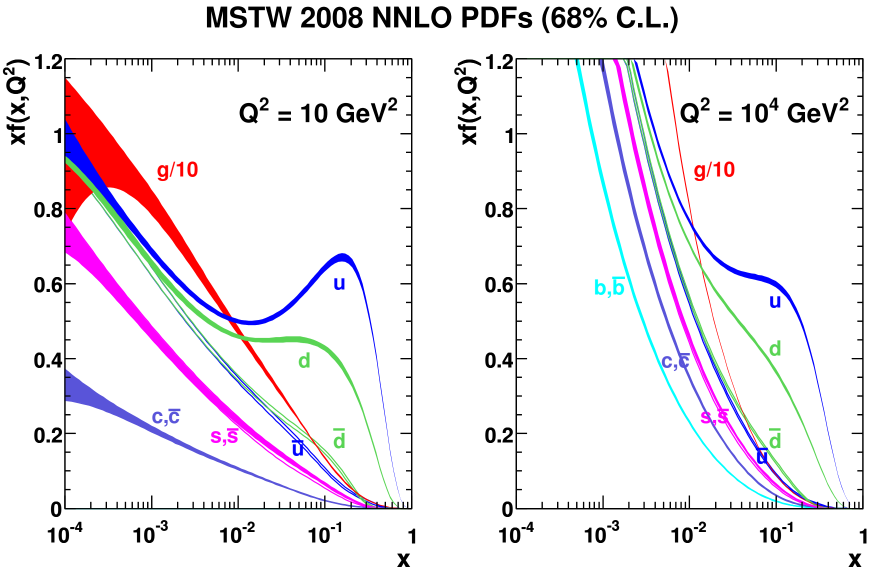


Figure 2.2: MSTW parton distribution functions [15] for gluons and quarks at $Q^2 = 10$ GeV (left) and $Q^2 = 10$ TeV (right).

The typical fraction of momentum x carried by each of the colliding partons in order to produce a top quark pair is defined by the relationship

$$\sqrt{x_1 x_2 s} \geq 2m_t \tag{2.3}$$

Therefore, assuming the partons carry a similar fraction of momentum $x_1 \approx x_2 \approx x$:

$$x = \frac{2m_t}{\sqrt{s}} \tag{2.4}$$

This corresponds to a typical value of $x \approx 0.05$ at the LHC for a center-of-mass energy $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. As shown in Fig. 2.2, the probability of gluon collisions is significantly larger than for any other parton in the corresponding range of x. Thus the production of top quark pairs at the LHC is dominated by the gluon fusion process ($\approx 90\%$). On the contrary, at the Tevatron, the typical value of $x \approx 0.2$ makes the valence quark-antiquark annihilation the prevalent $t\bar{t}$ production mode. The total $t\bar{t}$ production cross section at the LHC is calculated at next-to-next-to leading order in QCD to be $177.31^{+10.1}_{-10.8}$ pb at $\sqrt{s} = 7$ TeV and $252.89^{+13.30}_{-14.52}$ pb at $\sqrt{s} = 8$ TeV, for a top quark mass of 172.5 GeV [17–23].

Measurements have been performed by both ATLAS and CMS collaborations, yielding a combined value of 173.3 ± 10.1 pb at $\sqrt{s} = 7$ TeV [24, 25]. Preliminary measurements at $\sqrt{s} = 8$ TeV yield $\sigma_{t\bar{t}} = 242 \pm 9$ pb (ATLAS) [26] and $\sigma_{t\bar{t}} = 239 \pm 13$ pb (CMS) [27]. All measurements are in agreement with the SM predictions above.

2.2.2 Decay

The top quark has an extremely short lifetime of 5×10^{-25} s. Therefore decay occurs before hadronization can take place, and the decay products carry all of the information about 4 momentum and spin of the original particle. The top quark decays in almost all cases via electroweak charged current interaction into a b quark and a W boson, which, in turn, decays either leptonically, into a charged lepton and the corresponding antineutrino, or hadronically into a quark antiquark pair. Thus the final states corresponding to a $t\bar{t}$ pair can be classified in three categories, based on the decays of the two W bosons originating from the decays of the top and the antitop quarks:

- Full hadronic final state: both W bosons decay into quarks, leading to a $t\bar{t}$ final state with six quarks.
- Semileptonic final state: one W boson decays into quarks, while the other decays leptonically, leading to a final state with four quarks, one lepton and one neutrino.
- Dileptonic final state: both W bosons decay leptonically, leading to a final state with two b quarks, two leptons and two neutrinos.

Given that the W boson hadronic branching ratio is $\approx 2\ 3$, the full hadronic and semileptonic final states occur 4 out of 9 times each, while the dileptonic decay has the remaining 1 9 probability.

2.2.3 Properties

The properties of the top quark have been studied in detail at hadron collider experiments. The combination of Tevatron and LHC results brought the precision on the top mass measurement well below 1 GeV with $m_t = 173\ 3 \pm 0\ 8$ GeV [28]. An exotic electric charge of $4e\ 3$ for the top quark, for which the SM predicts a $2e\ 3$ charge, has been excluded by measurements at both Tevatron experiments [29, 30] and at the ATLAS experiment [31]. Since the top quark does not form bound hadronic states due to its short lifetime, it is the only quark whose spin properties, which are usually concealed by hadronization, can be measured. The final state particles in the decay carry information about the top quark spin; therefore, it s possible to measure the top polarization and the spin correlation of the top quark pair. At the current precision, these quantities have been found to be compatible with the SM predictions [32 34].

2.3 Charge asymmetry in top quark pair production

The measurements of the top quark properties listed in Sec 2.2.3 yield results compatible with SM predictions. A different trend appeared to follow the *forward-backward* (FB) asymmetry in the production of $t\bar{t}$ pairs in $p\bar{p}$ collisions at the Tevatron, The first precise measurements, using half of Tevatron dataset, showed large discrepancies between data and theory ($\gtrsim 3\sigma$). The asymmetry A_{FB} is defined in terms of the rapidity y of the top and antitop quarks in the laboratory frame[1]:

$$A_{FB} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)} \tag{2.5}$$

---

[1] The rapidity of a particle at colliders is given by $y = \frac{1}{2} log \frac{E+p_z}{E-p_z}$, where E is the energy of the particle and p_z the component of its momentum along the beam axis z.

with $\Delta y = y_t - y_{\bar{t}}$ and N number of events. At present, when the full Tevatron dataset has been analyzed, the discrepancies have been reduced with respect to previous results. The CDF collaboration reports a $1\ 7\sigma$ excess over the SM prediction [7], whereas the Dfi collaboration finds agreement within 1σ [6].

At the time when this thesis work began, a discrepancy between experimental data and the SM prediction over three standard deviations was reported by the CDF collaboration in the measurement of A_{FB} at high $t\bar{t}$ invariant mass $m_{t\bar{t}} > 450$ GeV, using half of the total dataset. This anomaly triggered a intense activity, both in developing new physics models which could explain the anomaly in $t\bar{t}$ production, and in calculating more precise and accurate SM predictions. The result also motivated studies of the $t\bar{t}$ production phenomenology aiming at identifying other observables where anomalies might appear if the Tevatron asymmetry were indeed a sign of new physics. Among these, the charge asymmetry at the LHC is one of the best candidates. In pp collisions a FB asymmetry with respect to a fixed direction, such as the one defined in Eq. 2.5, vanishes due to the symmetry of the initial state. However a *forward-central* charge asymmetry A_C can be defined,

$$A_C = \frac{N(\Delta\ y\ > 0) - N(\Delta\ y\ < 0)}{N(\Delta\ y\ > 0) + N(\Delta\ y\ < 0)} \tag{2.6}$$

with $\Delta\ y\ = \ y_t\ -\ y_{\bar{t}}\ $.

The need for a different observable becomes more clear by considering the asymmetry in the distribution of the angle θ_t between the outgoing top quark and the incoming quark (Fig. 2.3). This angle would be the fundamental observable to measure the charge asymmetry in $t\bar{t}$ production, but the direction of the incoming quark is not directly accessible at experiments. However, at Tevatron $p\bar{p}$ collisions, the quark is provided by the proton with very high probability ($\gtrsim 99\%$ at $\ s = 1\ 96$ TeV); therefore the top quark rapidity y with respect to the proton beam direction can be used in Eq. 2.5 to distinguish between *forward* and *backward* production with high efficiency. At LHC pp collisions the incoming quark has the same probability to be provided by either proton beam; thus the FB asymmetry of Eq. 2.5 vanishes by construction. However, valence quarks are more likely to participate in the interaction than sea quarks and carry, on average, a larger fraction of momentum than sea anti-quarks. Therefore, in the laboratory frame, the direction of flight of the $t\bar{t}$ system is likely to be the same as of the incoming quark. In this scenario, a forward top quark in the $t\bar{t}$ C.M. has on average a larger absolute rapidity $\ y\ $, in the laboratory frame, than the backward antiquark. Thus the asymmetry in Eq. 2.6 allows probing the charge asymmetry in $t\bar{t}$ production at the LHC.

2.3.1 SM charge asymmetry

The dominant contribution to the SM charge asymmetry originates from the QCD $q\bar{q}\quad t\bar{t} + X$ production. Specifically, it originates from the interference between the Born amplitude for $q\bar{q}\quad t\bar{t}$ and its one loop box correction, and from the one between initial and final state radiation (ISR, FSR) in $q\bar{q}\quad t\bar{t}g$ (Fig. 2.4). The terms of the angular differential cross section $d\sigma_{q\bar{q}\ \ t\bar{t}}\ d\cos\theta_t$ (Fig. 2.3) corresponding to these interferences are antisymmetric for the exchange of top and antitop quark momenta [35]. In particular, considering the definitions in Eq. 2.5 and 2.6, Born-box interference generates positive asymmetries, while the ISR-FSR interference generates negative asymmetries. The relative size of the two contributions depends on the transverse momentum of the $t\bar{t}$ system $p_{t\bar{t}}$. For $p_{t\bar{t}} \lesssim 25$ GeV the asymmetry is positive, while for $p_{t\bar{t}} \gtrsim 25$ GeV is negative. The overall effect is a positive asymmetry.

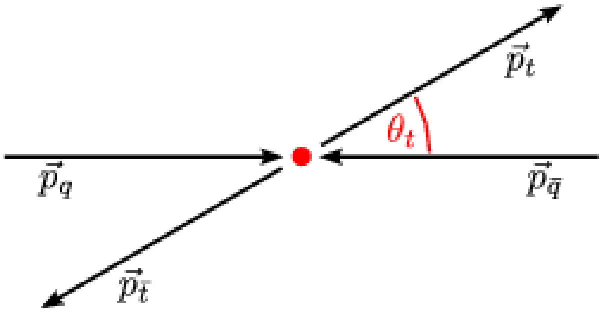
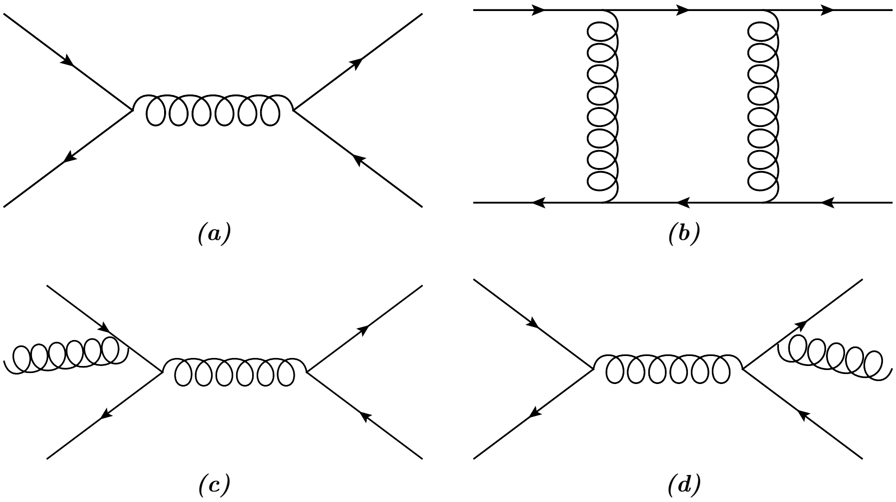


Figure 2.3: A representation of the $q\bar{q}$ $t\bar{t}$ process in the C.M. frame. The top quark t is produced at an angle θ_t with respect to the direction of the incoming quark q. For symmetric production the top quark is produced isotropically, while for a positive (negative) asymmetry, positive (negative) values of $\cos\theta_t$ are favored.



(a) *(b)*

(c) *(d)*

Figure 2.4: Main sources of the QCD charge asymmetry in $t\bar{t}$ production: interference of Born (a) and box (b) diagrams and interference of initial state (c) and final state (d) gluon radiation diagrams

Charge asymmetric contributions also arise from the interference of the QCD $q\bar{q}$ $t\bar{t}$ diagrams in Fig. 2.4 with the QED analogous where a gluon is replaced by a photon. The size of the QED corrections to the QCD asymmetric cross section depends on the relative importance of $u\bar{u}$ and $d\bar{d}$ annihilations, due to the difference in electric charge. At $p\bar{p}$ collisions (Tevatron) the QED contribution is estimated to be $\approx 18\%$, while at LHC pp collisions, with larger relative $d\bar{d}$ importance, it accounts for approximately 13% of the asymmetry [36]. Analogous contributions from weak interactions, where a Z boson is replaced to the photon, constitute only a $\approx 1\%$ correction at Tevatron and are negligible at LHC due to the smallness of the weak coupling.

The asymmetry definition in Eq. 2.5 and 2.6 can be generalized as the ratio between the asymmetric contribution σ_A to the cross section and the symmetric one σ_S:

$$A_{QCD} = \frac{\sigma_A}{\sigma_S} = \frac{\alpha_S^3\sigma_A^{(1)} + \alpha_S^4\sigma_A^{(2)} + \cdots}{\alpha_S^2\sigma_S^{(0)} + \alpha_S^3\sigma_S^{(1)} + \cdots} \tag{2.7}$$

Because the asymmetry vanishes at the tree level in the SM, a consistent fixed-order expansion at LO in perturbation theory involves the numerator at NLO and the denominator at LO:

$$\alpha_S A^{(0)}_{QCD} = \frac{\alpha_S \sigma^{(1)}_A}{\sigma^{(0)}_S} \tag{2.8}$$

Predictions of A_{FB} for Tevatron and A_C for LHC have thus been computed including the $\mathcal{O}(\alpha_s^3)$ QCD contribution and the mixed $\mathcal{O}(\alpha_s^2\alpha)$ QCD-QED and QCD-weak corrections discussed above. The SM forward-backward asymmetry at $p\bar{p}$ collision, evaluated at $\overline{s} = 1\ 96$ TeV, is $A_{FB} = 0\ 088 \pm 0\ 006$, while the SM predictions for the LHC asymmetry are, at $\overline{s} = 7$ TeV and $\overline{s} = 8$ TeV respectively, $A_C = 0\ 0123 \pm 0\ 0005$ and $A_C = 0\ 0111 \pm 0\ 0005$ [37].

The SM prediction for A_C is one order of magnitude smaller than for A_{FB} because of two effects. First, the symmetric $gg \quad t\bar{t} + X$ process is dominant at the LHC, accounting for 80% of the total cross section, as opposed to the 15% at the Tevatron. This process contributes exclusively to the denominator in Eq. 2.5 and 2.6, diluting the asymmetric $q\bar{q} \quad t\bar{t}$ production. In addition, the probability that the anti-quark carries a larger fraction of momentum than the quark in which case the assumption that the $t\bar{t}$ system is boosted in the direction of the incoming quark is voided is not negligible, and leads to a further dilution of the asymmetry. The impact of both dilutions can be reduced by applying a minimum requirement on the z-component of the $t\bar{t}$ system velocity $\beta_{z,t\bar{t}}$ [38]. Due to the average momentum imbalance between quark and antiquark, in $q\bar{q}$ annihilation the $t\bar{t}$ system is often produced with a large longitudinal momentum. Fig. 2.5 shows how the relative importance of the $q\bar{q}$ process increases with $\beta_{z,t\bar{t}}$.

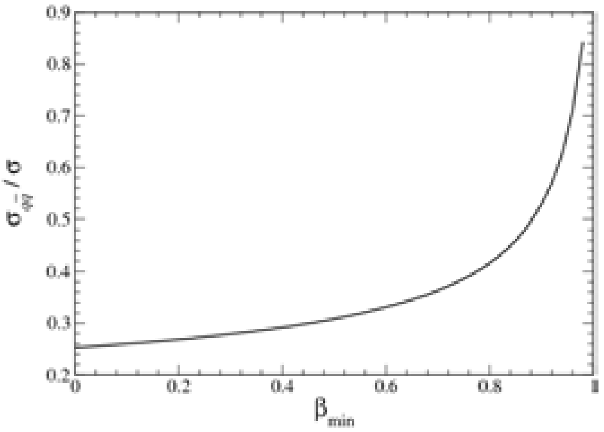


Figure 2.5: Relative fraction of $q\bar{q} \quad t\bar{t}$ events as a function of the minimum $t\bar{t}$ velocity.

2.3.2 bSM charge asymmetry

An appealing possibility is that the discrepancies between the experimental results and the SM predictions for A_{FB} are a signal of new physics in $t\bar{t}$ production. In the last few years various extensions of the SM have been proposed to explain the excess of the measured A_{FB}. The following models have been studied in detail by theorists with a focus on their impact on the charge asymmetry in $t\bar{t}$ production [39].

Color-octet vector G. Exchanged in the s channel via flavor-diagonal couplings, it gives an amplitude that interferes with the SM gluon-exchange diagram. The corresponding contribution to the charge asymmetry in $q\bar{q}$ $t\bar{t}$ is proportional to the product of axial couplings with light and top quarks $g_A^{u,d} g_A^t$. Depending on the relative sign of the couplings and on the mass M_G, the contribution to the asymmetry can be positive or negative.

Neutral Z boson. It contributes to the $u\bar{u}$ $t\bar{t}$ process in a neutralflavor changing t channel. A negative asymmetry is generated in the leading order interference; therefore higher order corrections are required to fit the Tevatron excess. However such tuning of the model is disfavored by measurements of other observables.

Charged W boson. This gauge boson couples to right-handed quarks and contributes in the t channel to the partonic process $d\bar{d}$ $t\bar{t}$. Larger couplings are needed to compensate the lower $d\bar{d}$ luminosity. This field also produces negative contributions to the asymmetry, and it s disfavored by current measurements.

Scalar isodoublet ϕ. Exchanged in the t channel with a flavor-changing coupling, it gives a positive contribution to the asymmetry via the interference with the tree level process. For small masses, an asymmetry consistent with the A_{FB} results can be obtained with relatively small couplings.

Color-triplet scalar . With charge $4e$ 3, it appears only in flavor changing u channel of $u\bar{u}$ $t\bar{t}$. The leading order contribution to the asymmetry is negative, so large couplings are required to have large higher order corrections. Consequently, cancellations of large leading order effects must occur to accomodate current measurements.

Color-sextet scalar Ω. Similar to the above, it also has charge $4e$ 3 and therefore contributes in the u channel diagram. However, with a positive asymmetry, it can easily fit the experimental results.

In all the models, the mass of the new particle and the couplings are free parameters which can be tuned to generate asymmetries of different size. The range of allowed A_C and A_{FB} values for each new physics model are shown in Fig. 2.6. Depending on the features of the model, the asymmetries show different dependence on the $t\bar{t}$ invariant mass. Therefore experimental input on this dependence in data is valuable for model discrimination.

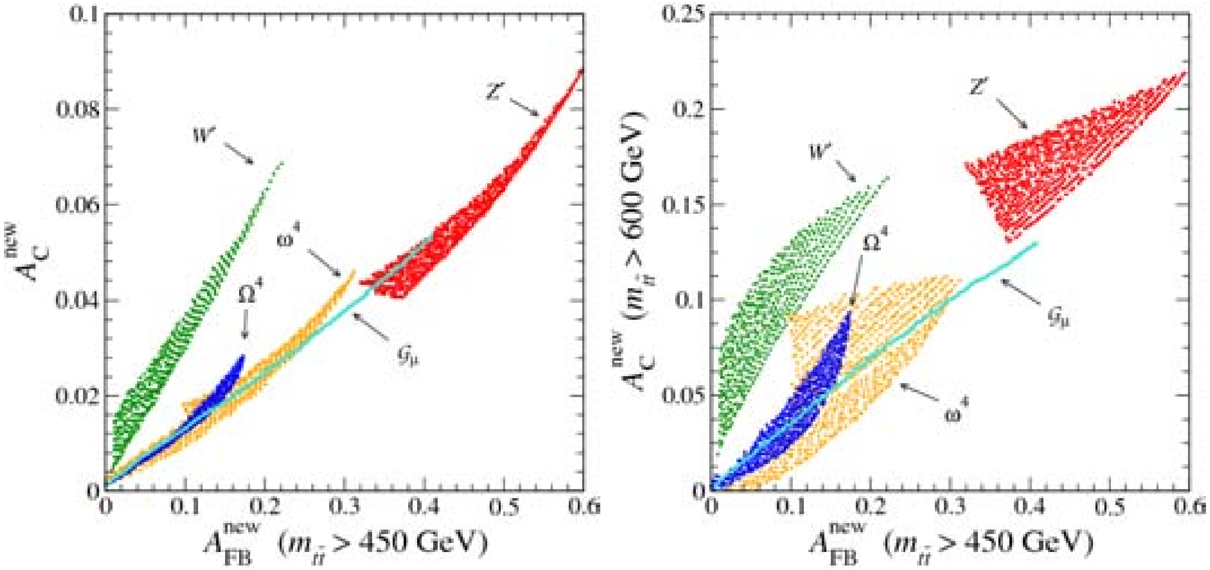


Figure 2.6: Allowed regions for the new physics contributions to the FB asymmetry at Tevatron and the inclusive charge asymmetry at LHC (left) and for $m_{t\bar{t}} > 600$ GeV [40].

Chapter 3

The ATLAS experiment at the LHC

The ATLAS detector is one of the two general-purpose experiments at the Large Hadron Collider (LHC), CERN, Switzerland. It has been built to pursue a broad particle physics program.

The LHC is described in Sec. 3.1, while the design and operation of the ATLAS detector are discussed in Sec. 3.2.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [41] is a circular proton-proton (*pp*) collider. It consists of a 27 km ring of superconducting magnets and accelerating cavities, where two beams of protons circulate in opposite directions. The magnets bend and focus the proton beams into a circular trajectory while radio-frequency cavities boost their energy. The beams undergo several acceleration steps in the pre-accelerators chain and in the LHC, which are both part of the CERN accelerator complex, shown in Fig. 3.1. Protons from the ionization of hydrogen atoms are first accelerated to 50 MeV in a linear collider (LINAC2). Then the energy is increased up to 450 GeV through three stages of synchrotrons: the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). Finally the protons are injected into the LHC.

In the LHC the protons are bunched together into up to 2808 bunches with a minimum spacing of 25 ns. Each bunch consists of about 10^{11} particles. These parameters will allow in the future to reach the design peak luminosity of 10^{34} cm^{-2} s^{-1}. However, most of the data collected so far were taken with a bunch spacing of 50 ns and up to 1308 bunches, yielding a maximum peak luminosity of $3\ 65 \cdot 10^{33}$ cm^{-2} s^{-1}. As shown in Fig. 3.2, the peak luminosity has increased steadily during the data taking.

Due to the high frequency of collisions and charge density of the bunches necessary to achieve such high luminosity, there is a non-zero probability that several events,originating from different *pp* collisions, may occur simultaneously. These pile-up events are categorized as *in-time* or *out-of-time* pile-up. In-time pile-up events are caused by multiple inelastic interactions of protons in the same bunch collision. The out-of-time pile-up occurs when the detector records events originated in different bunch-crossings. The mean number of interactions per bunch collision is shown in Fig. 3.3, with an average of about 9 in 2011 and 20 in 2012.

In 2011 each beam was accelerated to an energy of 3 5 TeV, resulting into a center-of-mass energy $\overline{s} = 7$ TeV, and an integrated luminosity of 5 32 fb^{-1} was delivered. In 2012, the

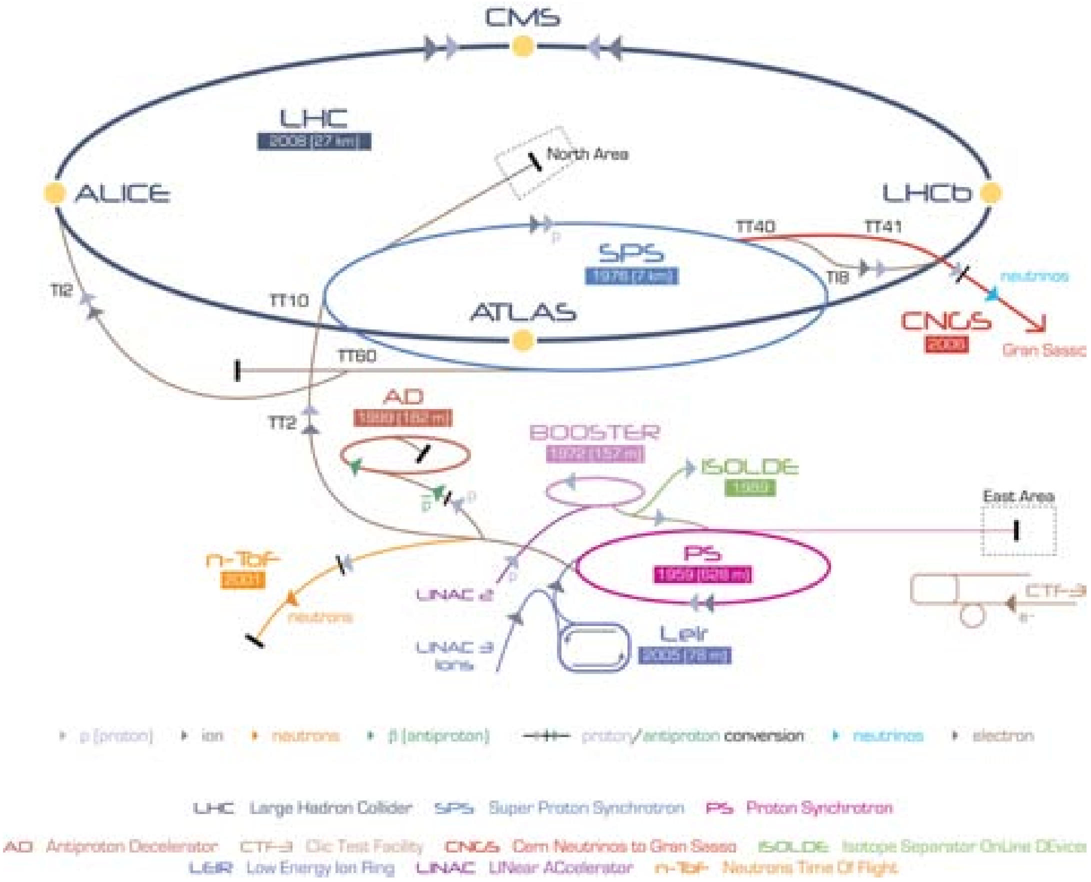


Figure 3.1: The CERN accelerator complex. The four main LHC experiments are shown at the four interaction points.

beam energy was increased to 4 TeV, corresponding to $\sqrt{s} = 8$ TeV. An integrated luminosity of 23 26 fb^{-1} was delivered. The LHC was designed to collide protons up to $\sqrt{s} = 14$ TeV. In 2015 the LHC operation will resume with proton collisions at $\sqrt{s} = 13$ TeV.

The beams are brought to collision in four interaction points along the ring, where four experiments are situated: ATLAS, described in Sec. 3.2 and CMS [42] are multipurpose experiments designed to study a broad range of physics processes; the LHCb [43] experiment is specialized in the detection of b hadrons, while the ALICE [44] collaboration focuses on the study of heavy ion collisions.

3.2 The ATLAS experiment

The ATLAS (A Toroidal LHC ApparatuS) detector is an experiment [45] designed to discover a wide range of new phenomena, if kinematically accessible, and to measure the properties of known particles produced at unprecedented energies and high rates. As shown in Fig. 3.4, the detector is built in several layers of sub-detectors, each devoted to the measurement of different properties for different types of particles. The sub-detectors are grouped into three main systems:

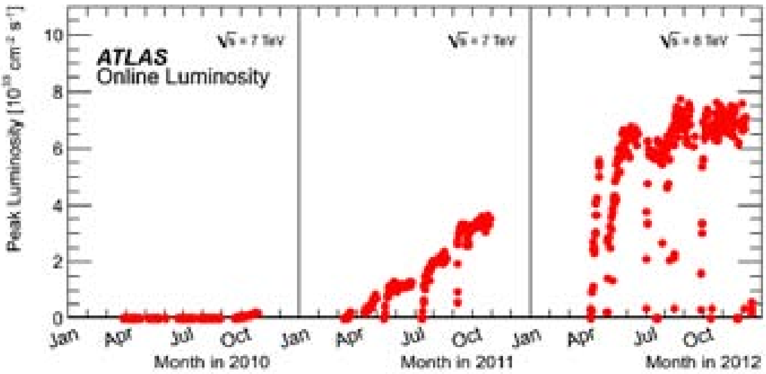


Figure 3.2: The peak instantaneous luminosity delivered to ATLAS per day versus time during the p-p runs of 2010,2011 and 2012.

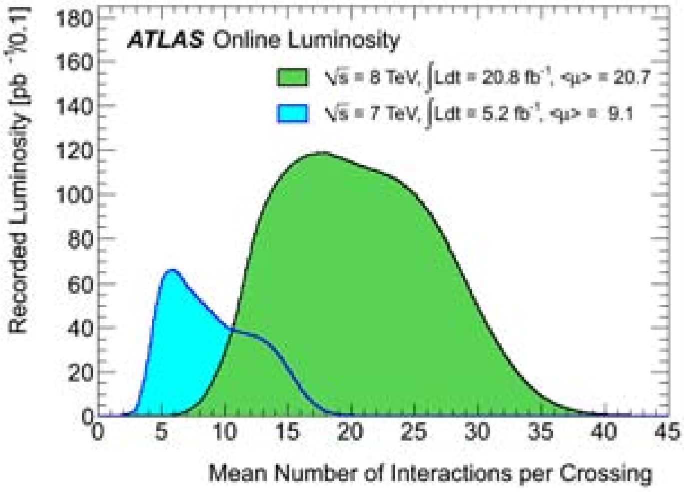


Figure 3.3: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data.

- The Inner Detector, described in Sec. 3.2.2, immersed in a solenoidal magnetic field, constitutes a tracking system used to identify and measure the momenta of charged particles and to identify the interaction vertices and the displaced vertices.
- The Calorimeters are used to identify and measure the energy of neutral and charged particles. They are designed to stop most types of particles, except for muons and neutrinos. Section 3.2.3 illustrates the design and operation of the calorimetry system.
- The Muon system is described in Sec. 3.2.4. Because muons minimally interact with the other parts of the detector and have long lifetimes, they are identified and measured in the outermost detector layer, which consists of muon spectrometers immersed in a toroidal magnetic field.

In $t\bar{t}$ events with +jets final state, electrons and jets are identified by the tracking and calorimetry systems, muons are identified by the tracking and muon systems, and the transverse components of the neutrino momenta are inferred from the imbalance in the total momenta measured in the detector.

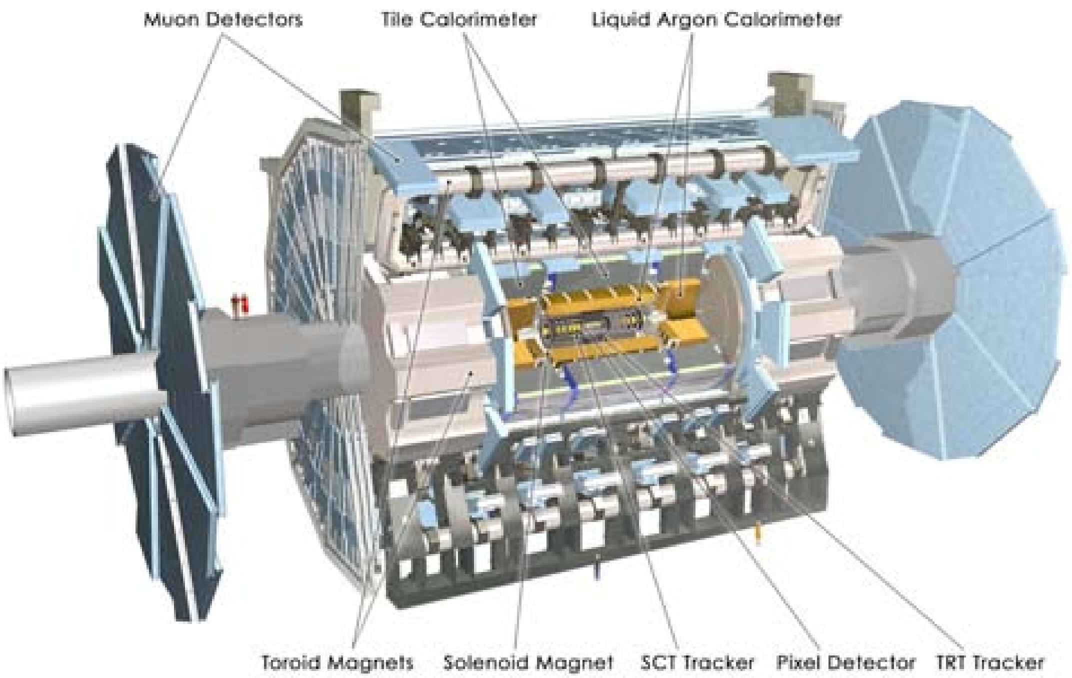


Figure 3.4: Drawing of the ATLAS detector showing the different detectors and the magnet systems.

3.2.1 Coordinate system

The ATLAS coordinate system is a cartesian right-handed coordinate system, with origin at the nominal collision point. The x axis points toward the center of the LHC ring, the y axis points upward, and the z axis is along the beam direction. The azimuthal angle ϕ is the angle with respect to the positive x axis in the x y plane, transverse to the beam direction. The polar angle θ is defined as the angle with the positive z axis. Its function pseudo-rapidity $\eta = -ln[tan(\frac{\theta}{2})]$ is more often used. For massless objects, this quantity is equivalent to the rapidity as $y = \frac{1}{2}ln[\frac{E+p_z}{E-p_z}]$. The distance in the η ϕ space is referred to as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3.2.2 Inner detectors

The ATLAS Inner Detector (ID) consists of three sub-detector systems: the Pixel detector and the SemiConductor Tracker (SCT), which use silicon semiconductor technology, and the Transition Radiation Tracker (TRT), which exploits the transition radiation produced in a gas mixture of Xe, CO_2 and O_2. Figure 3.5 shows a longitudinal and a transverse section of the ID. The whole ID is immersed in a 2 T solenoidal field (see Sec. 3.2.5).

With three concentric cylinders (barrel part), and three end-cap disks perpendicular to the beam axis, the Pixel detector covers a range of $\eta < 2\ 5$. Each of the 1744 sensors consists of a segmented silicon wafer with pixels of minimum area $50 \times 400\ \mu m^2$ and 46080 readout channels. The innermost pixel layer, called b-layer, is as close to the beam line as 50 5 mm and allows for extrapolation of tracks to the vertices with precision below 100 μm. Such precision is crucial for the identification of displaced secondary vertices.

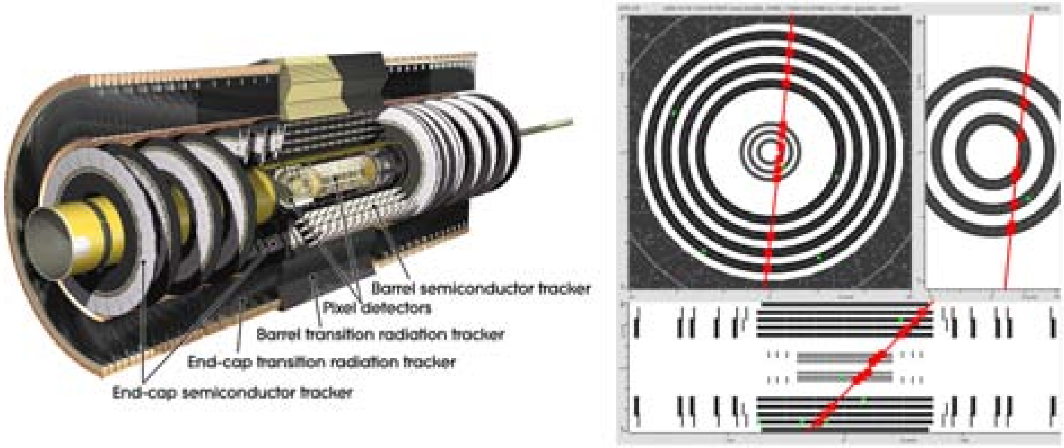


Figure 3.5: Overview of the Inner Detector: the left figure shows a longitudinal section of the Inner Detector with the different sub-detectors. The right figure displays a cosmic ray event as recorded by the pixel detector.

The SCT consists of a four layers barrel and two end-caps, each with nine disks. It covers the range $|\eta| < 2.5$. In the barrel, silicon strips are arranged parallel to the beam line, while in the disks, the strips are oriented radially. Modules are arranged back-to-back with a small stereo angle of 40 mrad to allow for a measurement of the azimuth angle in each layer. A typical track yields three space-points in the Pixel detector and eight in the SCT. Together, the silicon trackers ensure the measurement of the track momenta and the identification of primary and secondary vertices.

The TRT enhances the momentum resolution by providing additional track measurement points. The barrel is made of 73 planes of straw tubes filled with a gas mixture of 70% Xenon, 27% Carbon Dyoxide, 3% Oxygen. The tubes are arranged parallel to the beam axis. In the end-cap, there are 160 straw planes, oriented radially, covering a range of $|\eta| < 2.0$. The TRT also provides stand-alone electron/pion separation, based on the amount of transition radiation, much larger for electrons.

In order to limit the degradation of the energy measurement in the calorimeters, the ID is designed to absorb a minimum amount of energy. The total amount of material of the ID corresponds to roughly 0.5 electromagnetic radiation lengths[1] X_0 in the range $|\eta| < 0.6$. In the regions $0.6 < |\eta| < 1.37$ and $1.52 < |\eta| < 2.5$, the amount of material reaches up to $1.5X_0$. In the barrel-to-end-cap transition regions at $1.37 < |\eta| < 1.52$, the amount of material is even larger, with a significant degradation of the energy measurement. Therefore, electrons in the transition regions were not taken into account in this analysis.

3.2.3 Calorimeters

Figure 3.6 shows an overview of the different electromagnetic and hadronic calorimeters of the ATLAS detector. Both calorimeters are sampling calorimeters consisting of alternating layers of dense absorber material and active material. Only the active material is used for the energy measurement.

---

[1] The radiation length is defined as the typical amount of material traversed by an electron after which it has lost $\frac{1}{e}$ of its original energy by bremsstrahlung.

The hadronic calorimeter in the barrel (Tile) uses steel as absorber and scintillators as active material. The electromagnetic and the forward hadronic calorimeters use Liquid Argon (LAr) technology with different types of absorbers: lead in the ElectroMagnetic Barrel (EMB) and the ElectroMagnetic End-cap Calorimeter (EMEC), copper in the Hadronic End-cap Calorimeter (HEC) and the electromagnetic part of the Forward Calorimeter (FCal), and tungsten in the hadronic part of the FCal. The LAr calorimeters are placed in three cryostats: one for the barrel and one for each end-cap.

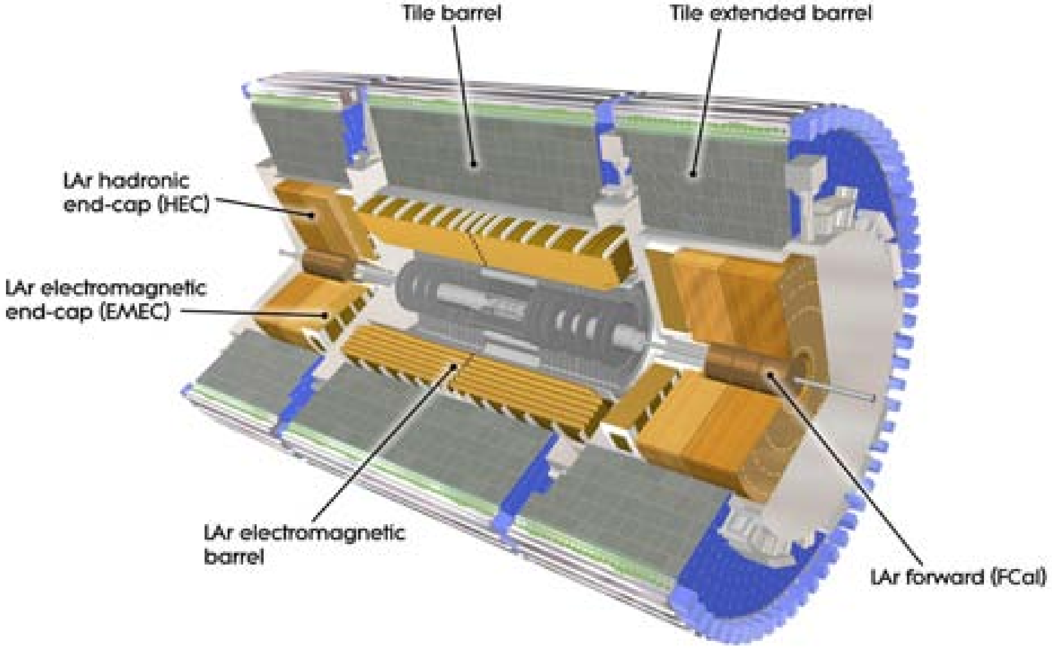


Figure 3.6: Overview of the calorimeter system: the different sub-detectors of the electromagnetic and hadronic calorimeter are shown.

The materials have been chosen to provide fast readout, radiation hardness, and high containment of electromagnetic and hadronic showers[2] to ensure a precise measurement of their energies.

All calorimeters are finely granulated and also segmented longitudinally to allow for a precise determination of the position of the showers and to distinguish different shower types by the use of shower shapes. This is particularly important for the central region, which is devoted to precision measurements of electrons and photons: the EMB (η $< 1\ 475$) is segmented into three longitudinal layers, where the first layer granularity is $\Delta\eta \times \Delta\phi = 0\ 0031 \times 0\ 098$. To ensure continuous coverage in azimuth and to enable fast readout, the lead absorbers are folded into an accordion-shaped structure.

A similar design as for the EMB has been used for the EMEC, which is divided into two wheels covering the ranges $1\ 375 <$ η $< 2\ 5$ and $2\ 5 <$ η $< 3\ 2$. The inner wheel has a coarser granularity in η and ϕ, limiting the region devoted to precision physics to η $< 2\ 5$. A thin

---

[2] A shower is a cascade of secondary particles produced as the result of a high-energy particle interacting with the dense matter of the calorimeter.

LAr layer (pre-sampler) is placed in front of the EMB and the EMEC at $\eta < 1\ 8$ to correct for energy lost in front of the calorimeter.

The Tile calorimeter is located behind the EMB and the EMEC and is divided into three longitudinal layers. It consists of a central barrel ($\eta < 1\ 0$) and an extended-barrel part ($0\ 8 < \eta < 1\ 7$).

The HEC is a traditional LAr sampling calorimeter covering the region ($1\ 5 < \eta < 3\ 2$), which is placed behind the EMEC in the same cryostat. It consists of two independent wheels, each of which is divided longitudinally into two parts.

Altogether, the calorimeters cover the range $\eta < 4\ 9$, thus providing enough hermeticity to ensure a precise measurement of the imbalance of the transverse momentum. Over the whole range in η, the total thickness of the calorimeter system ensures a good containment of electromagnetic and hadronic showers and limits punch-through effects to the muon spectrometer.

3.2.4 Muon detectors

The ATLAS muon system covers the range $\eta < 2\ 7$ and is designed to measure the momenta of muons with energy above ~ 3 GeV. The tracks of the muons are bent by the toroidal magnetic field (see Sec. 3.2.5), whose orientation is such that muon tracks in both barrel and end-caps are mostly orthogonal to the field lines.

The muon system, shown in Fig. 3.7, consists of high-precision tracking chambers as well as trigger systems. In the barrel part, Monitored Drift Tubes (MDTs) are used for tracking and Resistive Plate Chambers (RPCs) for triggering. In the end-caps, tracking information is provided by Cathode Strip Chambers (CSCs), and Thin Gap Chambers (TGCs) are used for triggering. In the barrel as well as in the end-caps, muons typically cross three longitudinal layers of the muon spectrometer. The muon system is divided into eight octants with overlaps in ϕ to avoid gaps in the detector coverage.

The technologies for the tracking systems have been chosen according to level of particle flux. The MDTs in the barrel part follow a robust and reliable detector design. Since each tube contains only one sense wire, the simple geometry allows for the prediction of deformations as well as for a precise reconstruction. As the particle flux increases with η, the CSCs are more suited for the end-cap region: the higher granularity of the multi-wire proportional chambers facilitates to cope with the increasing rates.

The choice of the technologies for the trigger chambers was driven by the requirement for fast and highly efficient trigger capabilities given the different conditions present in the barrel and end-cap regions during data taking. Additionally, an adequate resolution of the transverse momentum of the tracks was required. In the barrel, RPCs provide good spatial and time resolution. However, in the region $1\ 05 < \eta < 2\ 4$ the particle flux is higher and TGCs, with higher granularity, are used. With RPCs and TGCs, a time resolution of $15 - 25$ ns can be achieved, which is sufficient for fast trigger decisions and a good association of tracks to bunch crossings.

3.2.5 Magnet system

The ATLAS magnets system consists of four superconducting magnets: a central solenoid and three toroidal magnets in the barrel and the two end-caps (see Fig. 3.8).

The central solenoid provides an axial field with a strength of 2 T. The solenoid was designed to be particularly lightweight and to minimize the amount of material in front of the calorimeter system.

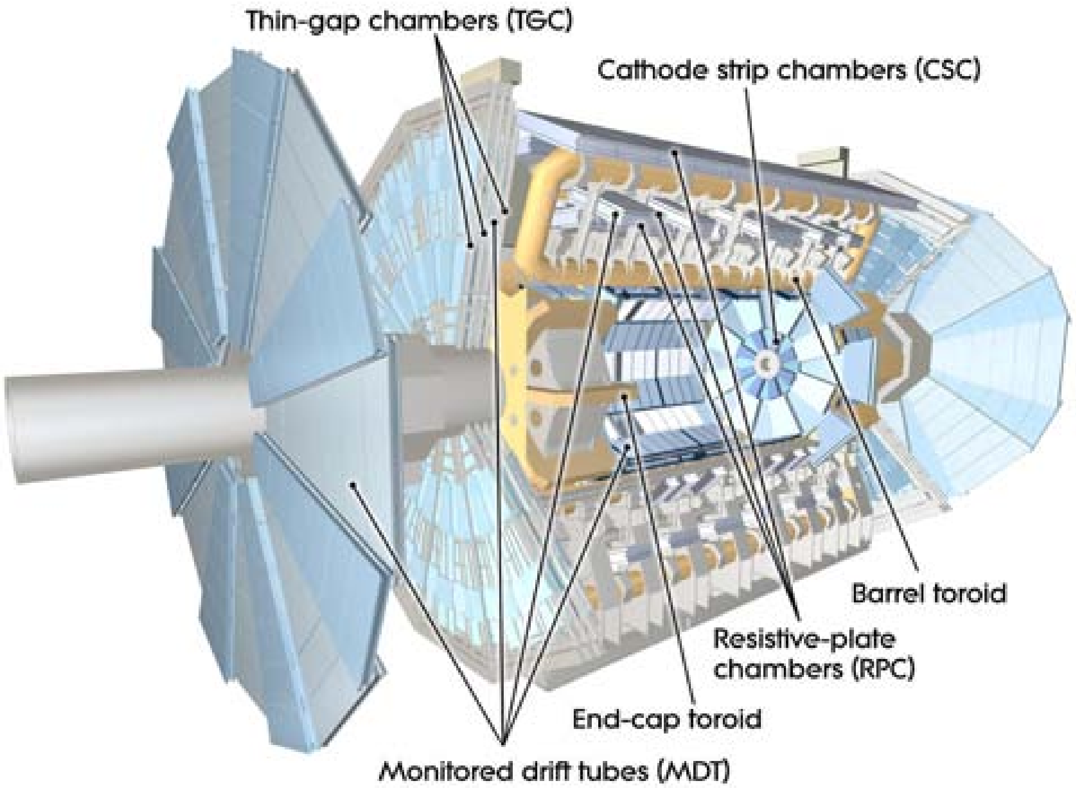


Figure 3.7: Overview of the muon system: the different types of tracking and trigger chambers are shown.

The toroids system provides magnetic fields with a bending power of 1 5 − 5 5 Tm in the barrel and 1 − 7 5 Tm in the end-cap regions. Each system consists of eight coils placed in aluminum housings. The toroidal fields contain non-uniformities which need to be known to high precision to allow for an accurate measurement of muon momenta. Hence, 1800 Hall sensors were installed in the muon spectrometer volume to enable the monitoring of the magnetic field.

3.3 Forward sub-detectors

ATLAS is equipped with sub detectors in the forward regions used for monitoring tasks. The Minimum Bias Trigger Scintillators (MBTS), embedded in the structure of TileCal extended barrel modules, are 32 scintillator paddles assembled in two disks covering the pseudo rapidity region $2\ 09 < \eta < 3\ 84$ and are used to detect minimum bias activity and to indirectly measure the luminosity.

Two other systems are specifically designed to determine the luminosity delivered to ATLAS: LUCID and ALFA. LUCID (LUminosity measurements using Cerenkov Integrating Detector) consists of 32 gas-filled tubes surrounding the beam pipe at 17 m from the interaction point on both sides of ATLAS and measures the luminosity bunch by bunch. ALFA (Absolute Luminosity For ATLAS) is used to calibrate the luminosity measurement with dedicated runs, and consists of 8 scintillating fibers placed at 240 m from the interaction point inside roman pots, above and below the beam pipe.

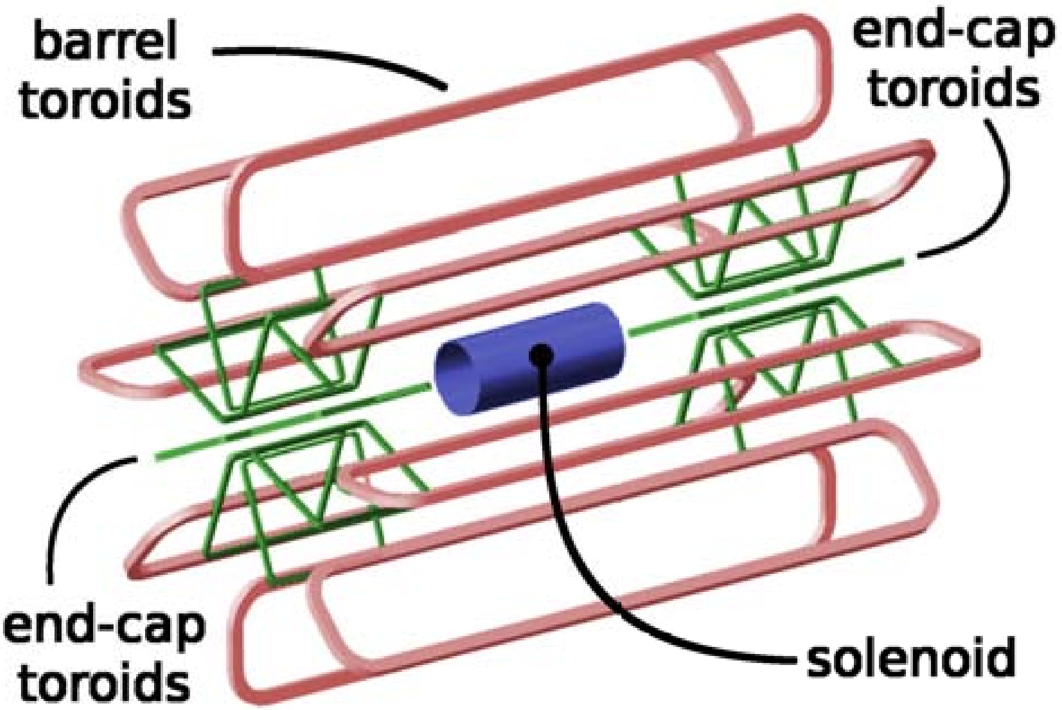


Figure 3.8: Overview of the magnet system: the solenoid and the toroids are shown.

Auxiliary monitoring of the luminosity is provided by the Zero-Degree Calorimeter, whose main purpose is to determine the centrality of heavy-ion collisions. Placed at 140 m from the interaction point on both sides of the beam axis, it is made of quartz rods alternated with tungsten plates.

The Beam Condition Monitor (BCM) consists of two sets of diamond sensors located 184 cm from the interaction point along the beam and at a radial distance of R =5.5 cm. It is devoted to the detection of beam losses and provides fast signals to abort the beam safely, if the loss rate becomes dangerous for the integrity of the ATLAS detector.

3.3.1 Trigger system

With a nominal bunch spacing of 50 ns and ≈ 1 MB event size, ATLAS generates more information than can be recorded to disk. In 20 million interactions per second, approximately one $t\bar{t}$ event is produced. It is therefore crucial to have an efficient trigger system for selecting events of interest.

The ATLAS trigger system, shown schematically in Fig. 3.9, has a three tier structure with increasing levels of information used in reconstruction, and hence refinement of the selection criteria, at each stage.

At the first stage, Level 1 (L1), hardware triggers use coarse calorimeter and muon information for the trigger decision. At this level the event accept rate is reduced to a maximum of 75 kHz with a latency on decision of ≈ 25 μs. In the cases where the trigger is passed, the raw event data are sent to the readout stream for the next trigger level. The L1 trigger defines one or more *regions-of-interest* (RoIs) in η and ϕ where the L1 trigger has identified interesting features.

The Level 2 (L2) trigger is software based. At this level the full detector granularity is used for the trigger decision but only within the RoIs identified at L1. Thus the throughput is reduced to $\approx$ 2kHz within the allowed latency of $\approx$ 40 ms.

The final trigger level is the Event Filter (EF). At this stage the RoIs information is treated using the same algorithms as the off-line reconstruction. The EF reduces the output rate to $\approx$ 200Hz, with a latency of $\approx$ 4 s. Once accepted by an EF trigger, an event is written to mass storage. The combination of the two software steps L2 and EF is referred to as High Level Trigger (HLT).

During 2011 the RoI seeded approach at EF was replaced by a *full-scan* strategy. It was found that sufficient time was available to read the information from the entire ATLAS calorimetry system, which allows to improve the on-line selection efficiency.

The trigger menu

The trigger system is configured via a trigger *menu* which defines trigger *chains*. Each *chain* is defined by a sequence of steps at each trigger level. Thresholds are increasingly tightened at each level in order to maximize selection efficiencies. For each HLT level, a sequence of read-out, reconstruction and selection steps is specified. A trigger is defined as the combination of the L1 seeding item and the two HLT (L2 and EF) chains.

Triggers are categorized based on the object reconstructed (jets, muons, etc..), and the selection requirements, such as energy thresholds, multiplicity or isolation. This categorization, together with other distinctions, depending on the reconstruction algorithm and its configuration, are encoded in the trigger name. The naming convention is as follow:

[LEVEL][N][TYPE(S)][THRESHOLD][ISOLATION][QUALITY]

where the components, from left to right, are: the trigger level used; the multiplicity of the type; the object candidate; the threshold applied to the transverse momentum or energy of the object candidate; the object isolation; additional requirements related to the type of algorithm used.

In order to keep the rate within the available bandwidth ($\approx 200\ Hz$), prescales are applied to the triggers with loose requirements so that only a fraction of the events passing the selection are actually recorded.

Prescales are dynamically adjusted as the instantaneous luminosity decreases at each collision. Data taking time units are defined as *Luminosity Blocks* (LB), corresponding to few minutes of data taking, with associated information on beam conditions, detector performance and trigger configuration. All the LB between the start and the end of a stable beam collision period constitute a *run*. Runs are grouped in *Data Periods*, labelled with capital letters (Period A, Period B, *etc.*), which correspond to the same general detector conditions, machine configuration and trigger menu.

3.4 Data Quality

Not all collision events recorded by ATLAS are used for data analysis. Each sub-detectors maintains a record of its performance across the run. Only the data collected with sub detectors meeting quality requirements are considered for the analysis. Therefore, for each dataset *Good*

Runs Lists (GRL) are compiled recording for each LB which sub detectors satisfied the requirements. For the measurements presented in this dissertation, all ATLAS subsystems are needed as the physics objects used in the analyses are reconstructed using the information from the full detector.

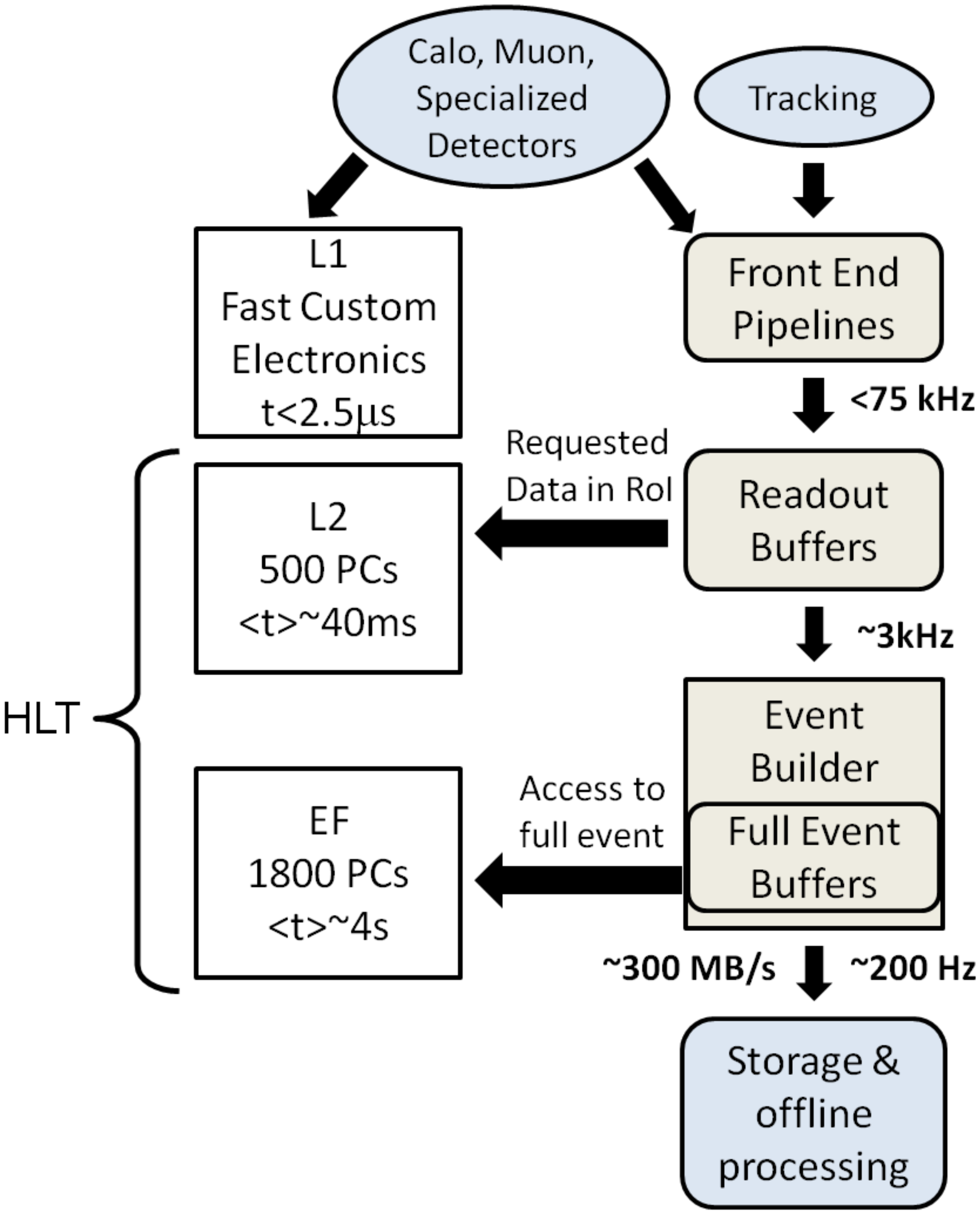


Figure 3.9: Schematic diagram of the ATLAS trigger system [46].

Chapter 4

Event simulation

An accurate simulation of the physics processes and of the interaction of particles with the detector is necessary to model the impact of the analysis procedure on the measured quantities, and to estimate the background composition expected in data. A set of computer programs known as Monte Carlo (MC) event generators simulates the physical processes. Pseudo-random numbers are used to simulate the event-by-event fluctuations intrinsic of quantum processes. MC generators make use of the factorization principle (see Section **??** reference to theory chapter); therefore, the different phases of the *pp* collision are considered independently.

4.1 Simulation of pp collisions

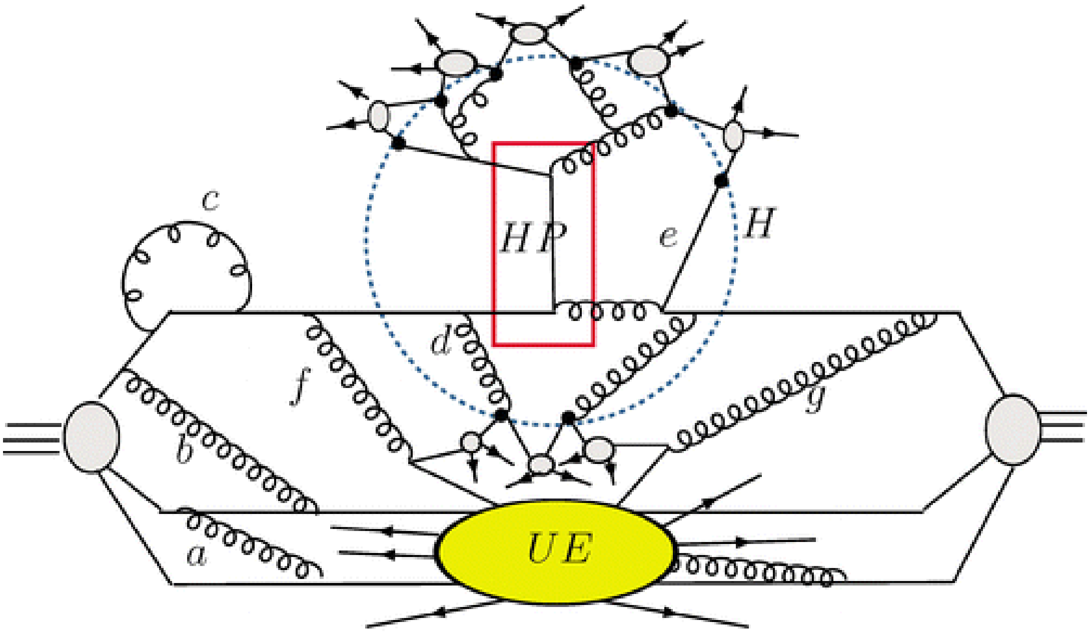
A pp collision event in Monte Carlo simulation is the combination of different sub-processes, illustrated in Figure 4.1. The event is simulated in several steps. Two protons collide and undergo a deep inelastic interaction with a large momentum transfer. The process of interest is generated by the *hard interaction* of two partons within the protons and it is computed at a fixed order (LO or NLO) in perturbation theory. The lower energy interactions between the proton remnants, referred to as *underlying event* (UE), are described with phenomenological models.

Since the partons involved in the hard interaction are color charged, they can radiate gluons. Emission associated with the two colliding partons is referred to as *Initial State Radiation* (ISR), while *Final State Radiation* (FSR) is emitted by the partons produced in the collision. The emitted gluons can emit further gluons or split in quark/anti-quark pairs, leading to the formation of *parton showers*.

The radiation process is effectively described by perturbative QCD until the showers develop into processes at energy below ≈ 1 GeV. At this stage, *hadronization* takes place, where partons are bound into colorless hadrons. Phenomenological models are used to describe the decay of hadrons into the final state particles that interact with the detector.

4.1.1 Hard interaction

The event simulation begins with the collision, with large transfer of momentum, of two partons within the protons. At high energy scale, the partons behave as asymptotically free, and a



Mangano ML, Stelzer TJ. 2005.
Annu. Rev. Nucl. Part. Sci. 55:555–88

Figure 4.1: The general structure of a hard *pp* collision: the hard process (HP) is described as an interaction among fundamental, freely moving constituents; the radiation process continues until the hadronization scale (H) is reached. The underlying event (UE) is described by the soft, multiple interactions among partons not involved in the hard process [47].

perturbative description is applied. The QCD cross section for a generic process $pp \quad X$ is defined (Section **??**-reference to theory chapter) as

$$\sigma_{\mathrm{pp} \quad \mathrm{X}} = \sum_{a,b} \int \mathrm{d}x_a \mathrm{d}x_b \quad \int \quad f_a(x_a) f_b(x_b) \mathrm{d}\sigma_{\mathrm{ab}}(x_a p_a \quad x_b p_b) \tag{4.1}$$

where $x_{a(b)}$ is the fraction of momentum carried by the colliding parton within the proton $p_{a(b)}$, with PDF $f_{a(b)}$.

The fractions of momentum and flavors of the colliding partons are selected by sampling the PDFs of the proton at the energy scale of the process. The cross section for the partonic process $\sigma_{\mathrm{ab}}(x_a p_a \quad x_b p_b)$ is computed explicitly at the lowest relevant order in perturbation theory.

4.1.2 Parton shower

The parton showers represent higher-order corrections to the hard interaction, corresponding to the production of additional partons. Since radiative corrections at a fixed perturbative order are divergent at low energies (*infrared divergence*) or small angles (*collinear divergence*), the explicit calculation is not possible, and an approximation scheme is used where only the dominant contributions are considered.

There are three possible processes for QCD emission (*splitting*): $q \quad gq$, $g \quad gg$ and $g \quad q\bar{q}$. The cross section for each of these processes corresponds to the product of the production cross section for the original parton and a factor accounting for the splitting probability. Hence, for each splitting process i, the $(n+1)$-parton differential cross section is defined, at the LO, by

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_S}{2\pi} \frac{d\theta^2}{\theta^2} dz\ d\phi\ P_i(z\ \ \phi) \tag{4.2}$$

where θ and ϕ are the opening angle and azimuthal angle of the splitting, and P_i is the splitting function, which describes the distribution of the fraction z of energy of the original parton, assigned to the new parton. The simulation algorithm develops the shower by applying Eq. 4.2 iteratively, for each parton involved in the hard interaction.

In order to define the starting and final stage of the evolution of the parton shower, the *virtuality* q^2 of the parton undergoing the splitting is defined as the invariant mass of the two partons produced. The initial virtuality is required to be smaller than the momentum transfer of the hard process, and the shower is terminated when the virtuality has fallen below the hadronization scale ($q^2 = Q_0^2 \quad 1$ GeV^2). The parton shower takes into account virtual emissions that are reabsorbed in quantum loops as a probability of *not* splitting in a given virtuality range $[q_1^2\ \ q_2^2]$. Such probability is referred to as *Sudakov form factor*

$$\Delta_i(q_1^2\ \ q_2^2) = \exp\left[- \int_{q_2^2}^{q_1^2} \frac{dq^2}{q^2} \frac{\alpha_S}{2\pi} \int_{\frac{Q_0^2}{q^2}}^{1-\frac{Q_0^2}{q^2}} dz \int_0^{2\pi} d\phi P_i(z\ \ \phi) \right] \tag{4.3}$$

The evolution of the parton shower is therefore governed by the Sudakov factor. Given the initial scale Q^2, the MC generator solves the equation $\Delta_i(Q^2\ \ q_1^2) = R_1$, where R_1 is a random number uniform on the interval [0,1], for the virtuality q_1^2 of the first splitting. If the condition $q_1^2 < Q_0^2$ is met, the shower development is terminated and hadronization takes place. Otherwise, the procedure is repeated for each new parton produced by the splitting, taking q_1^2 as initial scale. For each splitting the variables z and ϕ are generated according to the distribution defined by the splitting function.

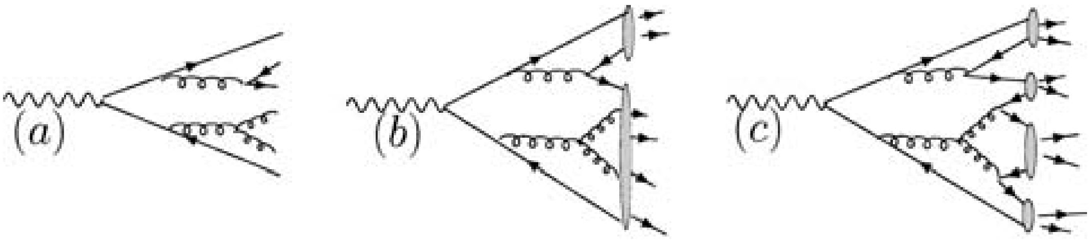
ISR and FSR showers

The description above applies to the development of showers associated with partons produced in the hard interaction, starting at a high energy scale Q^2 and progressively reaching the hadronization scale. This process is typical of FSR parton showers that are generated from outgoing partons of the hard interaction.

In the case of ISR parton showers, the radiation is emitted by the colliding partons, and there is an important difference in the shower evolution, as the final energy of the showering is set by the hard interaction energy scale. MC generators implement a mechanism of *backward evolution* that first sets the correct parton momentum fractions for the hard scatter, and then develops the showers backward, with the intermediate partons gaining energy at each emission. The Sudakov form factors are then slightly different from Equation 4.3, being rescaled by a factor that takes into account the PDFs of the parton before and after splitting. This procedure ensures a highly efficient simulation, as only ISR showers compatible with the energy scale of the hard interaction are generated.

4.1.3 Hadronization

When the shower evolution brings the parton virtuality q^2 below the hadronization scale Q_0^2 1 GeV^2, the dynamics of the parton enters a non-perturbative phase, which leads to the formation of the final-state colorless hadrons. The hadronization process cannot therefore be described with perturbative QCD, and MC generators rely on phenomenological models.



Mangano ML, Stelzer TJ. 2005.
Annu. Rev. Nucl. Part. Sci. 55:555–88

Figure 4.2: Possible radiation pattern from a $q\bar{q}$ pair (a), and illustration of string fragmentation (b) and cluster hadronization (c) [47].

Two phenomenological hadronization models are typically used to bound partons into hadrons. In the *Lund string model*, the confinement between partons induced by the color force is represented by a gluonic string. In the case of a quark-antiquark pair, as the color charges move apart, the string is stretched, and its potential energy grows. When the energy becomes of the order of hadron masses, it becomes energetically favorable for the string to break and create a new quark-antiquark pair. The two segments of string will stretch and break again, until all the energy has been converted into quark-antiquark pairs connected by short strings. In the case of more complicated color structures, multiple strings are considered with as many endpoints as the color charges available.

The other hadronization scheme is the *cluster model*, where final state gluons are forced to split into quark-antiquark pairs, and partons are grouped to form colorless clusters. At the hadronization scale, most clusters have masses below 3 GeV, and their decay into hadrons is simulated with three-body models with intermediate resonances (*quasi-two-body decay*). Clusters with higher masses are decomposed using a string-like mechanism.

4.1.4 Underlying event

In *pp* collision events containing a hard interaction, an additional hadron production mechanism arises from the softer interaction of spectator partons. Because of the low energy scale of these processes, phenomenological models, whose parameters are tuned based on experimental data, are used. The dominant subprocess of the underlying event is gluon-gluon scattering, with a cross section larger than the total *pp* scattering cross section, indicating that multiple gluon scatterings per proton collision are likely. For this reason the generic soft scattering of partons is referred to as *multiple parton interactions* (MPI) and is modeled in MC generators as the production of back-to-back jet pairs with little total transverse momentum. The color connection with the beam remnants that are not interacting is also simulated with phenomenological models.

4.2 Generators

Generators are classified as either *multi-purpose* generators, capable of performing the full simulation chain described above, or as *specialized* generators, optimized for an accurate simulation of specific aspects. The following sections summarize the characteristics of the MC generators used in this work.

PYTHIA

`PYTHIA` [48, 49] is a multi-purpose MC generator using LO calculations for 2 n ($n \leq 3$) processes and PS with emissions ordered in transverse momentum. The Lund string model is used for hadronization, and UE simulation is included.

HERWIG

`HERWIG` [50] is a multi-purpose MC generator using LO calculations for 2 2 processes and PS with emissions ordered in opening angle. The cluster model is used for hadronization and for the UE description, `HERWIG` is typically interfaced with the standalone software `JIMMY` [51] that simulates UE as MPI.

ACERMC

`ACERMC` [52] is a MC generator computing LO cross sections and typically interfaced either with `PYTHIA` or `HERWIG` for the modeling of PS, hadronization and UE.

ALPGEN

`ALPGEN` [53] is a MC generator specialized for LO calculations of 2 n ($n \leq 9$) processes. It is interfaced with either `PYTHIA` or `HERWIG` for PS development and hadronization. UE is simulated through `PYTHIA`.

MC@NLO

`MC@NLO` [54] is a MC generator using NLO calculations. The full NLO correction provides precise cross section estimates, but higher-multiplicity parton emissions are simulated via `HERWIG` PS with a poor description of hard emissions. Hadronization and UE are simulated through `HERWIG` and `JIMMY`.

POWHEG

`POWHEG` [55] is a MC generator computing NLO cross sections and typically interfaced either with `PYTHIA` or `HERWIG` for the modeling of PS, hadronization and UE.

4.3 ATLAS detector simulation

The MC generators create a list of four-vectors of all stable particles produced in the *pp* collision. The simulation of the experimental setup [56] then propagates all final state particles through the ATLAS detector and converts the energy depositions into electronic signals simulating the readout system. The interaction of particles with the detector, taking into account its materials, geometry and readout system, is modeled using the `GEANT4` [57] package.

The `GEANT4` parameters are tuned using test-beam data and the accuracy of the detector simulation in *pp* collision data. The detector simulation is based on the information from the *geometry database*, which contains the description of the detector volumes in terms of dimensions, geometry, position and material composition, while the *conditions database* provides the information on the detector real-time conditions like dead channels, misalignments, temperatures. Since conditions vary from run to run, it is important that the detector simulation reproduces as close as possible the real status of ATLAS during a particular data period. To ensure this, simulation samples are reprocessed for each *data releases*.

4.4 Monte Carlo simulation weighting and corrections

In order to compare with the distributions observed in data, the simulated samples are normalized to the number of events expected based on the theoretical cross section and the integrated luminosity. An event weight w is applied, defined as:

$$w = \frac{\sigma \times k}{N} L \tag{4.4}$$

where σ is the process theoretical cross section, N is the number of simulated events, L the integrated luminosity and k a correction to the LO cross section to reproduce a higher-order (e.g. NLO) calculation.

In addition, a weight to account for pile-up effect is applied so that the number of interactions per bunch crossing $< \mu >$ matches real data-taking conditions.

In order to ensure an accurate modeling of the detector effects, data-driven corrections are applied to the simulated samples. The reconstruction efficiency, energy scale and resolution of the different physics objects are calibrated based on precise measurements.

Chapter 5

Physics Objects

The ATLAS detector records events as raw data, which correspond to bits of electric signal collected when particles interact with the detectors. The goal of object identification is to reconstruct the particle four-momenta by combining the information from the different sub-detectors. This task is performed by algorithms optimized for the reconstruction of electrons, muons, jets and the energy imbalance left by the passage of neutrinos.

5.1 Tracks

In the solenoidal magnetic field of the inner detectors, a charged particle moves along a helicoidal trajectory with a curvature proportional to its momentum. Tracks are the reconstruction of these trajectories from the electric signals induced in the detectors by ionization. Therefore, tracks are used to identify charged particles and measure their momenta. In addition, the extrapolation of the trajectories allows the identification of the interaction vertices.

The parameters describing a track are shown in Fig. 5.1: θ, the angle with respect to the Z axis in the RZ plane measured from the perigee[1]; ϕ_0, the angle with respect to the X axis in the XY plane measured from the perigee; d_0, the impact parameter, or perigee with respect to the Z axis in the XY plane; z_0, Z component of the perigee.

In order to reconstruct the track, the first step is to retrieve the information from the ID hits, which are converted into three-dimensional space points. Then, the *inside-out* algorithm [58] iteratively builds a track by combining space points one by one, starting from a seed of three aligned hits in the pixel detector or in the SCT. For each new point, a Kalman filter algorithm [59] checks the compatibility between the track and the new point. A cleaning procedure prunes the track collection, removing tracks sharing hits with other tracks, and tracks reconstructed from noise hits. Finally, the track quality is improved by taking into account the signal from the TRT and the effects from the interaction of the charged particle with the detector material.

For signals in the TRT that are not associated to any track candidate by the inside-out reconstruction, a second algorithm, referred to as *outside-in*, is applied in order to reconstruct tracks from secondary charged particles. The algorithm uses as seeds hits in the TRT and extrapolates back to the SCT and pixel detector.

---

[1] The perigee is the point of the track closest to the origin.

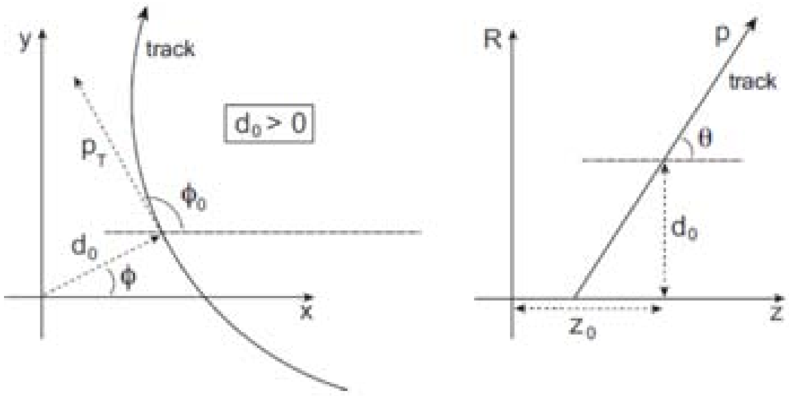


Figure 5.1: Track parameters in the XY (left) and RZ (right) planes where the origin is the beam spot, i.e. the region in the where the protons collide.

5.2 Primary Vertices

The reconstruction of the interaction points, referred to as *Primary Vertices* (PVs), is essential to identify which one corresponds to the hard scattering process and reconstruct the physics objects accordingly. The reconstruction of primary vertices is divided in two steps: first vertex candidates are identified by association of reconstructed tracks; then, the actual vertex position is precisely determined. The two steps are repeated iteratively [60].

The initial position for a PV is the maximum in the distribution of the z_0 parameter of reconstructed tracks. The PV is then estimated with an iterative η^2 fitting algorithm that down-weighs the contribution of incompatible tracks. Tracks incompatible with the PV by more than 7σ are used to reconstruct a new vertex. The procedure is repeated until there are no tracks left to create a new vertex. Among the PV candidates found with this procedure, the ones with less than two tracks associated are ignored.

The hard scatter PV is assumed to be the one with the highest sum of squared transverse momenta of the tracks. The rest of the PVs are considered pile-up interactions.

Vertices incompatible with the beam collision region are considered secondary vertices. Also referred to as displaced vertices, they typically originate from decays of long-lived particles. The reconstruction of secondary vertices is useful to identify B-hadrons, as it will be described in Section 5.5.1.

5.3 Electrons

An electron candidate object [61] is selected by searching for a narrow, localized cluster of energy deposits in the EM calorimeter, with at least one ID track associated to it. A *sliding-window* clustering algorithm is used to identify electron clusters. The algorithm performs a scan of the calorimeter, searching for local maxima of energy within a window of dimensions $\Delta\eta\times\Delta\phi = 0\ 075\times 0\ 125$. The scan is performed in the range $\eta_{\mathrm{cluster}} < 2\ 47$, which corresponds to the ID coverage for reconstructing tracks. Tracks within a window $\Delta\eta \times \Delta\phi = 0\ 05 \times 0\ 10$ are associated with the cluster.

The electron four-momentum is built from the cluster energy and the direction of the associated ID track. When multiple tracks are associated with the cluster, the closest to the cluster[2] is considered. The track momentum is required to be compatible with the cluster energy, which is calibrated to the electromagnetic scale. The calibration is derived from simulation, test-beam studies, and Z ee data events [62].

In order to suppress the mis-identification of other particles as electrons, selection criteria based on cluster shape, track-cluster matching, track quality, and isolation are applied. The shower development is narrower for electrons than for hadrons, and the hadronic leakage[3] is smaller. Track quality requirements reduce the impact of accidental track association with photons, energetic π^0 or η mesons with electromagnetic decays reconstructed as a single energy cluster. Finally, the isolation helps rejecting electrons from semi-leptonic decays of heavy hadrons. Six electron definitions are used in ATLAS in order to discriminate real electrons from misidentified ones. Each definition applies additional requirements with respect to the one preceding.

`Loose` electrons are defined by applying requirements of low hadronic leakage and on the shower shape. The identification efficiency is above 95%, with a jet contamination rate of about 1 500.

`Loose++` electrons are `loose` electrons whose track has at least one hit in the pixel detector and at least 7 hits in the combined silicon detectors. The $\Delta\eta_{\mathrm{firstEM}}$ between the track extrapolated to the first EM layer and the matched cluster must be lower than 0.015. The identification efficiency is similar as for `loose` one with a ten times higher jet rejection.

`Medium` electrons are `loose++` electrons with additional requirements on shower shape track quality: d_0 <5 mm and Δ η_{firstEM} $<$ 0 01. The efficiency is about 88% and the rejection is better than for `loose++` electrons.

`Medium++` electrons are `medium` electrons whose track has at least one hit in the first pixel detector layer and a high fraction of high-threshold TRT hits. In addition, Δ η_{firstEM} is required to be smaller than 0 005, and more stringent requirements are applied to the shower shape of clusters at η $<$ 2 01. The efficiency is about 85% and rejection is a bit less 50×10^3.

`Tight` electrons are `medium++` electrons with additional requirements on the distance between the track and the matched cluster ($\Delta\phi$ $<$ 0 02, $\Delta\eta$ $<$ 0 005). An higher fraction of high-threshold TRT hits is required, as well as an impact parameter d_0 <1 mm. The efficiency is about 75% and the rejection is slightly better than for `medium++`.

`Tight++` electrons are `tight` electrons with asymmetric $\Delta\phi$ requirements, yielding better efficiency (about 80%) and rejection (about 50×10^3) with respect to `tight`. This definition is used to identify electrons in this work.

The identification efficiencies depend on the electron p_{T} and pseudo rapidity, while they are not strongly affected by pileup. As shown in Fig. 5.2, the modeling in simulation differs slightly from what observed in data, therefore a calibration is applied in MC samples.

The `medium++` selection, where isolation and track-cluster matching requirements are relaxed, is also used, with the purpose of studying and modeling the contamination from QCD multijet production among the tight candidates. The procedure to estimate this background contribution is detailed in Section 6.3.3.

In order to ensure high purity, in addition to the identification requirements, the E_{T} of the electron used in the analysis is required to be larger than 25 GeV, and electrons in the transition region 1 37 $<$ η_{cluster} $<$ 1 52 are not considered. Electrons are also required to be isolated both

---

[2] The position of a cluster is computed as a weighted average of the η–ϕ positions of the calorimeters cells in the cluster, based on the absolute value cell's energy.

[3] The hadronic leakage is the fraction of energy reconstructed in the first layer of the hadronic calorimeter

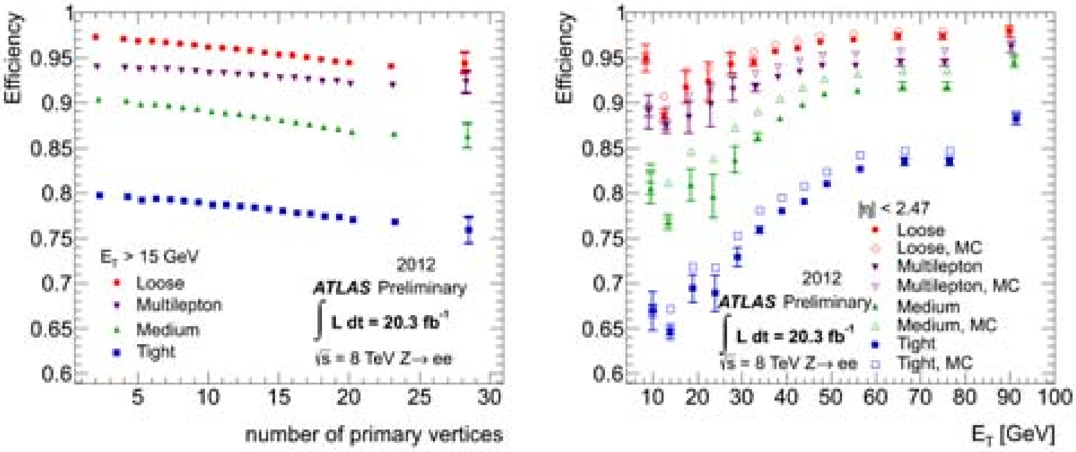


Figure 5.2: Electron identification efficiencies measured from $Z \quad ee$ events in data and simulation as a function of the number of PVs (left) and E_{T} (right).

at the calorimeter level, considering the energy within $\Delta R(\ ) < 0\ 2$, and at the track level, considering the scalar sum of p_{T} of tracks within $\Delta R(\ ) < 0\ 3$

5.4 Muons

Muons are reconstructed in the region $\eta \ < 2\ 5$ using tracks measured in the Muon Spectrometer (MS) and in the ID. The information from the two systems is used by matching the tracks with the *MuId* algorithm [63] to build *combined* muon candidates. The algorithm searches for track segments in the RPC and TGC in $\Delta\eta \times \Delta\phi = 0\ 4 \times 0\ 4$ regions where the trigger fired. A single MS track is built with a least-square fitting method, and the trajectory is extrapolated back to the interaction point, taking into account the energy losses in the calorimeter material. The MS track is combined with the ID track that provides the best match, based on a η^2 test. If no track is found, no combined muon candidate is built. The momentum of the muon candidate is computed as a weighted average of ID and MS measurements and calibrated using $Z \quad \mu\mu$ events. Muons are reconstructed with ≈ 98% efficiency, as shown in Fig. 5.3.

The p_{T} of the muon used in the analysis is required to be larger than 25 GeV and the muon is required to be isolated: muons overlapping with reconstructed jets (see Section 5.5) within a 0.4 cone in ΔR are rejected. In addition, a *mini–isolation* [64] requirement is applied to reject non isolated muons and reduce sensitivity to the high pile-up conditions of $\overline{s} = 8$ TeV collision events. The mini isolation is defined as

$$I^{\mu}_{mini} = \sum_{tracks} p_T^{tracks} \quad p_T^{\mu} \tag{5.1}$$

where p_T^{μ} is the transverse momentum of the reconstructed muon and the sum runs over all tracks found within a p_{T} dependent cone radius:

$$\Delta R(\mu\ \ track) < \frac{10\ \mathrm{GeV}}{p_T^{\mu}} \tag{5.2}$$

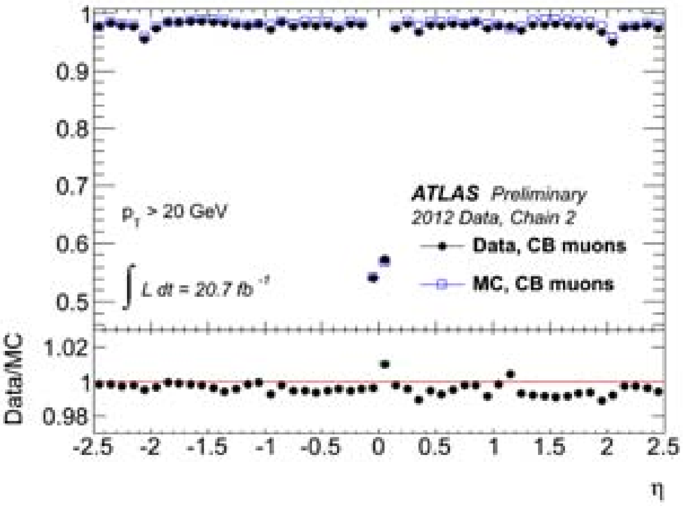


Figure 5.3: `Combined` muon reconstruction efficiencies using the `Muid` algorithm measured from Z $\mu\mu$ events in data and simulation as a function of η [64].

The mini isolation requirement is $I^{\mu}_{mini} < 0$ 05, yielding a 97% efficiency for identifying hard scatter muons.

5.5 Jets

Jets are collimated showers of particles from the hadronization of quarks or gluons produced in the collision. The resulting stable particles leave tracks in the ID, if charged, and clusters of energy deposits in the calorimeters. Reconstructing the total energy and direction of a cluster allows an estimation of the four momentum of the jet originating the shower.

Neighboring calorimeter cells are grouped into topological clusters (*topoclusters*) based on the significance of the energy deposit in the calorimeter cells E_{cell} with respect to their noise level σ. The noise level σ is defined as the RMS of the energy distribution measured in events triggered at random bunch crossings. Cells with E_{cell} $\sigma > 4$ are considered as seeds, and all the neighboring cells with E_{cell} $\sigma > 2$ are included in the topocluster. Cells adjacent to the selected ones are also included without any energy requirement. Topoclusters are calibrated at the electromagnetic (EM) scale as the calorimeter signals originates from the electromagnetic interaction of particles with the detector material. For $\overline{s} = 8$ TeV collision events an additional correction, referred to as Local Cluster Weighting (LCW) [65], is applied. The LCW calibration scheme classifies the clusters as *mainly electromagnetic* or *mainly hadronic*, based on their shape, and applies dedicated corrections, derived from simulation, accounting for non compensation, out of cluster energy and dead material effects.

Jets are then reconstructed using the *anti-k_t* algorithm [66], which combines topoclusters iteratively, based on a distance parameter criterium. The distance parameter is defined as:

$$d_{ij} = min(\frac{1}{p^2_{T_i}}\ \frac{1}{p^2_{T_j}})\frac{\Delta R^2_{ij}}{R^2} \tag{5.3}$$

where p_{Ti} is the transverse momentum of topocluster i, $\Delta R_{ij}=\sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2}$ the distance between topoclusters i and j, R a parameter of the algorithm that approximately controls the size of the jet and is chosen to be 0.4. The algorithm computes d_{ij}, the distance between two topocluster inputs i and j, and $d_{iB} = 1\ p^2_{T_i}$, the distance between the input i and the beam axis, for the whole list of topoclusters found in the event. If d_{iB} is the smallest distance, the topocluster i is considered a jet and removed from the list. Then the algorithm repeats the procedure with the remaining input objects. Otherwise, the i and j topoclusters corresponding to the smallest distance d_{ij} are combined, and the list is updated for a new iteration. The procedure is repeated until the list is empty.

The jet energy is computed as the sum of the energy of the cells forming the jet. A η dependent correction is applied to subtract the additional energy from in-time and out-of-time pileup, based on the number of primary vertices in the event and the average number of pp interactions. The corrected energy corresponds to the energy of the jet as reconstructed in a scenario without pileup collisions and, therefore, a single primary vertex in the event. Finally, a correction factor to extrapolate the particle level energy, referred to as *jet energy response*, is applied. The jet energy response is derived from the simulation of di jet events from single pp interactions. The jets used for the calibration are isolated jets[4] reconstructed in the calorimeter and matched to a truth-level jet within $\Delta R < 0\ 3$. The jet energy response is the ratio between the energy measured in the reconstructed jets (E^j_{LC}) and the truth jet energy (E^j_{truth}). The calibration is derived as a function of the jet energy E^j_{LC} and pseudo rapidity η_{det}, as shown in Fig. 5.4.

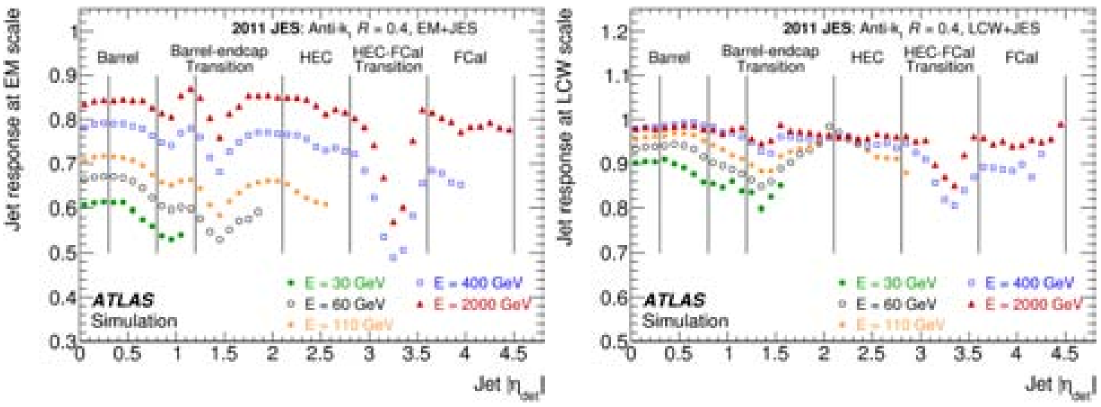


Figure 5.4: Average jet response for topoclusters at EM scale (left) and at LCW scale (right). The response is shown separately for various truth-jet energies as function of the jet pseudorapidity $_{\mathrm{det}}$. Also indicated are the different calorimeter regions.

Only jets with $p_{\mathrm{T}} > 25$ GeV and $\eta\ < 2\ 5$ are considered in the analysis. In addition, jets found within $\Delta R = 0\ 3$ of a reconstructed electron are not considered, in order to avoid double counting the energy deposit of the electron shower. In order to ensure that the selected jets originate from the hard scattering process, the information of tracks associated with the jets is exploited. It is required that the total p_{T} of tracks originating from the primary vertex for the signal process be at least 75% of the total p_{T} of tracks with $p_{\mathrm{T}} > 1$ GeV associated with the jet. This requirement, referred to as *Jet Vertex Fraction* (JVF), is slightly different in the analysis of the dataset at $\ \overline{s} = 8$ TeV, where only jets with $p_{\mathrm{T}} < 50$ GeV are required to have a $JVF > 50\%$.

---

[4] A jet is considered isolated when no other jet with $p_{\mathrm{T}} > 7$ GeV is found within a ΔR cone of $2.5R$, where $R = 0.4$ is the jet radius.

5.5.1 b tagging

When a bottom quark is produced in an event, it hadronizes into a B hadron, which has a lifetime of the order of 10^{-12} s and hence can travel about 3 mm before decaying. This long decay length leads to displaced secondary vertices and large impact parameters for the decay products. Therefore the measurement and identification of these objects in the ID allows to identify jets corresponding to bottom quarks.

The identification of b–jets (b-tagging) is performed by the combination of three algorithms, referred to as *JetFitter*, *IP3D*, and *SV1*, which exploit different properties to determine a probability (b–tag weight) for the jet to originate from the fragmentation of a b–quark. The IP3D algorithm estimates the b–tag weight by defining a likelihood based on the impact parameters, along the beam axis and in the transverse plane, of the tracks associated with the jet. The SV1 algorithm reconstructs the secondary vertex and computes a likelihood ratio to discriminate between b–jets and light jets[5] using the number of track pairs in the secondary vertex, their total invariant mass and the fraction of momentum corresponding to the secondary vertex. The JetFitter algorithm performs a reconstruction of the full decay chain of B and C hadrons by using a Kalman filter to determine a common path between the primary vertex and the vertices from b and c hadrons inside the jet. The significance of the flight length is used as discriminant, in addition to the observables used by the SV1 algorithm. A neural network, referred to as *MV1* algorithm, is used to combine the output of the three algorithms into a single b–tag weight. The working point used to tag b–jets corresponds to a 70% efficiency, as measured in $t\bar{t}$ simulated events and shown in Fig. 5.5, while the mis–identification probability is $\sim 0.8\%$ for light jets, and the one from c–jets $\sim 20\%$.

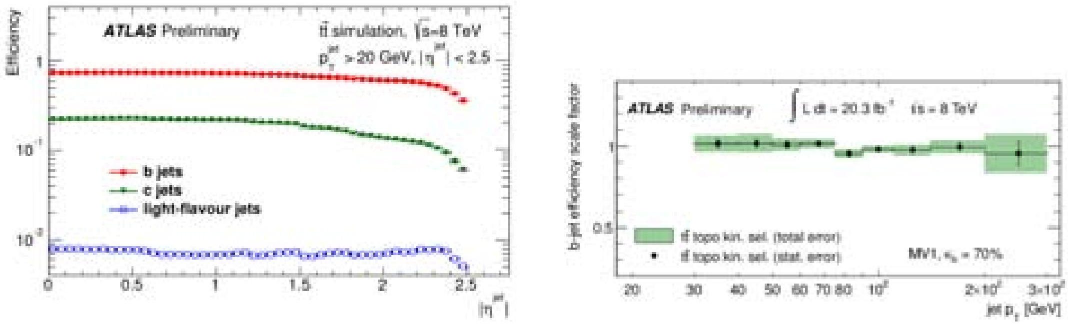


Figure 5.5: Efficiency of the MV1 tagger to select b–, c–, and light–flavor jets, as a function of jet $|\eta|$ (right) and the p_T–dependent data–to–simulation scale factor for data at $\sqrt{s} = 8$ TeV (left).

The b–tagging efficiencies are measured for b–, c–, and light–flavor jets [67–69] and the simulation is calibrated with the appropriate scale factors, p_T and η–dependent (Fig. 5.5). For b–jets, the efficiency collision data at $\sqrt{s} = 7$ TeV is derived from di–jet samples with muons in the final states. The muon momentum component transverse to the axis of the system formed by the muon and the jet is exploited to distinguish muons from b–jets from muons from c– and light–jets. The large dataset at $\sqrt{s} = 8$ TeV allows the use of a pure $t\bar{t}$ sample where both W bosons decay leptonically. Events with exactly 2 jets are selected, where one of the two is b–tagged. By construction, the other jet originates from a bottom quark as well, and can be used to probe the b–tagging efficiency.

---

[5] Jets originating from the fragmentation of light quarks or gluons.

The efficiency for c jets is measured in samples of jets with D^* mesons, where the yields of D^* mesons with and without b tagging requirements are compared. The efficiency for b tagging light jets is measured in an inclusive jet sample, using the *negative tag method*, where a small significance of tracks impact parameter and of decay length of secondary vertices is required.

Tag Rate Function method

When requiring ≥ 1 b-tagged jet, the amount of simulated events is significantly reduced for some particular background processes, leading to large fluctuations in the predicted distributions. To overcome this problem, the Tag Rate Function (TRF) method is introduced. Events from simulation are not rejected based on the b-tagging count requirement. Instead, they are assigned a weight corresponding to the probability for the required number of b-jets to be present. Appendix A describes the TRF method in greater detail.

5.6 Missing Transverse Energy

At the LHC the overall momentum of the pp collision is zero, as the colliding protons have equal energies and opposite direction. However, for inelastic scattering, spectator partons travel down the beam pipe without interacting with the detector. Hence, while the z momentum of the colliding partons is unknown, the momentum transverse to the beam pipe is very close to zero. Since the detector provides a near 4π coverage of the solid angle, the requirement of transverse momentum conservation can be used to estimate the transverse momentum of neutrinos, which do not interact with the detector, as the missing transverse energy (E_{T}).

The E_{T} is computed by combining all the topoclusters found in the calorimeters. Each topocluster is associated with reconstructed objects: electrons, photons, jets and muons. The topoclusters are then calibrated according to the reconstructed physics objects to which they are associated. Topoclusters which are not associated with any physics object are also considered and a dedicated calibration is applied. The E_{T} and its components are therefore defined as:

$$\begin{aligned} E_T^{\mathrm{miss}} &= \left| -\sum p_T \right| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \\ E_x^{\mathrm{miss}} &= -\sum p_x \\ E_y^{\mathrm{miss}} &= -\sum p_y \end{aligned} \tag{5.4}$$

where the sums run over all the topoclusters in the event.

Chapter 6

Analysis strategy

This chapter introduces the strategies implemented to measure the asymmetries in the datasets collected at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. Due to the different amount of data and data taking conditions for the two datasets (see Sec. 3), slightly different procedures are adopted, in order to obtain the best precision achievable in both scenarios.

6.1 General strategy

The goal of the analysis is the measurement of the charge asymmetry A_C integrated over the kinematics of the $t\bar{t}$ system, referred to as *inclusive* A_C, and its differential spectra as a function of the invariant mass of the top quark pair $m_{t\bar{t}}$, its transverse momentum $p_{t\bar{t}}$, and the absolute value of its rapidity $y_{t\bar{t}}$. In order to achieve this goal, three steps are required:

- Events where top quark pairs are produced are selected and the background composition in the sample is estimated.
- The kinematics of the $t\bar{t}$ system is reconstructed in order to access the top quark and antiquark directions and compute the $\Delta\ y$ observable as described in Sec. **??**.
- The information about the distortions introduced by the two steps above is used to estimate the parton level asymmetries with an *unfolding* procedure.

While the same reconstruction and unfolding procedures, described in the following chapters, are used for both datasets, the selection strategy and the background estimation is driven by the amount of data available. The dataset at $\sqrt{s} = 8$ TeV is approximately six times larger than the one at $\sqrt{s} = 7$ TeV, with two implications: on one hand the statistical uncertainties are greatly reduced, so that the constraint of systematic uncertainties becomes crucial to perform precise measurements; on the other hand, the large dataset can be split in subsamples without significant loss in statistical power, and the properties of the background can be exploited to determine its composition. Therefore, the measurements at $\sqrt{s} = 7$ TeV are performed in a single pure sample of $t\bar{t}$ events, while the measurements at $\sqrt{s} = 8$ TeV are performed using six subsamples with different background and signal compositions. The following sections describe the details of the selection requirements applied to the reconstructed objects in order to maximize the $t\bar{t}$ signal purity. In addition, the chapter illustrates the techniques used to estimate the background composition of the sample and the comparison between data and predictions used to validate such estimations.

	$\overline{s} = 7$ TeV			
channel	$\mu +$ jets	$e +$ jets		
trigger	*EF_mu18* *EF_mu18_medium*	*EF_e20_medium*	*EF_e22_medium* *EF_e22vh_medium1*	*EF_e45_medium1*
trigger $p_{\mathrm{T}min}$	18 GeV	20 GeV	22 GeV	45 GeV
offline $p_{\mathrm{T}min}$	20 GeV	25 GeV		

	$\overline{s} = 8$ TeV			
channel	$\mu +$ jets		$e +$ jets	
trigger	*EF_mu24i_tight*	*EF_mu36_tight*	*EF_e24vhi_medium1*	*EF_e60_medium1*
isolation	$p_{\mathrm{T},tracks}^{0.2}\ \ p_{\mathrm{T}\mu} < 0\ 12$		$p_{\mathrm{T},tracks}^{0.2}\ \ E_{\mathrm{T}e} < 0\ 1$	
trigger $p_{\mathrm{T}min}$	24 GeV	36 GeV	24 GeV	60 GeV
offline $p_{\mathrm{T}min}$	25 GeV		25 GeV	

Table 6.1: Lepton p_{T} and isolation requirements for the single lepton triggers during 2011 and 2012 data-taking.

6.2 Event selection

As discussed in Sec. 2.2.2, top quarks decay mainly into a W boson and a b quark. The W boson decays into two quarks or into a lepton and a neutrino. Therefore, the semi-leptonic decay channel corresponds to an event topology with at least four reconstructed jets, exactly one isolated lepton (muon or electron) and missing transverse energy.

A preliminary selection of the events is performed during data acquisition by the single lepton triggers. Events are recorded when at least one high-p_{T} muon or electron is reconstructed by the trigger algorithms. The combination of p_{T} and isolation requirements is optimized to ensure a good selection efficiency in different luminosity conditions. Therefore several trigger configurations are considered, and events satisfying at least one of the trigger decisions are used in the analyses. In the $\overline{s} = 7$ TeV dataset, the *EF_e20_medium*, *EF_e22_medium*, *EF_e22vh_medium1*, *EF_e45_medium1*, configurations are used to select $e +$ jets events, while the *EF_mu18* and *EF_mu18_medium* triggers select $\mu +$ jets events. In the $\overline{s} = 8$ TeV dataset, events accepted by either *EF_e24vhi_medium1* or *EF_e60_medium1* are used in the $e +$ jets channel, while *EF_mu24i_tight* and *EF_mu36_tight* triggers are used for $\mu +$ jets events. The p_{T} and isolation requirements for each trigger are detailed in Table 6.1.

The events used in the analysis are required to contain exactly one lepton reconstructed offline as described in Sec. 5. The lepton selected offline is required to match within $\Delta R < 0\ 15$ the corresponding lepton at the trigger level. Minimum p_{T} requirements ensure that the trigger efficiencies are maximized and constant as a function of the lepton p_{T} for the selected events. The electron p_{T} requirement is also used to reduce the amount of multijet background in the sample. The offline thresholds for the various scenarios are detailed in Table 6.1.

In addition to the lepton, a minimum of four jets with p_{T} larger than 25 GeV are required to be reconstructed in the event. Only jets reconstructed in the range $\eta < 2\ 5$ are considered (see Sec. 5.5). The number of jets tagged as b jets is used to define $t\bar{t}$ (with at least one b jet) or background (without b jets) enriched regions. The measurements at $\overline{s} = 7$ TeV are performed

	$\overline{s} = 7$ TeV	$\overline{s} = 8$ TeV		
channel	≥ 1 b-jets	0 b-jets	1 b-jet	≥ 2 b-jets
$\mu +$ jets	$E_{\mathrm{T}} > 20$ GeV $E_{\mathrm{T}} + m_{\mathrm{T}}^{W} > 60$ GeV	$E_{\mathrm{T}} > 40$ GeV $E_{\mathrm{T}} + m_{\mathrm{T}}^{W} > 60$ GeV	$E_{\mathrm{T}} > 20$ GeV $E_{\mathrm{T}} + m_{\mathrm{T}}^{W} > 60$ GeV	
$e +$ jets	$E_{\mathrm{T}} > 30$ GeV $m_{\mathrm{T}}^{W} > 30$ GeV	$E_{\mathrm{T}} > 40$ GeV $E_{\mathrm{T}} + m_{\mathrm{T}}^{W} > 60$ GeV	$E_{\mathrm{T}} > 20$ GeV $E_{\mathrm{T}} + m_{\mathrm{T}}^{W} > 60$ GeV	

Table 6.2: Minimum E_{T} and m_{T}(W) requirements.

in a $t\bar{t}$ enriched sample with at least one b jet. For the measurements at $\overline{s} = 8$ TeV the b jet multiplicity is exploited to estimate precisely the background composition considering three categories of events: without b jets, with exactly 1 b jet, with at least 2 b jets.

In order to suppress the QCD multijet and Z+jets backgrounds, requirements on the E_{T} and the transverse mass m_{T} of the leptonically-decaying W boson are applied.[1] Since the sample composition varies with the lepton flavor and b-jet multiplicity, different requirements are applied for each scenario, as detailed in Table 6.2. This is particular important for the measurements at $\overline{s} = 8$ TeV, where the suppression of the $Z +$ jets background allows a more precise estimation of the $W +$ jets background.

6.3 Signal and background modeling

The requirements described in Sec. 6.2 are designed to select semi-leptonic decays of the $t\bar{t}$ pair. However, fully leptonic decays can also be selected when one of the two leptons is not reconstructed. These events are considered as part of the signal sample, notwithstanding that the top quark pair kinematics cannot be properly reconstructed under the semi-leptonic assumption.

Even though the requirements are designed to select the $t\bar{t}$ topology, other processes contaminate the samples: W and Z boson production in association with jets (W+jets, Z+jets), single top, QCD multijet, and diboson (WW,ZZ,WZ production constitute a non-negligible fraction of the selected events. In the case of $W +$ jets events, the largest background, the leptonic W decay produces a high-p_{T} isolated lepton, while the additional jets production, including b-jets, mimics the hadronic top quark decay. Analogously the leptonic decays of a boson in $Z +$ jets and diboson processes feature high-p_{T} leptons reconstructed with high efficiency. While the misidentification of jets as reconstructed leptons is very rare, the large cross section to QCD multijet still results in a significant amount of background events, due to the large cross section.

The background composition of the data samples is estimated with Monte Carlo simulation and data-driven techniques. In particular the expected yield of $W +$ jets events and the flavor composition of the associated jets is calibrated in-situ, as well as the QCD multijet background.

6.3.1 Simulated samples

Different MC generators, PDF sets and parton shower and fragmentation algorithms are used to model signal and backgrounds in the data samples collected at $\overline{s} = 7$ TeV and at $\overline{s} = 8$ TeV,

---

[1] $m_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\ell} \not{E}_{\mathrm{T}} (1 - \cos \Delta\phi)}$, with p_{T}^{ℓ} being the transverse momentum (energy) of the muon (electron) and $\Delta\phi$ the azimuthal angle separation between the lepton and the direction of the missing transverse momentum.

with the goal of providing the most reliable possible prediction in each scenario. In the following the different choices for each process are listed.

The $t\bar{t}$+jets process is simulated using the `ALPGEN` generator with the `CTEQ6L1` PDF set, interfaced to the `HERWIG` parton shower (2011) and using `POWHEG` with the `CT10` PDF set, interfaced to the `PYTHIA` parton shower (2012).

Simulated samples of $W\ Z$ boson production are generated with up to five additional partons using the `ALPGEN` generator and the `CTEQ6L1` PDF set, interfaced to `HERWIG` (2011) or `PYTHIA` (2012) for parton showering and fragmentation. Dedicated samples are produced to simulate W production in association with bottom quark pairs ($Wb\bar{b}$+jets), charm quark pairs ($Wc\bar{c}$+jets), single charm quarks (Wc + jets), and light flavor partons, including gluons (W + light jets). The Z+jets samples are generated separately for Z+light jets, $Zb\bar{b}$+jets, and $Zc\bar{c}$+jets and normalized to the inclusive NNLO theoretical cross section [70]. Overlap between $W\ ZQ\bar{Q}$+jets ($Q = b\ c$) events generated from the matrix element calculation and those generated from parton-shower evolution in the $W\ Z$+light jets samples is avoided by using the matrix element prediction only for $\Delta R(Q\ \bar{Q}) > 0\ 4$.

Simulated samples of single top quark backgrounds corresponding to the s-channel and Wt production mechanisms are generated with `MC@NLO` interfaced with the `HERWIG` parton shower (2011) and `POWHEG` interfaced with `PYTHIA` (2012), using the `CT10` PDF set. The simulation of t-channel single top quark production is generated with the `ACERMC` LO generator with the `MRST LO**` PDF set (2011) and `POWHEG` interfaced with `PYTHIA` with the `CT10` PDF set (2012). Single top samples are normalized to the corresponding NLO cross sections [71 73]. Finally, diboson production is modelled using `HERWIG` with the `MRST LO**` PDF set, and is normalized to the NLO theoretical cross sections [74].

6.3.2 $W+\text{jets}$ background normalization

The normalization of the $W+\text{jets}$ background is measured in data in order to constrain its uncertainty. The procedure exploits the difference in production cross section at LHC between W^+ and W^- to estimate the $W+\text{jets}$ yield. Due to the higher density in protons of u quarks with respect to d quarks, the cross sections $\sigma(u\bar{d}\quad W^+)$ and $\sigma(d\bar{u}\quad W^-)$ are different, with a larger production rate for W^+. The prediction for the W boson charge asymmetry in $W+\text{jets}$ production is less affected by theoretical uncertainties [75] and can be measured in data to derive the correct overall normalization for the MC prediction. The total number of $W+\text{jets}$ events in the selected data sample $N_W = N_{W^+} + N_{W^-}$ is estimated as

$$N_W = \left(\frac{N_{W^+} + N_{W^-}}{N_{W^+} - N_{W^-}}\right)_{\text{MC}} (N_{W^+} - N_{W^-})_{\text{meas}} \tag{6.1}$$

where positive and negative W bosons are identified from the charge of the reconstructed lepton.

The W boson charge asymmetry observed in simulation depends on the flavor composition of the sample, as the size and sign of the asymmetry varies for $Wb\bar{b}+\text{jets}$, $Wc\bar{c}+\text{jets}$, $Wc+\text{jets}$ and W + light jets production. Therefore a calibration of the flavor composition is derived simultaneously with the estimation of the total normalization in Eq. 6.1. The relative fractions are estimated in a $W+\text{jets}$ enriched control region where exactly two reconstructed jets are required and no b tagging requirement is applied (*pretag* region). The additional requirement of at least one b tagged jet is applied to define a $Wb\bar{b}+\text{jets}$ enriched region (*tag* region) and

channel	$K_{b\bar{b}/c\bar{c}}$	K_c	K_{light}
$\mu +$ jets	$1\ 2 \pm 0\ 4$	$1\ 0 \pm 0\ 4$	$0\ 97 \pm 0\ 09$
$e +$ jets	$1\ 4 \pm 0\ 4$	$0\ 7 \pm 0\ 4$	$1\ 00 \pm 0\ 10$

Table 6.3: Calibration factors for flavor composition and overall normalization of the $W +$jets background as measured in the 2011 dataset.

the $W +$ jets event yield is given by:

$$N^{W,\text{tag}} = N^{W,\text{pretag}} \sum_{x=b\bar{b},c\bar{c},c,light} F_x P_x \tag{6.2}$$

where F_x are the flavor fractions $N_x^{\text{pretag}}\ N^{\text{pretag}}$ and P_x is the selection efficiency of the b tagging requirement for each flavor type $x = b\bar{b}\ c\bar{c}\ c\ light$. With the assumption that $Wb\bar{b}+$jets and $Wc\bar{c} +$ jets simulations require a fully correlated calibration, given the similarity of the processes, three calibration factors $K_{b\bar{b}/c\bar{c}} = F_{b\bar{b}}^{\text{data}}\ F_{b\bar{b}}^{\text{MC}} = F_{c\bar{c}}^{\text{data}}\ F_{c\bar{c}}^{\text{MC}}$, $K_c = F_c^{\text{data}}\ F_c^{\text{MC}}$ and $K_{light} = F_{light}^{\text{data}}\ F_{light}^{\text{MC}}$ are estimated to fit data. The calibration factors are then extrapolated to the signal region, where at least four jets are required, using the MC prediction to renormalize to unity the sum of the flavor fractions. Table 6.3 summarizes the flavor fraction and normalization calibration factors derived in the $e+$jets and $\mu+$jets channel for the 2011 dataset at $\overline{s} = 7$ TeV.

For the analysis of the 2012 dataset a similar approach is used by performing the *in-situ* calibration embedded in the unfolding procedure described in Sec. 8. The three b jet multiplicity subsamples are further split according to the lepton charge, thus obtaining six independent samples. The b jet multiplicity provides information about the heavy-flavor composition of the $W +$ jets background, while the lepton charge is used to determine the normalization of each component. Data and predictions for the six channels used for the inclusive A_C measurement are compared in Fig. 6.1. The two background enriched subsamples without b jets provide little information about the $t\bar{t}$ asymmetry, therefore no shape information is used. The three calibration factors for $Wb\bar{b} +$ jets/$Wc\bar{c} +$ jets, $Wc +$ jets and $W +$ light jets obtained from the inclusive A_C measurements are respectively $1\ 50 \pm 0\ 11$, $1\ 07 \pm 0\ 27$ and $0\ 80 \pm 0\ 04$ for the combined $+$ jets sample.

6.3.3 Multijet background

Multijet events can pass the selection criteria when a lepton is reconstructed in the event. The dominant sources of spurious leptons are:

- semi leptonic b-jet decays;
- long lived weakly decaying states such as $\pi^{\pm}$ or K mesons;
- reconstruction of π^0 showers as electrons;
- reconstruction of electrons from conversions or direct photons.

While the probability of a QCD multijet event being selected is very low, the production cross section for multijet events is orders of magnitude larger above that of $t\bar{t}$ production. Because of this and the fact that this probability depends on the detector configuration and

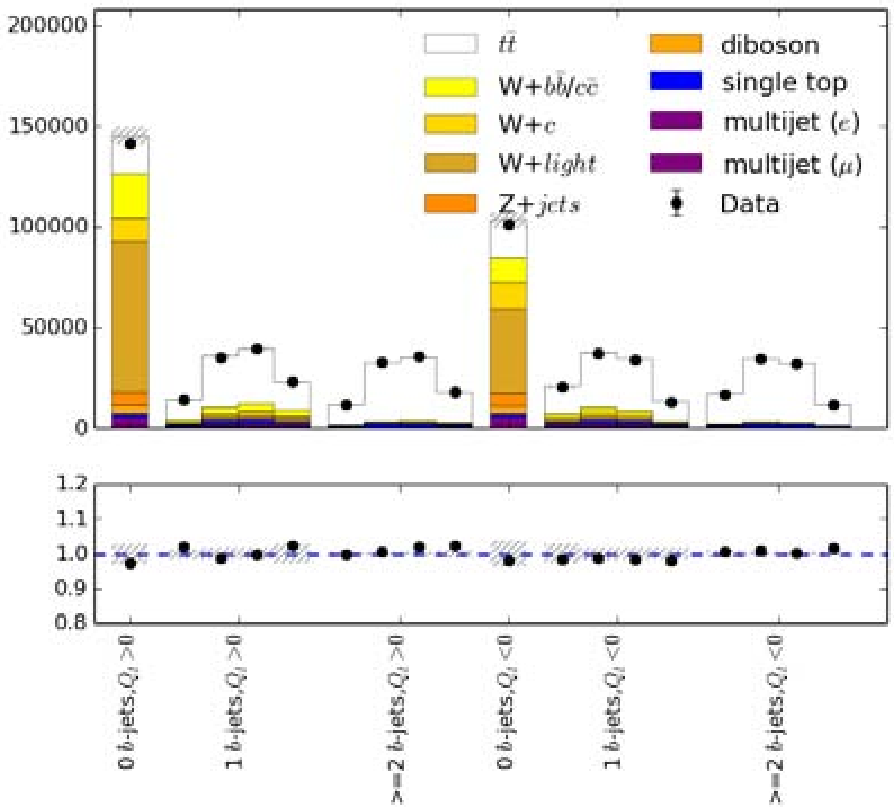


Figure 6.1: Comparison between data and predictions in the six channels considered for the measurement at $\sqrt{s} = 8$ TeV.

geometry, it is more efficient to determine the multijet background in the selected sample from data with a technique referred to as *matrix method* [76].

The matrix method is based on the selection of two categories of events: the ones that satisfy loose lepton selection requirements, and the ones that satisfy tight lepton selection requirements. The tight requirements are the ones used in the analysis (see Sec. 5.3 and Sec. 5.4), while the loose selection is obtained by applying looser isolation and identification criteria. Therefore, the tight lepton sample constitutes a subset of the loose lepton one. The number of events with one loose lepton and the number of events with one tight lepton can be written as:

$$\begin{aligned} N^{\text{loose}} &= N^{\text{loose}}_{\text{real}} + N^{\text{loose}}_{\text{fake}} \\ N^{\text{tight}} &= N^{\text{tight}}_{\text{real}} + N^{\text{tight}}_{\text{fake}} \end{aligned} \tag{6.3}$$

where $N^{\text{loose}}_{\text{real}}$ is the number of events with a real lepton satisfying loose lepton requirements, $N^{\text{loose}}_{\text{fake}}$ is the number of multijet events without a real lepton but still satisfying loose lepton requirements, and the same are $N^{\text{tight}}_{\text{real}}$ and $N^{\text{tight}}_{\text{fake}}$ for the tight lepton-requirements. The efficiency $\epsilon = N^{\text{tight}} \;\; N^{\text{loose}}$ is different for real leptons and for fake leptons. The signal efficiency

$$\epsilon_{\text{real}} = \frac{N^{\text{tight}}_{\text{real}}}{N^{\text{loose}}_{\text{real}}} \tag{6.4}$$

is measured in data events with Z boson decays in two leptons. The fake efficiency

$$\epsilon_{\text{fake}} = \frac{N^{\text{tight}}_{\text{fake}}}{N^{\text{loose}}_{\text{fake}}} \tag{6.5}$$

is measured in data control regions with small E_{T} and m^{W}_{T}, where the contribution from fake leptons is larger. While ϵ_{real} is close to unity and does not depend on the event topology, ϵ_{fake} is parametrized as a function of kinematic observables such as lepton η and $\Delta R(\;\; jet)$. With these two efficiencies, one can solve the two linear equations (6.3) for $N^{\text{tight}}_{\text{fake}}$ as a function of N^{loose} and N^{tight}, obtaining:

$$N^{\text{tight}}_{\text{fake}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (N^{\text{loose}} \epsilon_{\text{real}} - N^{\text{tight}}) \tag{6.6}$$

The sample of multijet events is therefore estimated by the weighted data events, where the weight for tight events is $\epsilon_{\text{fake}}(\epsilon_{\text{real}} - 1) \;\; (\epsilon_{\text{real}} - \epsilon_{\text{fake}})$ and the weight for loose-not-tight events is $(\epsilon_{\text{fake}} \cdot \epsilon_{\text{real}}) \;\; (\epsilon_{\text{real}} - \epsilon_{\text{fake}})$.

6.4 Comparison between data and prediction

The expected and observed number of selected events in both the $\mu +$ jets and the $e +$ jets channels are reported in Table 6.4 for the 2011 dataset, and in Tables 6.5 and 6.6 for 2012.

The modeling of the main background process is validated by comparing distributions as observed in the data samples with the corresponding predictions from simulation. Before applying b tagging requirements, the sample composition can be studied without biases toward

Channel	μ + jets pretag	μ + jets tag	e + jets pretag	e + jets tag
$t\bar{t}$	34900 ± 2200	30100 ± 1900	21400 ± 1300	18500 ± 1100
W + jets	28200 ± 3100	4800 ± 900	13200 ± 1600	2300 ± 900
Multi jets	5500 ± 1100	1800 ± 400	3800 ± 1900	800 ± 400
Single top	2460 ± 120	1970 ± 100	1530 ± 80	1220 ± 60
Z + jets	3000 ± 1900	480 ± 230	3000 ± 1400	460 ± 220
Diboson	380 ± 180	80 ± 40	230 ± 110	47 ± 22
Total background	40000 ± 4000	9200 ± 1000	21700 ± 2900	4800 ± 1000
Signal + background	74000 ± 4000	39300 ± 2100	43100 ± 3100	23300 ± 1600
Observed	70845	37568	40972	21929

Table 6.4: Numbers of expected events for the $t\bar{t}$ signal and the various background processes and observed events in data for the dataset at $\sqrt{s} = 7$ TeV. The uncertainties include statistical and systematic components.

Channel	μ + jets pretag	μ + jets 0-tag	μ + jets 1-tag	μ + jets 2-tag
$t\bar{t}$	197000 ± 400	28480 ± 170	88230 ± 300	93810 ± 310
Single top	16100 ± 130	3260 ± 60	8070 ± 90	5560 ± 70
W+jets	171700 ± 400	147800 ± 400	21150 ± 150	3310 ± 60
Z+jets	17820 ± 130	15270 ± 120	2160 ± 50	623 ± 25
Diboson	6140 ± 80	5220 ± 70	855 ± 29	90 ± 10
QCD	15600 ± 120	9030 ± 100	4930 ± 70	2590 ± 50
Total background	227400 ± 500	180500 ± 400	37170 ± 200	12170 ± 110
Total expected	424400 ± 600	209000 ± 500	125390 ± 350	105980 ± 330
Observed	427138	207057	126105	111248

Table 6.5: Numbers of expected events for the $t\bar{t}$ signal and the various background processes and observed events in data for the μ + jets dataset at $\sqrt{s} = 8$ TeV. The uncertainties include statistical and systematic components.

Channel	$e+$ jets pretag	$e+$ jets 0-tag	$e+$ jets 1-tag	$e+$ jets 2-tag
$t\bar{t}$	159800 ± 400	23040 ± 150	71440 ± 270	77120 ± 280
Single top	13190 ± 110	2670 ± 50	6590 ± 80	4660 ± 70
W+jets	126900 ± 400	109150 ± 330	15720 ± 130	2450 ± 50
Z+jets	32710 ± 180	28400 ± 170	3690 ± 60	1178 ± 34
Diboson	5420 ± 70	4590 ± 70	765 ± 28	100 ± 10
QCD	24970 ± 160	18540 ± 140	6020 ± 80	1560 ± 40
Total background	203200 ± 500	163300 ± 400	32780 ± 180	9950 ± 100
Total expected	362900 ± 600	186400 ± 400	104220 ± 330	87070 ± 300
Observed	354865	181295	99236	89462

Table 6.6: Numbers of expected events for the $t\bar{t}$ signal and the various background processes and observed events in data for the $e+$ jets dataset at $\sqrt{s} = 8$ TeV. The uncertainties include statistical and systematic components.

a specific flavor composition. As shown in the yield tables, the purity of the $t\bar{t}$ signal is about 50%, while the largest background comes from the $W+$ jets process. Fig. 6.2 shows that the various backgrounds cluster in different kinematic regions depending on their properties. For processes with real W bosons, such as top quark(s) and $W+$ jets production, the m_{T}^{W} distribution peaks at ≈ 90 GeV, while processes without neutrinos, such as $Z(\quad)$+jets and the multijet background, are more prominent at low E_{T} values.

The fraction of $t\bar{t}$ events can be enhanced by requiring one or at least two b tagged jets, and the pure samples is used to check the modeling of the kinematic quantities of the reconstructed objects used for selection and reconstruction, as shown in Fig. 6.3. The distributions in data of the lepton p_{T} and pseudo-rapidity, the leading[2] jet p_{T} and pseudo-rapidity are in good agreement with the predictions.

6.5 Systematic uncertainties

The precision of the measurements depends on two aspects: the amount of signal events in the sample, determining the statistical component of the uncertainty, and the systematic uncertainties. Sources of systematic uncertainties are the finite precision of the calibration of the reconstructed objects, the imperfections of signal and background modeling, the limited description of the experimental conditions (e.g. luminosity, pile-up). Systematics affect both the normalization of the total event yield and the shape of the observed distributions.

The individual sources of systematic uncertainties are treated as uncorrelated from each other, while the effect of each source is fully correlated across processes and channels. The sources of systematic uncertainties considered in this work are discussed in the following.

---

[2] Jets are conventionally ordered in decreasing p_{T}; hence the "leading jet" is the reconstructed jet with the highest transverse momentum

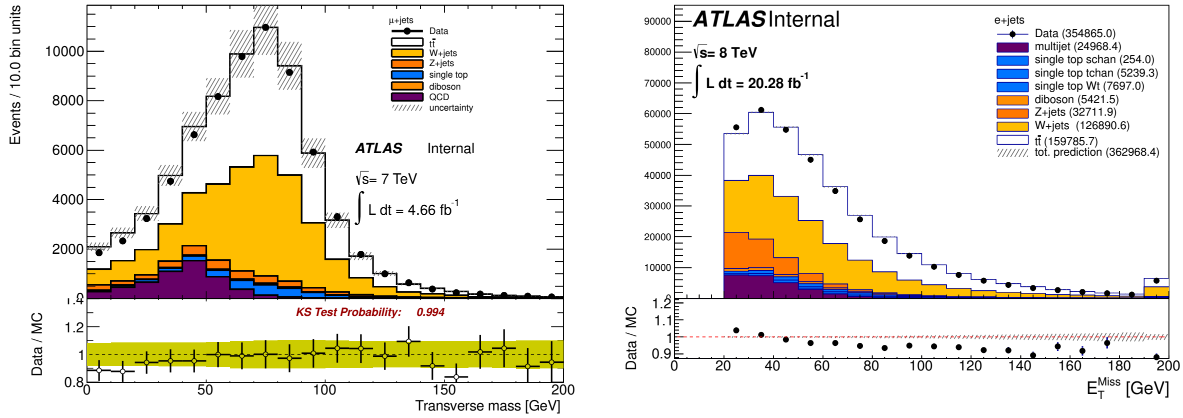
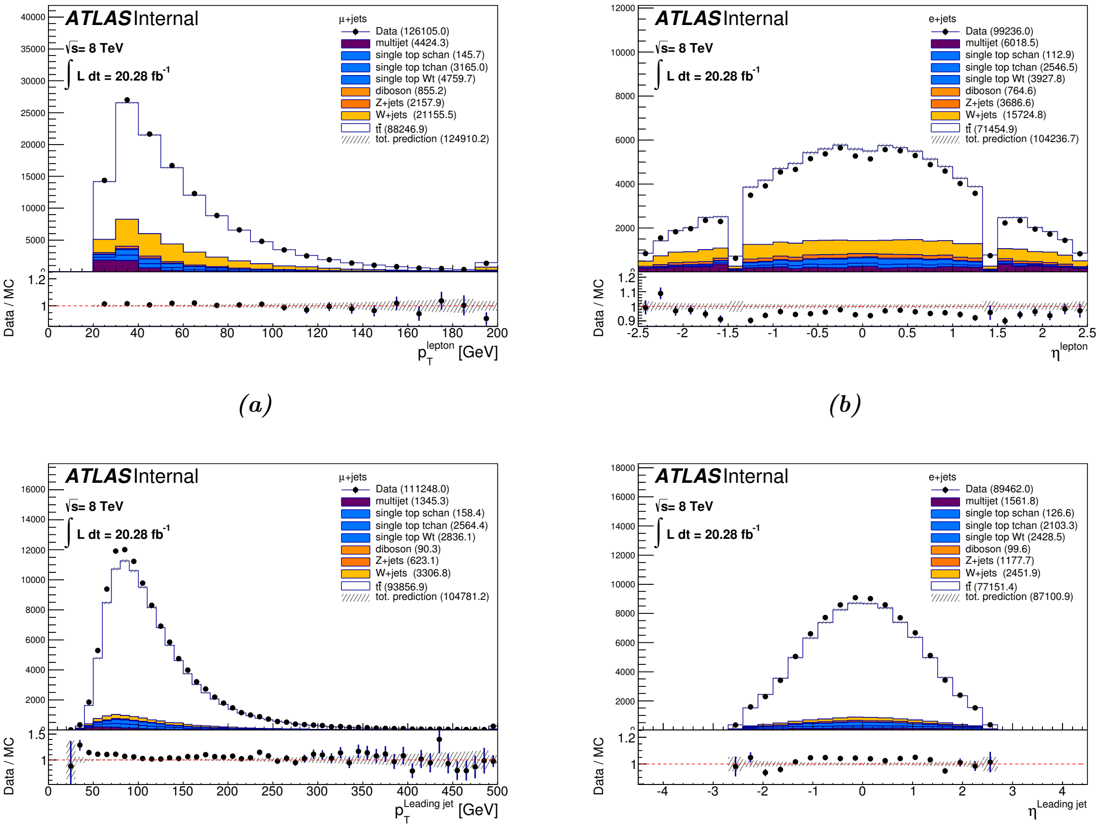


Figure 6.2: Comparison between data and prediction in the sample without b tagging requirements: the m_{T}^{W} distribution in the $\mu +$ jets channel at $\overline{s} = 7$ TeV (left), and the E_{T} distribution in the $e +$ jets channel at $\overline{s} = 8$ TeV (right)



(a) *(b)*

(c) *(d)*

Figure 6.3: Comparison between data and prediction in the $\overline{s} = 8$ TeV sample: distributions of (a) the lepton p_{T} in the $\mu +$ jets channel with exactly 1 b jet, (b) the lepton pseudo rapidity in the $e +$ jets channel with exactly 1 b jet, (c) the leading jet p_{T} in the $\mu +$ jets channel with at least 2 b jets and (d) the leading jet pseudo rapidity in the $e +$ jets channel with exactly 2 b jets.

6.5.1 Luminosity

The uncertainty on the absolute integrated luminosity is estimated to be of 1.8% at $\overline{s} = 7$ TeV [77] and 2.8% at $\overline{s} = 8$ TeV. This systematic uncertainty affects all processes for which the event yield from simulation is used, $Z+$jets, diboson, single top.

6.5.2 Object definitions

The object reconstruction and calibration introduces uncertainties associated with the definition of leptons, jets, E_{T} and on the jet flavor-tagging. In the following the corresponding systematic uncertainties are discussed.

Lepton reconstruction, identification and trigger efficiencies

The efficiencies of triggering, reconstructing and identifying leptons (Sections 5.3 and 5.4) are calibrated in the simulation with scale factors to match the ones in data. The calibrations are derived with tag-and-probe techniques in $Z \quad ^{+} \, ^{-}$ ($\; = e \;\, \mu$) data and simulated samples. The associated systematic uncertainty is taken into account by comparing the simulated samples with scale factors varied within their uncertainties.

In the $e+$jets channel, the overall effects on predicted yields are 0.3%, 1.1% and 0.2%, respectively for electron reconstruction, identification and trigger efficiency uncertainties. In the $\mu+$jets channel, the effects are 0.2%, 1.1% and 1.4%, respectively for muon reconstruction, identification and trigger efficiency uncertainties.

Lepton momentum scale and resolution

The lepton momentum scale and resolution is calibrated using simulated samples of $Z \quad ^{+} \, ^{-}$ and $J \qquad ^{+} \, ^{-}$. The calibration for muons is applied by adjusting the muon momentum scale and resolution in the simulation to match data. In the case of electrons, the energy resolution is smeared in simulated samples to match data, while the energy scale corrections are applied to data in all detector regions and to simulation only in the calorimeter transition region. The systematic uncertainties associated with these calibrations vary slightly the selection acceptance, with effects on the event yield below the percent level.

JVF efficiency

The efficiency of the JVF requirement (Sec. 5.5) is calibrated in simulation using scale factors derived in $Z(\quad ^{+} \, ^{-})$+1-jet events. The scale factor for pileup jets efficiency is constant and ~ 1, while the hard scatter jets efficiency is scaled up by a variable amount between $\approx 3\%$ for jets with $p_{\mathrm{T}} = 25$ GeV and down to 1% for jets with $p_{\mathrm{T}} > 150$ GeV

The uncertainties on the scale factors applied is propagated to the predicted event yields with effects of $\sim 2 \;\, 5\%$.

Jet energy scale

The Jet Energy Scale (JES) is derived from measurements at test-beam and measurements using data and simulation [78]. The associated p_{T} and η dependent uncertainty is propagated to the predicted distributions by varying up and down the energy of all reconstructed jets of each event. For each variation, the jet four momenta and the missing transverse momentum E_{T} are recomputed consistently to the varied p_{T} of the jets.

Pile-up activity introduces an additional source of systematic uncertainty that depends on the number of primary vertices and on the average number of interactions per bunch crossing $< \mu >$. Momentum balance techniques in Z+jets, γ+jets and multi-jet events are combined to derive a small residual correction for jets in the transverse momentum range 20 GeV $< p_{\mathrm{T}} \lesssim$ 1 TeV to take account of this effect.

The overall variation due to JES systematic uncertainty evaluated in the central detector region is ~4% for jets with $p_{\mathrm{T}} = 25$ GeV and improves to ~1% for jets with $p_{\mathrm{T}} = 500$ GeV.

Jet energy resolution

The Jet Energy Resolution (JER) is measured in data and simulation as function of the jet transverse momentum and pseudo rapidity. The measurement shows compatible resolutions in data and simulation, therefore no correction is applied. However, the quadratic difference between the JER in data and in simulated samples is used to smear the energy of jets in the simulation. The symmetrized templates with respect to the nominal case are used as positive and negative variations.

Flavor tagging

The efficiencies of flavor (b, c, or light) jets identification with the b-tagging algorithm (Sec. 5.5.1) are measured in data for each flavor [67 69]. In simulation b (c) jet efficiencies are calibrated with p_{T} and η dependent scale factors in the range 0.9 1.0 (1.1 1.2), the scale factor for light tagging efficiency is ~1.3. The uncertainty on these scale factors is between 7% and 13% for b jets, between 15% and 39% for c jets, and ~25% for light jets.

The systematic uncertainty on flavor tagging efficiency is divided into six independent components corresponding to the p_{T}bins used for the efficiency measurement. Therefore a total of 18 uncorrelated systematic uncertainties six per flavor is considered. For each component, a per-jet weighting procedure [79] is applied to simulated events in order to propagate the calibration uncertainties.

6.5.3 Background normalizations

Cross sections

The single top, $Z+$jets and diboson processes constitute a small fraction of the background, which is estimated using simulation only. Therefore the corresponding normalization uncertainty is determined by the precision of the associated theoretical cross section. Uncertainties of ±4 7%, ±4% and ±5% are assumed for the theoretical cross sections of the single top [**stopxs**, **stopxs˙2**], $Z+$jets and diboson [74] backgrounds respectively. In addition, a 48% Berends scaling uncertainty [80] is associated to the $Z+$jets and diboson normalizations.

$W+$jets calibration

For the $\overline{s} = 7$ TeV measurements, the overall $W+$jets normalization uncertainty results from the calibration procedure described in Sec. 6.3.2. The total uncertainty on the $W+$jets yield in the signal region is about 40% for e + jets channel and 20% in μ + jets. However the correlation with other sources of uncertainties is properly handled by applying the corresponding normalization and heavy flavor scale factor for each source of uncertainty considered. Such scale factors are determined by applying the calibration procedure using the varied templates for each sample.

For the $\sqrt{s} = 8$ TeV measurements, the calibration of the $W+$jets background is embedded in the unfolding procedure; therefore no a priori uncertainty is assigned. The overall normalization uncertainties on $Wb\bar{b}+$jets/$Wc\bar{c}+$jets, $Wc+$jets and $W+$light jets, obtained with the combined $+$jets measurement are about 10%, 25% and 5%, respectively.

PDF The charge asymmetry in W boson production, on which the calibration of the $W+$jets background relies, is due to the different density of positive and negative incoming partons. Therefore the modeling of the PDFs in the simulated W + jetssamples is relevant. In order to check the impact of the choice of the PDFs parametrization and its uncertainty, the measurements are repeated using W + jets simulations with variations of the CT10, MSTW, and NNPDF PDFs sets [81]. A small effect ($\sim 0\ 001$) is seen in the inclusive A_C measurement at $\sqrt{s} = 7$ TeV, while the effect is negligible for all other measurements.

Multijet

Systematics uncertainties affecting the multijet background originate from the difference between estimates obtained using different control regions and from the calibration of the method using simulated multijets events. For the $\sqrt{s} = 7$ TeV measurements, a 50% (20%) normalization uncertainty is assigned for the $e+$jets ($\mu+$jets) channel. In the $\sqrt{s} = 8$ TeV measurements a 50% uncertainty is assigned to the normalization of the multijet templates for each of the b-tagging multiplicities considered.

Chapter 7

Kinematic Reconstruction

In order to access the kinematic quantities of the $t\bar{t}$ system, an event by event estimation of the four momenta of $t\bar{t}$ pair is required. This section describes the kinematic fit performed to associate the reconstructed objects to the final state particles of the semi leptonic $t\bar{t}$ decay and obtain the four momenta of the top quark and antiquark.

7.1 Kinematic fit

A maximum likelihood approach [82] is used to determine the four-momenta of the top and of the antitop quarks from the four-momenta of the reconstructed decay products: at least four reconstructed jets, one lepton, and the missing transverse momentum. The kinematic likelihood fit finds the most likely association between reconstructed objects and $t\bar{t}$ final state particles: two b quarks, two light quarks, one lepton and one neutrino. The free parameters of the likelihood are the parton energies, the lepton transverse momentum and the three neutrino momentum components. The likelihood is defined as

$$
\begin{aligned}
L \;=\; & \mathcal{B}(\widetilde{E}_{\mathrm{p,1}}\;\; \widetilde{E}_{\mathrm{p,2}}\; m_W\;\; \Gamma_W) \cdot \mathcal{B}(\widetilde{E}_{\mathrm{l}}\;\; \widetilde{E}_{\nu}\; m_W\;\; \Gamma_W) \cdot \\
& \mathcal{B}(\widetilde{E}_{\mathrm{p,1}}\;\; \widetilde{E}_{\mathrm{p,2}}\;\; \widetilde{E}_{\mathrm{p,3}}\; m_t\;\; \Gamma_t) \cdot \mathcal{B}(\widetilde{E}_{\mathrm{l}}\;\; \widetilde{E}_{\nu}\;\; \widetilde{E}_{\mathrm{p,4}}\; m_t\;\; \Gamma_t) \cdot \\
& \mathcal{W}(E_x^{miss}\;\; \widetilde{p}_{x,\nu}) \cdot \mathcal{W}(E_y^{miss}\;\; \widetilde{p}_{y,\nu}) \cdot \mathcal{W}(E_{\mathrm{lep}}\; \widetilde{E}_{\mathrm{lep}}) \cdot \\
& \prod_{\mathrm{i}=1}^{4} \mathcal{W}(E_{\mathrm{jet,i}}\; \widetilde{E}_{\mathrm{p,i}}) \cdot P(b \text{ tag } \quad \text{quark flavor})
\end{aligned}
\tag{7.1}
$$

where:

- the $\widetilde{\square}$ are the partonic quantities to be estimated and $\square$ are the measured values;
- $\mathcal{B}(\widetilde{E}\; m\;\; \Gamma) \quad ((E^2 - m^2)^2 + m^2\Gamma^2)^{-1}$ is the Breit-Wigner parametrization of the parton energies $\widetilde{E}_{\mathrm{p,i}}$, lepton energy $\widetilde{E}_{\mathrm{lep}}$ and neutrino energy $\widetilde{E}_{\nu}$, in the W boson and top quark decays.
- $\mathcal{W}$ is the transfer function mapping the measured energy of a reconstructed object to the energy of the corresponding final state particle.

- The mass and width of the W boson and of the top quark are set to their measured values $m_W = 80\ 4$ GeV, $\Gamma_W = 2\ 1$ GeV, $m_t = 172\ 5$ GeV, and $\Gamma_t = 1\ 5$ GeV.

- The term $P(b - \text{tag}\quad \text{quark flavor})$ is used to weight the jet parton association based on b tagging information, and it is described in the following.

The likelihood is evaluated for all permutations of four reconstructed jets assigned to the four partons. In events where five or more jets were reconstructed, only the five highest-p_{T} jets are considered. The most probable combination, out of all the possible four-jets permutations, is chosen. The role of the probability term $P(b - \text{tag}\quad \text{quark flavor})$ is to favor the association of b jets with the b quarks from the top quark decays. The probability is defined, for a given permutation of four jets, as the product of b-tag efficiencies ϵ_b or mis tag rates ϵ_{light} of each jet j, based on the b tagging decision and the flavor of the associated quark:

$$P(b\text{−tag}\quad \text{quark flavor}) = \prod_j \begin{cases} \epsilon_b \text{ if tagged jet associated with } b \text{ quark} \\ (1-\epsilon_b) \text{ if non-tagged jet associated with } b \text{ quark} \\ \epsilon_{light} \text{ if non-tagged jet associated with quark from } W \text{ boson} \\ (1-\epsilon_{light}) \text{ if tagged jet associated with quark from } W \text{ boson} \end{cases} \tag{7.2}$$

The average values for ϵ_b and ϵ_{light} are 70% and 0 8% respectively, even though they mildly depend on jet p_{T} and pseudo rapidity.

The transfer functions $\mathcal{W}(E_{truth}\ E_{reco})$ describe the probability that the reconstructed energy E_{reco} corresponds to a true value E_{truth}. The transfer functions are parametrized as double gaussians

$$\mathcal{W}(E_{\text{truth}}\ E_{\text{reco}}) = \frac{1}{\overline{2\pi}(p_2 + p_3 p_5)} \left[\exp\left(-\frac{(\Delta E - p_1)^2}{2p_2^2} \right) + p_3 \exp\left(-\frac{(\Delta E - p_4)^2}{}2p_5^2 \right) \right] \tag{7.3}$$

where $\Delta E = (E_{truth} - E_{reco})\ E_{truth}$, and the p_i free parameters are estimated in different regions of E_{truth} and η for each particle. Figure 7.1 show, as example, the transfer functions for b quarks and for electrons.

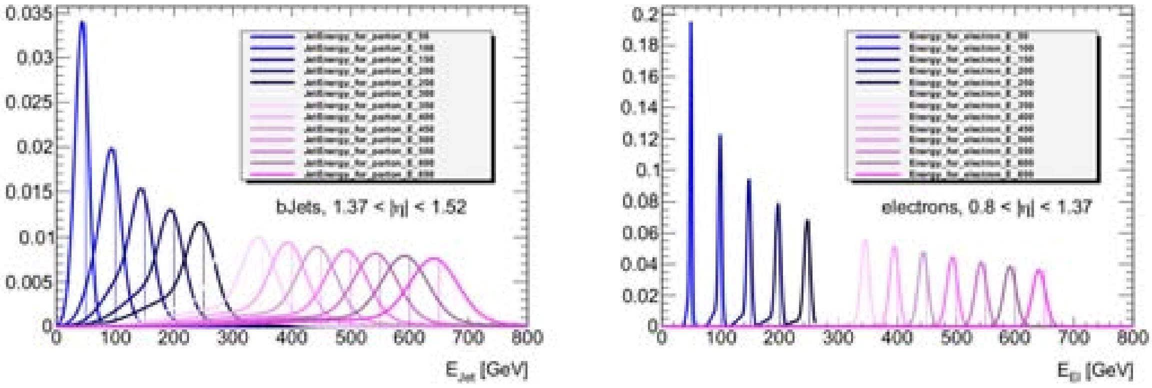


Figure 7.1: Energy evolution of the transfer functions for electrons in the range $0\ 8 < \eta < 1\ 37$ (left) and for b quarks in the range $1\ 37 < \eta < 1\ 52$ (right).

7.2 Performance

Due to selection, resolution and combinatorics effects, the efficiency of reconstructing a $t\bar{t}$ pair matching[1] the original partons is 25%. The dominant cause of mis reconstruction is the wrong jet assignment to the corresponding $t\bar{t}$ final state parton. This typically occurs in three cases:

- *jets outside acceptance*: the jet originating from a given parton does not satisfy the reconstruction requirements, such as $p_{\mathrm{T}} > 25$ GeV or pseudo rapidity within the $\eta < 2\ 5$ range. The fraction of events where all $t\bar{t}$ final state jets are reconstructed is $25 - 40\%$, depending on the b jets requirement.
- *jets outside selection*: not all four jets corresponding to the $t\bar{t}$ decay are used in the reconstruction algorithm. This occurs approximately 50% of the times for events with more than five jets and all matched jets within acceptance.
- *jets mis-assigned*: one or more jets are not assigned to the correct parton. It occurs about $15 - 25\%$ of the times for events where the correct jets where used in the reconstruction algorithm, depending on the b jet requirement.

The impact of mis reconstruction is reflected on the resolution of the kinematic quantities of the reconstructed $t\bar{t}$. The plots in Fig. 7.2 show the two dimensional distributions of the reconstructed quantities in relation to their true values. The resolution on $m_{t\bar{t}}$ and $p_{t\bar{t}}$ improves at higher values, while the top rapidity shows a small spread over the whole range.

For what concerns the asymmetry measurement, the relevant benchmark of the reconstruction performance is the efficiency of reconstructing correctly the Δy sign, defined as

$$\epsilon = \frac{N(\Delta y_{truth} \cdot \Delta y_{reco} > 0)}{N_{tot}} \tag{7.4}$$

Events where the Δy sign is measured incorrectly can occur when the top and antitop quarks are not reconstructed from the correct final state particles (mis-reconstruction). Even when the reconstruction algorithm associates the correct particles, small Δy values can be reconstructed with the wrong sign due to resolution effects. As a result the asymmetry of the distribution of reconstructed Δy is diluted by a factor

$$D = 2 \times \epsilon - 1 \tag{7.5}$$

The efficiency ϵ, measured in the $t\bar{t}$ simulated sample, is $\approx 75\%$ corresponding to a dilution factor $D = 2 \times 0\ 75 - 1 = 0\ 5$. The efficiency ϵ depends on the kinematic region considered. As shown in Fig. 7.3, the Δy reconstruction performs better at high $m_{t\bar{t}}$.

7.3 Comparison between data and prediction

The modeling of the reconstructed quantities is validated by comparing the distributions as observed in data with the prediction. Figure 7.4 shows the comparison of the distributions of the top quark pair invariant mass $m_{t\bar{t}}$, transverse momentum $p_{t\bar{t}}$, rapidity $y_{t\bar{t}}$, and velocity along the beam axis $\beta_{z,t\bar{t}}$ for the 2011 dataset at $\overline{s} = 7$ TeV.

---

[1] A reconstructed object i is considered matched with a parton j if the condition $\Delta R\,(i,j) < 0.2$ is satisfied.

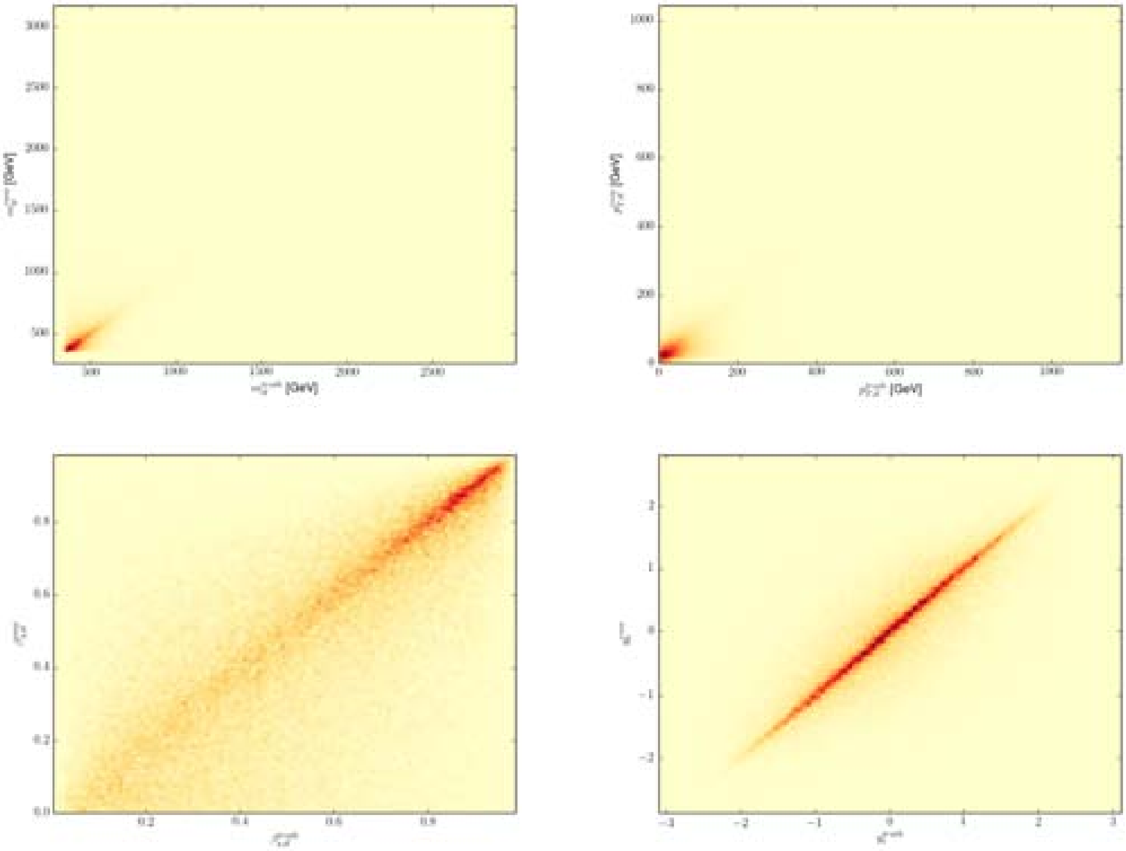


Figure 7.2: Distributions of kinematic quantities of the reconstructed $t\bar{t}$ pair in relation to their true values.

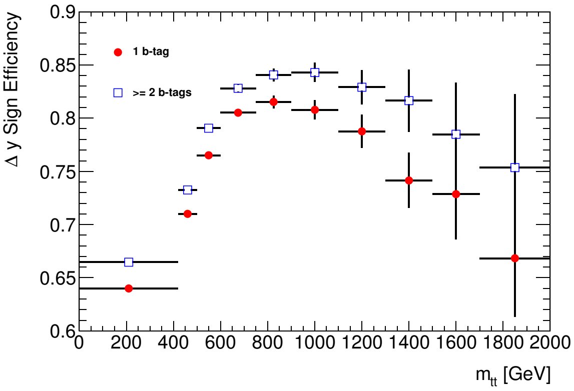


Figure 7.3: Probability to reconstruct the correct $\Delta|y|$ sign as a funcion of the $t\bar{t}$ invariant mass m_{tt} for the $\mu +$jets sample.

7.3.1 Likelihood requirement

For the analysis of the 2011 dataset collected at $\sqrt{s} = 8$ TeV, a requirement on reconstruction quality is applied to improve the resolution on reconstructed quantities. A 20% reduction on the

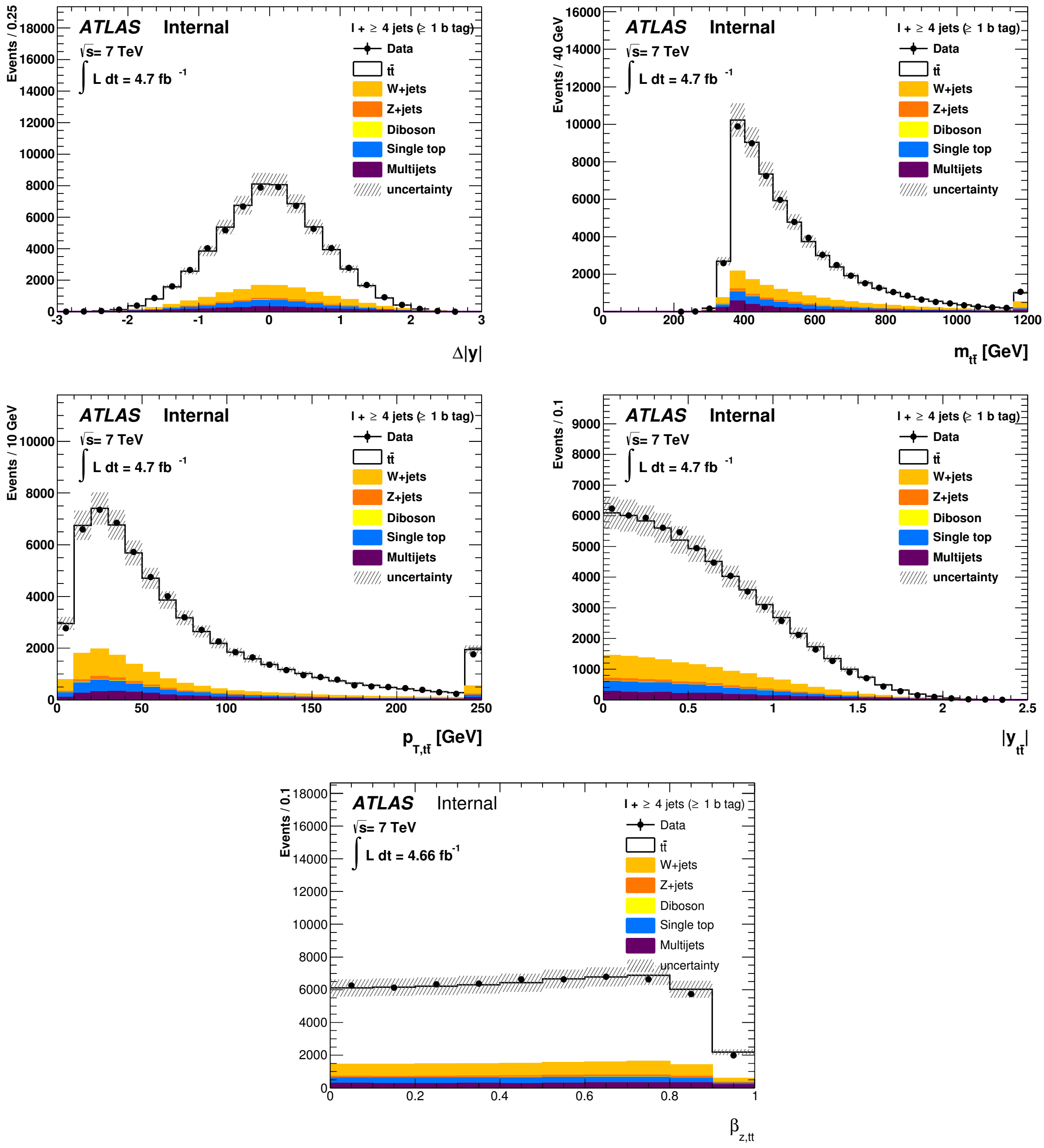


Figure 7.4: Control plots for $\Delta\ y$ (top left), invariant mass $m_{t\bar{t}}$ (top right), transverse momentum $p_{t\bar{t}}$ (centre left), rapidity $y_{t\bar{t}}$ (centre right) and velocity $\beta_{z,t\bar{t}}$ (bottom) distributions for the $e+$jets and $\mu+$jets channels combined in the tagged sample at $\overline{s} = 7$ TeV. The uncertainty on the total prediction includes both the statistical and the systematic components.

A_C uncertainty in the highest mass bin ($m_{t\bar{t}} > 750$ GeV) is achieved by requiring the logarithm of the likelihood (*log-likelihood*) to be > -55, as shown in Figure 7.5.

The modeling of the log-likelihood is validated by comparing its distribution in data with the prediction, as shown in Figure 7.6.

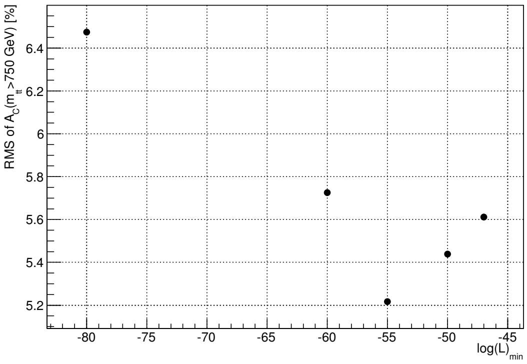


Figure 7.5: Scan of the statistical uncertainty on A_C for $m_{t\bar{t}} > 750$ GeV with different requirement on log-likelihood.

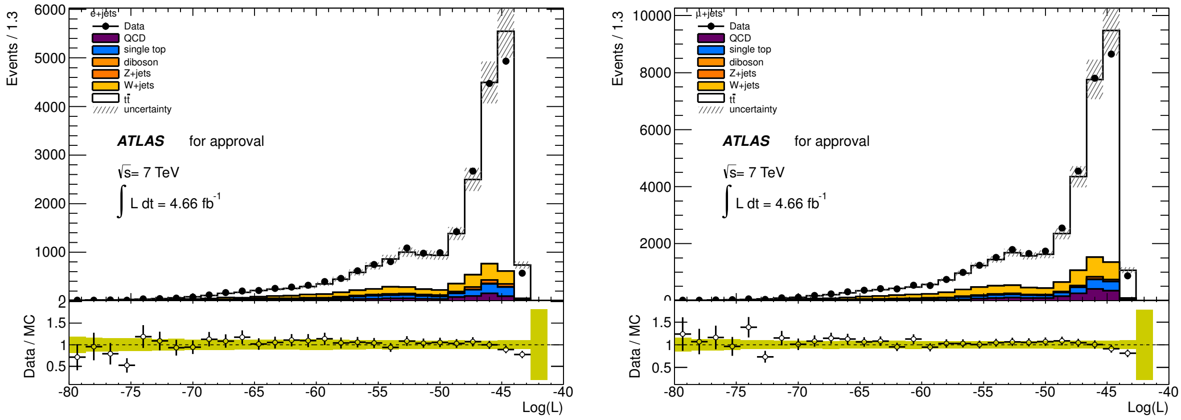


Figure 7.6: log-likelihood distributions for $e+$jets (left) and for $\mu+$jets (right).

Chapter 8

Unfolding

In order to allow a direct comparison between the experimental results and the theoretical predictions, the parton level asymmetry spectra need to be estimated. As discussed in Sec. 7 the asymmetry computed from the reconstructed kinematics of the $t\bar{t}$ system is affected by the efficiency of correctly reconstructing the $\Delta\ y$ sign. The distortion induced on the A_C values depends on the kinematic region considered and it can dilute the asymmetry up to half of its original (*true*) value. The $t\bar{t}$ MC simulation is used to map the distortions, bin dependent *efficiencies* and *migrations*, caused by resolution and acceptance effects. The unfolding procedure consists in combining this information with the distributions observed in data in order to estimate the parton level distributions.

8.1 Fully Bayesian Unfolding

The Fully Bayesian Unfolding [83] consists in the application of Bayes theorem to the problem of unfolding. This application can be stated in the following terms: given an observed spectrum $\boldsymbol{D} \quad \mathbb{N}^{N_r}$ and a response matrix $\mathcal{M} \quad \mathbb{R}^{N_r} \times \mathbb{R}^{N_t}$, the posterior probability of the true spectrum $\boldsymbol{T} \quad \mathbb{R}^{N_t}$ follows the probability density

$$p\,(\boldsymbol{T}\ \boldsymbol{D}\ \mathcal{M}) \quad \mathcal{L}\,(\boldsymbol{D}\ \boldsymbol{T}\ \mathcal{M}) \cdot \pi\,(\boldsymbol{T}) \tag{8.1}$$

where $\mathcal{L}\,(\boldsymbol{D}\ \boldsymbol{T}\ \mathcal{M})$ is the conditional likelihood for $\boldsymbol{D}$ given $\boldsymbol{T}$ and $\mathcal{M}$, and π is the prior probability density for the true spectrum $\boldsymbol{T}$.

8.1.1 Likelihood

Under the assumption that the data are poissonian counts, the likelihood $\mathcal{L}\,(\boldsymbol{D}\ \boldsymbol{T}\ \mathcal{M})$ can be computed from the following two pieces of information, contained in the response matrix $\mathcal{M}$:

- the probability $P(r\ t)$ of an event to be produced in the true bin t and to be observed in the reconstructed bin r;
- the efficiency ϵ_t for an event produced in the true bin t and that is reconstructed in any bin r.

The above quantities allow the extrapolation of the reconstructed spectrum $\boldsymbol{R}$ corresponding to a given true spectrum $\boldsymbol{T}$ as in

$$r_i(\boldsymbol{T}\ \mathcal{M}) = \sum_{j=0}^{N_r} \epsilon_j \cdot p(r_i\ t_j) \tag{8.2}$$

The likelihood is then defined by comparing the observed spectrum $\boldsymbol{D}$ with the expected one, which includes the background prediction $\boldsymbol{B}$:

$$\mathcal{L}\left(\boldsymbol{D}\ \boldsymbol{T}\right) = \prod_{i=1}^{N_r} Poisson(d_i\ \ r_i(\boldsymbol{T}\ \mathcal{M}) + b_i) \tag{8.3}$$

8.1.2 Prior and sampling

While the response matrix can be estimated from the simulated sample of signal events, the prior probability density $\pi\left(\boldsymbol{T}\right)$ is to be chosen according to what we know about $\boldsymbol{T}$ before the measurement is performed. The simplest choice is an *uninformative* prior that assigns equal probabilities to all $\boldsymbol{T}$ spectra in a subset of N_t $[T^{\ulcorner}\ T^{\urcorner}]$:

$$\pi\left(\boldsymbol{T}\right) \qquad \begin{cases} 1 & \text{if } T_t \quad [T^{\ulcorner}\ T^{\urcorner}] \quad t \quad [1\ N_t] \\ 0 & \text{otherwise} \end{cases} \tag{8.4}$$

A more general definition for the prior is given by

$$\pi\left(\boldsymbol{T}\right) \qquad \begin{cases} e^{\alpha S(\boldsymbol{T})} & \text{if } T_t \quad [T^{\ulcorner}\ T^{\urcorner}] \quad t \quad [1\ N_t] \\ 0 & \text{otherwise} \end{cases} \tag{8.5}$$

where α is an arbitrary strength parameter, and $S(\boldsymbol{T})$ is a regularization function. The choice of α determines the impact of the prior on $p\left(\boldsymbol{T}\ \boldsymbol{D}\right)$, while $S(\boldsymbol{T})$ determines which additional information is used to constrain the parameter space.

Having chosen the prior, the posterior probability density $p\left(\boldsymbol{T}\ \boldsymbol{D}\right)$ is determined by sampling the N_t dimensional parameter space, and evaluating for each point $\mathcal{L}\left(\boldsymbol{D}\ \boldsymbol{T}\right)$ and $\pi\left(\boldsymbol{T}\right)$, thus performing a numerical integration. Combining this set of points with the weight given by $\mathcal{L}\left(\boldsymbol{D}\ \boldsymbol{T}\ \mathcal{M}\right) \cdot \pi$, one can determine not only the posterior probability density distribution for each bin of the spectrum, but also the posterior probability density distribution for any quantity that is computed from the spectrum, such as A_C. The mean and RMS of the posterior probability density for A_C are taken as central value and statistical uncertainty respectively. As an example, the unfolding input, the response matrix, and posterior probability densities for the $\Delta\ y$ distribution and the corresponding A_C are illustrated in Fig. 8.1.

8.1.3 Binning choice and bias

The choice of the binning for the $\Delta\ y$ distribution is driven by two factors:

- The number of parameters to estimate (the $\Delta\ y$ bin yields) affects their variance; with fewer bins, the relative statistical errors on the bin contents are reduced. Therefore the

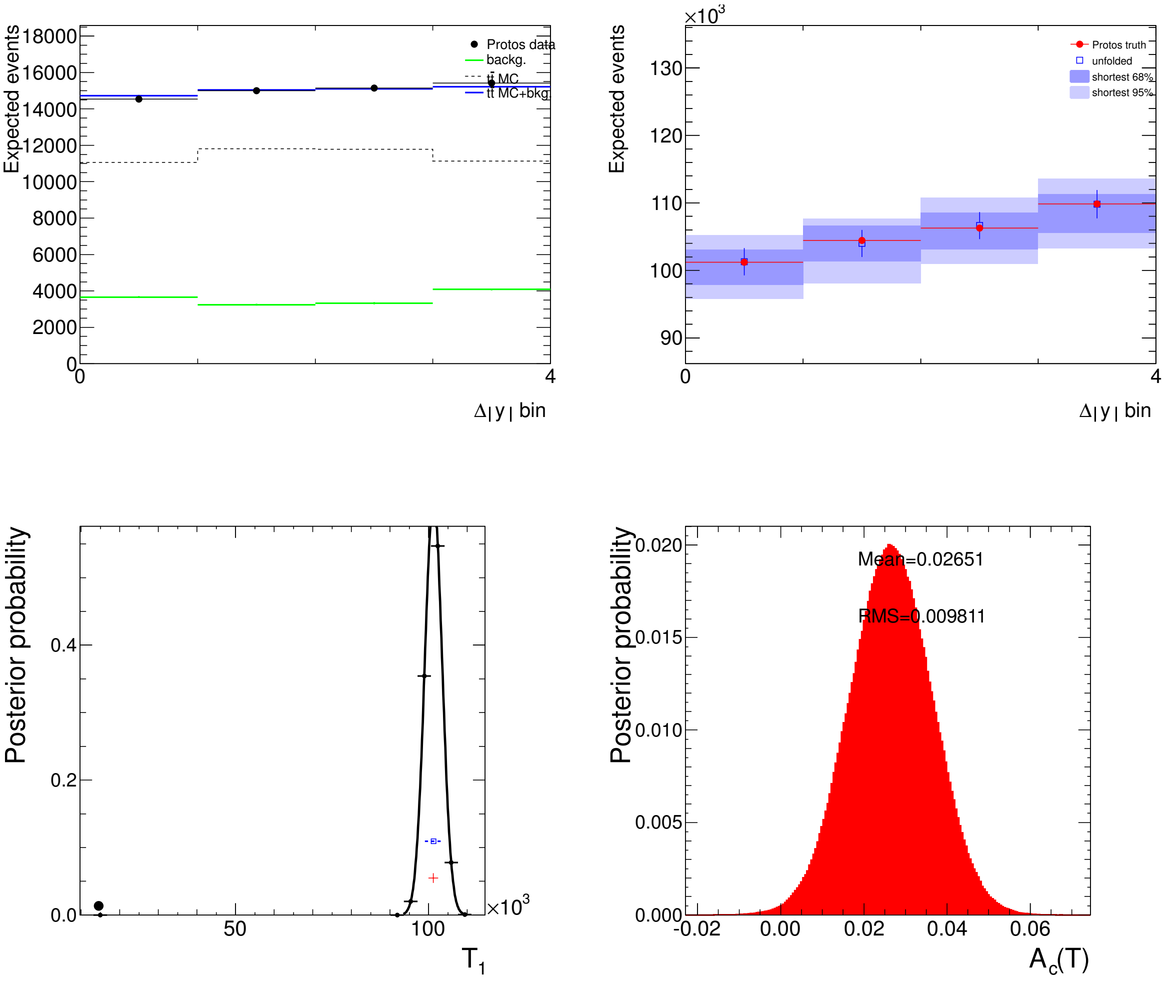


Figure 8.1: Unfolding input and output (top), and posterior distributions (bottom). In this figure we use a pseudo-data sample with a known positive asymmetry. The inputs (top-left) are the expected Δy distributions at the reconstruction level for the backgrounds. The FBU output (top-right) is the estimated spectrum at the parton level (blue bands and markers), that is computed unfolding the `PROTOS` reconstructed Δy distribution with the efficiency and transfer matrix estimated from the `ALPGEN` sample. The content of each bin is estimated from the bin posterior, shown for example in the bottom-left plot for the first bin (the full circle is the reconstructed value, the black line is the bin posterior, the red cross is the true bin content, and the blue empty square is the mean value of the posterior). From the posterior of each bin, the A_C posterior (bottom-right) can be computed. In this case we measure an unfolded asymmmetry of 2 7 ± 1 0 from a 2 5 injected asymmetry.

resulting statistical error on A_C is smaller with fewer bins. At least two bins are necessary to compute A_C (positive and negative side of the Δy distribution).

- A large number of bins allows an accurate mapping of the migrations, yielding unbiased estimates for each bin content. However, only migrations that change the Δy sign affect the computation of A_C. Such migrations are more likely for small Δy values, therefore a fine binning of the central Δy region ensures an unbiased measurement of A_C.

The benchmarks driving the choice of the binning are thus the statistical error and the bias on A_C.

The statistical error on A_C is validated by performing unfolding in an ensemble of pseudo-experiments where statistically independent pseudo-data distributions are generated based on Poisson statistics. The distribution of $({A_C}^{unf} - {A_C}^{true})\ \sigma^{unf}$ (*pull*) in the ensemble is considered, where ${A_C}^{unf}$ and σ^{unf} are the unfolded asymmetry value and its uncertainty, while ${A_C}^{true}$ is the parton level asymmetry of the sample used to generate the pseudo-data. As shown in Fig. 8.2, the RMS of the pull distribution is ≈ 1, indicating that the uncertainty is correctly estimated.

The bias in the unfolding response is measured by studying the unfolded asymmetry in pseudo-data samples for which the true asymmetry is known. The asymmetric samples are built by reweighing the baseline $t\bar{t}$ simulation to the parton level $\Delta\ y$ spectrum of BSM axi gluon samples (see Sec. 2) corresponding to $\pm2\%$, $\pm4\%$ and $\pm6\%$ asymmetries. In order to minimize the effect of statistical fluctuations, the unfolding procedure for each reference point is repeated in N_{PE} pseudo-experiments. For each reference point with true asymmetry ${A_C}^{true}$, the mean of the unfolded values ${A_C}^{unf}$, with error $\sigma\ \ \overline{N_{PE}}$, is then obtained. The set of $({A_C}^{true}, {A_C}^{unf})$ pairs is interpolated with a straight line parametrized as:

$$ {A_C}^{unf} = a \cdot {A_C}^{true} + b \tag{8.6} $$

where a and b are the *slope* and *offset* parameters. The unfolding response is considered unbiased when the distance $\ 1 - a\ $ is much smaller than the relative statistical error and the offset b is much smaller than the absolute statistical error. For both 2011 and 2012 measurements, four is the minimum number of bins which allows an unbiased response. Figure 8.2 shows the linearity test for the inclusive A_C measurement at $\ \overline{s} = 8$ TeV.

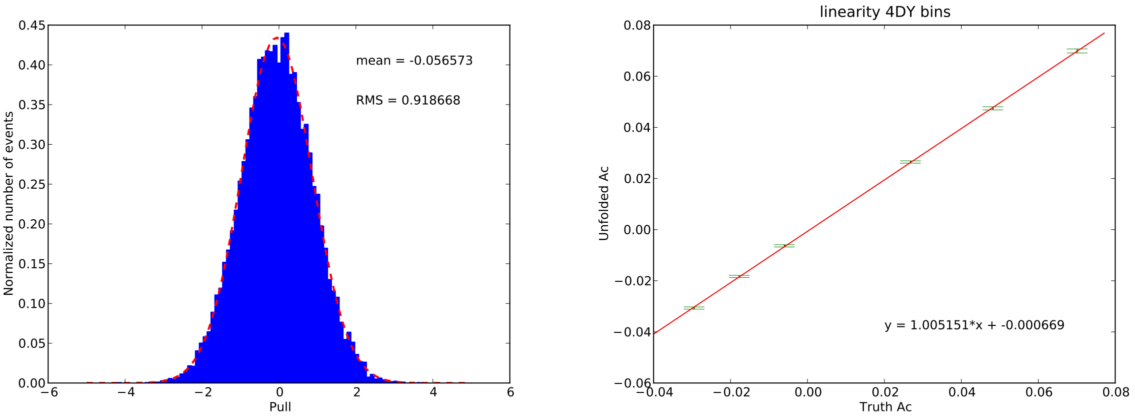


Figure 8.2: Pull distribution (left) and linearity test (right) for the inclusive A_C measurement at $\ \overline{s} = 8$ TeV

8.2 Marginalization

The treatment of systematic uncertainties is naturally included in the Bayesian inference approach by extending the likelihood $\mathcal{L}(\boldsymbol{D}\ \boldsymbol{T})$ with nuisance parameters terms. The *marginal* likelihood is defined as

$$\mathcal{L}\,(\boldsymbol{D}\;\boldsymbol{T}) = \int \mathcal{L}\,(\boldsymbol{D}\;\boldsymbol{T}\;\;\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{8.7}$$

where $\boldsymbol{\theta}$ are the nuisance parameters, and $\pi(\boldsymbol{\theta})$ their prior probability densities, which are assumed to be Normal distributions.

A nuisance parameter is associated to each of the uncertainty sources listed in Sec. **??**. Two categories of nuisances are considered: the normalizations of the background processes ($\boldsymbol{\eta}$), and the uncertainties associated to the objects identification, reconstruction and calibration ($\boldsymbol{\theta}$). While the first ones only affect the background predictions, the latter, referred to as object systematics, affect both the reconstructed distribution $\boldsymbol{R}(\boldsymbol{T};\boldsymbol{\theta})$ and the backgrounds $\boldsymbol{B}(\boldsymbol{\theta}\;\;\boldsymbol{\eta})$. The marginal likelihood becomes then

$$\mathcal{L}\,(\boldsymbol{D}\;\boldsymbol{T}) = \int \mathcal{L}\,(\boldsymbol{D}\;\boldsymbol{R}(\boldsymbol{T};\boldsymbol{\theta})\;\;\boldsymbol{B}(\boldsymbol{\eta})) \cdot Normal(\boldsymbol{\theta}) Normal(\boldsymbol{\eta}) d\boldsymbol{\theta} d\boldsymbol{\eta} \tag{8.8}$$

The marginal posterior probability density for $\boldsymbol{T}$ is computed by sampling the $N_t + N_{np}$ parameter space, where N_{np} is the total number of nuisance parameters, and projecting the sample over the $\boldsymbol{T}$ parameter space. The projections over each nuisance parameter gives the corresponding posterior probability density, which matches the Normal prior for unconstrained nuisance parameters, while it has a narrower shape for nuisance parameters that can be measured in the dataset (see Fig. 8.3). The posterior for A_C is computed as described in Sec. 8.1.2 with the difference that the RMS of the marginal posterior represents the total uncertainty.

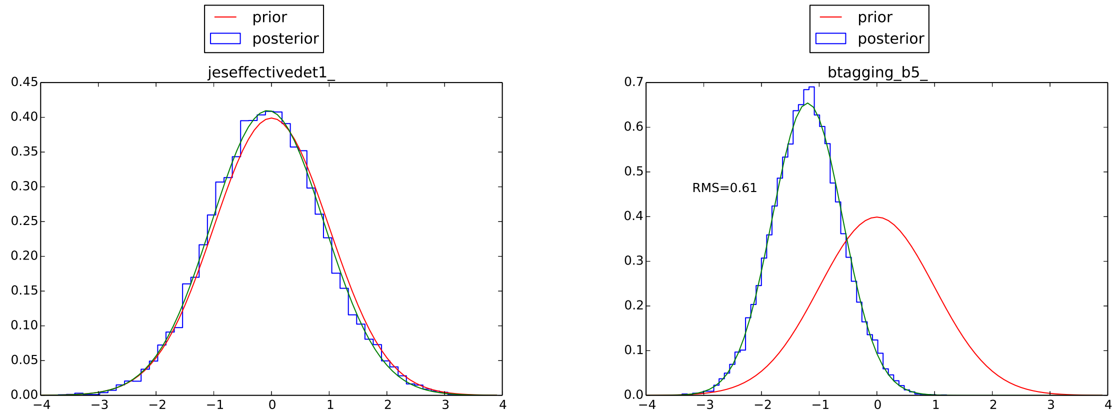


Figure 8.3: Prior and posterior distributions for nuisance parameters corresponding to a component of the JES uncertainty (left) and of the b-tagging efficiency (right) in the $\overline{s} = 8$ TeV measurement (see Sec. **??**).

8.2.1 Channel combination

As discussed in Sec. **??**, the combination of orthogonal channels with different background compositions is crucial to estimate precisely the $W+$jets contamination in the $\overline{s} = 8$ TeV data sample. The marginalization approach provides a natural framework to treat simultaneously unfolding and background estimation using multiple data regions. Given the distributions $\boldsymbol{D}_i$ measured in N_{ch} independent channels, the likelihood definition 8.8 can be extended to the product

$$\mathcal{L}\left(\;\boldsymbol{D}_1 \cdots \boldsymbol{D}_{N_{ch}} \quad \boldsymbol{T}\right) = \int \prod_{i=1}^{N_{ch}} \mathcal{L}\left(\boldsymbol{D}_i \; \boldsymbol{T};\boldsymbol{\theta}\right) \cdot Normal(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{8.9}$$

where the nuisance parameters are evaluated simultaneously for all the factors.

Chapter 9

Results

This section presents the unfolded A_C measurements at $\overline{s} = 7$ TeV and eighttev, both inclusive and differential in bins of $m_{t\bar{t}}$, $p_{t\bar{t}}$ and $y_{t\bar{t}}$. The results from data at $\overline{s} = 7$ TeV have been published on the Journal of High Energy Physics [84], while the $\overline{s} = 8$ TeV ones are internal results reported here for the purpose of this thesis. The publication results are being finalized and are expected before the end of 2014.

9.1 A_C measurements at $\overline{s} = 7$ TeV

The measurements are performed in the $+ \geq 4$ jets sample with at least 1 b jet. The observed $\Delta\ y$ distribution in data and the background predictions used as input for the inclusive measurement is shown in Fig. 9.4.

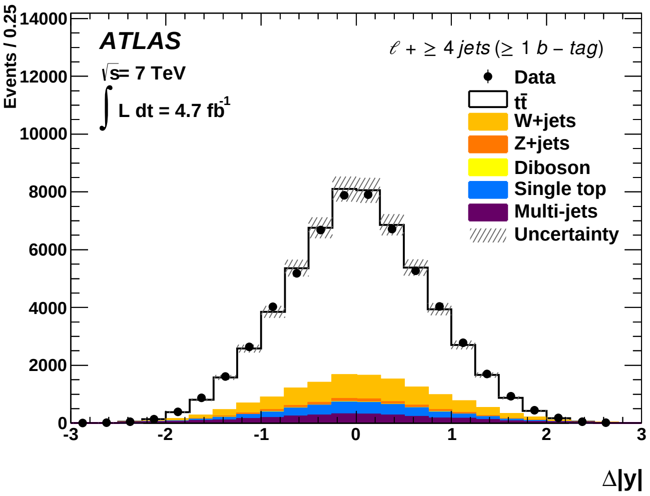


Figure 9.1: Observed $\Delta\ y$ distributions for the $+ \geq 4$ jets sample with at least 1 b jet. Data (dots) and SM expectations (solid lines) are shown. The uncertainty on the total prediction includes both the statistical and the systematic components.

The asymmetry A_C in $t\bar{t}$ production at $\overline{s} = 7$ TeV is measured to be $A_C = 0\ 006 \pm 0\ 010$, compatible with the SM prediction $A_C = 0\ 0123 \pm 0\ 0005$ (see Sec. 2.3). Table 9.1 summarizes the measurements and predictions for the inclusive asymmetry, at high invariant mass of the top quark pair ($m_{t\bar{t}} > 600$ GeV), and at high z component of the velocity of the $t\bar{t}$ system ($\beta_{z,t\bar{t}} > 0.6$). The quoted uncertainty corresponds to the RMS of the marginal posterior probability density $p(A_C|\boldsymbol{D})$ described in Sec. 8.2; therefore it accounts for both statistical and systematic components.

	Data	Theory
A_C	0.006 ± 0.010	0.0123 ± 0.0005
$A_C(m_{t\bar{t}} > 600$ GeV$)$	0.018 ± 0.022	$0.0175^{+0.0005}_{-0.0004}$
$A_C(\beta_{z,t\bar{t}} > 0.6)$	0.011 ± 0.018	$0.020^{+0.006}_{-0.007}$

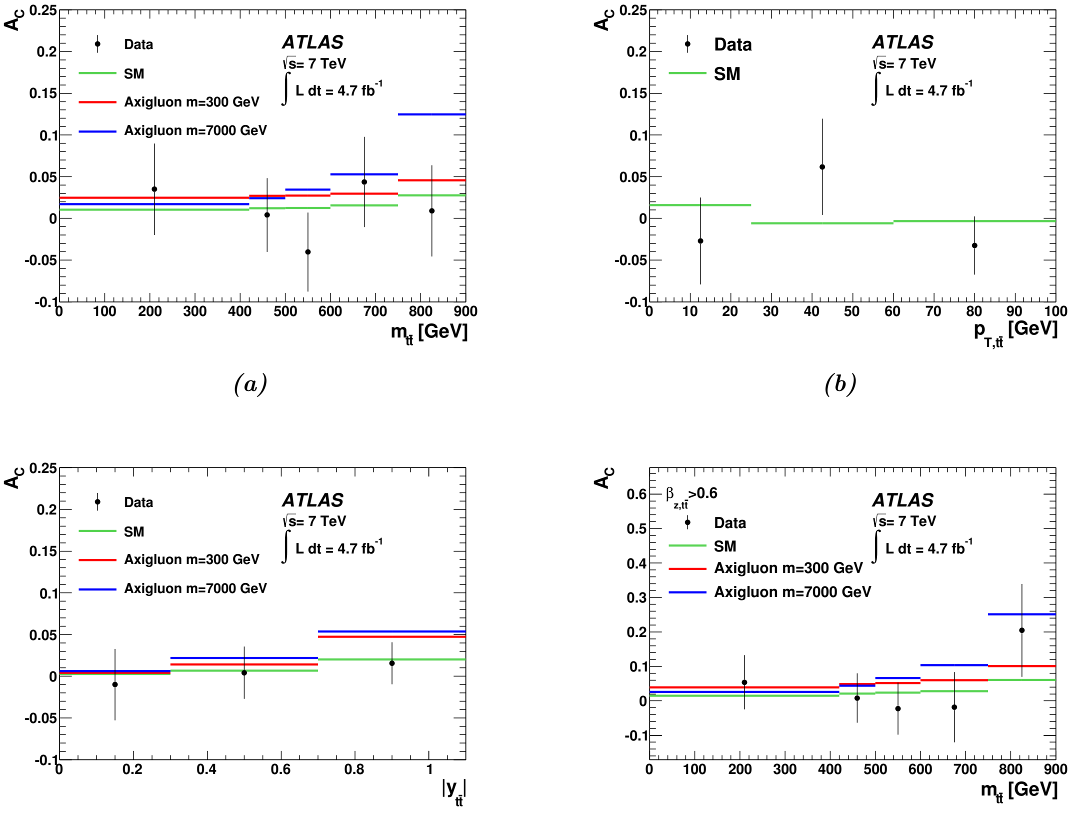
Table 9.1: Measured charge asymmetry A_C compared with SM predictions. The measurements with the $q\bar{q}$-enhancing $\beta_{z,t\bar{t}} > 0.6$ requirement and at high $t\bar{t}$ invariant mass $m_{t\bar{t}} > 600$ GeV are also included.

In order to study the impact on A_C of each systematic uncertainty, the unfolding procedure is performed for each variation without using marginalization. For each source of uncertainty, the background templates and response matrix corresponding to $\pm1\sigma$ variation are considered. Table 9.2 shows the average asymmetry variations in the asymmetry computed, for each source of uncertainty, as $|A_C(+1\sigma) - A_C(-1\sigma)|/2$. While the precision is dominated by the statistical uncertainty, the sources of largest systematic uncertainties are the ones with a large impact on the size of the $W+$jets background, such as the uncertainty on the energy scale and resolution of lepton and jets.

Source of systematic uncertainty	δA_C		
	Inclusive	$m_{t\bar{t}} > 600$ GeV	$\beta_{z,t\bar{t}} > 0.6$
Lepton reconstruction/identification	< 0.001	< 0.002	< 0.002
Lepton energy scale and resolution	0.003	0.003	0.003
Jet energy scale and resolution	0.003	0.003	0.005
Missing transverse momentum and pile-up modeling	0.002	0.002	0.004
Multi-jets background normalisation	< 0.001	< 0.002	< 0.002
b-tagging/mis-tag efficiency	< 0.001	< 0.002	< 0.002
Signal modeling	< 0.001	< 0.002	< 0.002
Parton shower/hadronization	< 0.001	< 0.002	< 0.002
Monte Carlo statistics	0.002	< 0.002	< 0.002
PDF	0.001	< 0.002	< 0.002
$W+$jets normalisation and shape	0.002	< 0.002	< 0.002
Statistical uncertainty	0.010	0.021	0.017

Table 9.2: Impact of individual sources of uncertainty on the measured A_C and $A_C(m_{t\bar{t}} > 600$ GeV$)$. Variations below 10% of the statistical uncertainty are quoted as < 0.001 and < 0.002 are considered negligible.

The asymmetry spectra, measured as functions of $m_{t\bar{t}}$, $p_{t\bar{t}}$ and $y_{t\bar{t}}$ are compared in Fig. 9.2 with the theoretical SM predictions and found to be compatible. In addition, the results are compared with new physics predictions for color octets (*axigluons*) with masses $M_G = 300$ GeV and $M_G = 7$ TeV [85]. Both models would not be observable in direct searches as $t\bar{t}$ resonances: the light axigluon mass is below the threshold for production of top quark pairs, while a 7 TeV mass is beyond the kinematic reach of the current searches at $\overline{s} = 7$ TeV and $\overline{s} = 8$ TeV. In both scenarios the new physics couplings with top and light quarks are tuned to yield a forward backward asymmetry compatible with the results reported by Tevatron experiments.



(a) *(b)*

(c) *(d)*

Figure 9.2: Measured A_C spectra compared with predictions for SM and for a color octet axigluon with a mass of 300 GeV (red lines) and 7000 GeV (blue lines) as functions of $m_{t\bar{t}}$ (a), $p_{t\bar{t}}$ (b) and $y_{t\bar{t}}$ (c). The asymmetry as a function of $m_{t\bar{t}}$ with $\beta_{z,t\bar{t}} > 0\ 6$ is also shown (d).

The asymmetry values corresponding to the spectra in Fig. 9.2 are detailed in App. ??. The correlation of the asymmetry values $A_C{}^i$ across differential bins is fully described by the posterior probability density $p\left(A_C{}^1 \quad A_C{}^N\ \boldsymbol{D}\right)$, where N is the number of differential bins, and it is summarized in App. ??. The A_C spectrum as a function of the $t\bar{t}$ invariant mass $m_{t\bar{t}}$ with the additional requirement on the z component of high velocity of the $t\bar{t}$ system, $\beta_{z,t\bar{t}} > 0\ 6$, is also shown in Fig. 9.2. The enhancement of the $q\bar{q} \quad t\bar{t}$ process results in larger asymmetries for both SM and new physics predictions.

A comparison of the inclusive A_C measurements, together with the Tevatron A_{FB} ones, with predictions from a broad range of new physics model (see Sec. 2.3.2) is shown in Fig. 9.3. The A_C and A_{FB} measurements, inclusive and at high $m_{t\bar{t}}$ mass, define a region where the new physics prediction are compatible with the current measurements from both Tevatron and LHC experiments. Some models show only a limited range of parameter values which yield acceptable asymmetries.

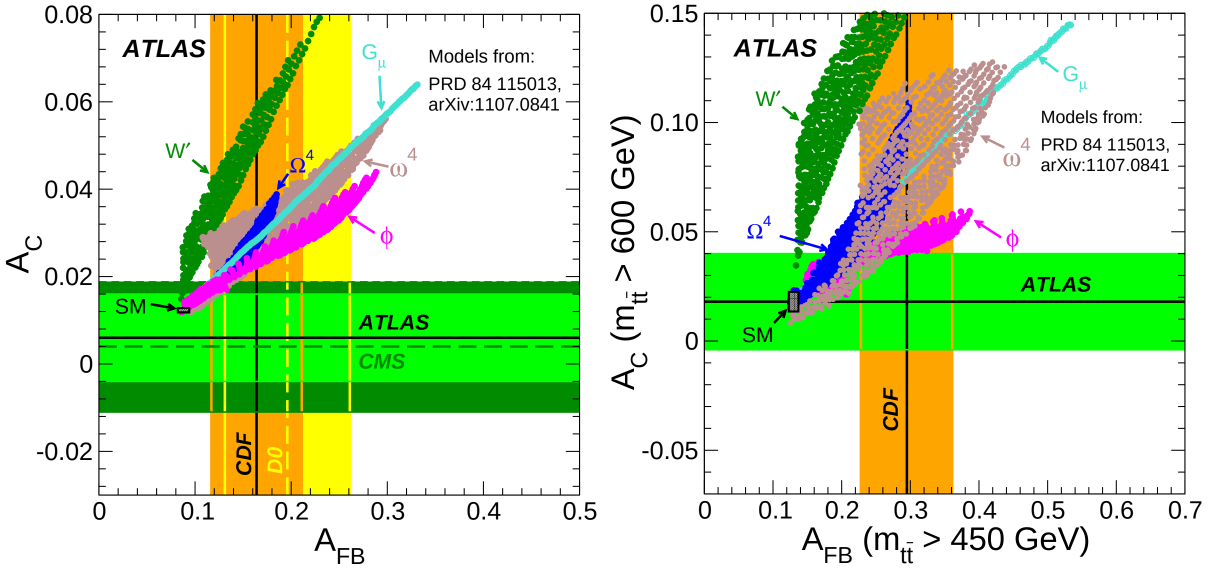


Figure 9.3: Measured forward backward asymmetries A_{FB} at Tevatron and charge asymmetries A_C at LHC, compared with the SM predictions (black box) and values in allowed new physics scenarios. The horizontal bands and lines correspond to the ATLAS (light green) and CMS (dark green) measurements, while the vertical ones correspond to the CDF (orange) and D0 (yellow) measurements. The inclusive asymmetry measurements are shown in the left plot. The right plot shows a comparison with the A_{FB} measurement by CDF for $m_{t\bar{t}} > 450$ GeV and the ATLAS A_C measurement for $m_{t\bar{t}} > 600$ GeV.

9.2 A_C measurements at $\overline{s} = 8$ TeV

The measurements at $\overline{s} = 8$ TeV are performed using six sub-samples of $+ \geq 4$ jets events, where, simultaneously, the background normalizations are fitted and the parton level $\Delta\ y$ distribution is estimated. Fig. 9.4 shows the fitted distribution used for the inclusive A_C measurement compared with the observed one.

The asymmetry A_C in $t\bar{t}$ production at $\overline{s} = 8$ TeV is measured to be $A_C = 0\ 011 \pm 0\ 005$, compatible with the SM prediction $A_C = 0\ 0111 \pm 0\ 0004$.

Since the background estimation is part of the bayesian inference procedure described in Sec. 8, it s not possible to study the impact of systematic uncertainties by repeating unfolding on data with varied templates, without using marginalization. Instead, the expected impact of systematic uncertainties is studied with pseudo-data distributions corresponding to the sum of the background and signal predictions. For each source of uncertainty, the $\pm 1\sigma$ variations of the predictions are used to build the pseudo data, and the unfolding procedure is repeated. The baseline background templates and response matrices, as in the actual measurements, are used. Table 9.3 shows the average asymmetry variation δA_C computed, for each source of uncertainty, as $A_C(+1\sigma) - A_C(-1\sigma)\ \ 2$. As in the measurements at $\overline{s} = 7$ TeV, the precision is dominated by the statistical uncertainty, and the main sources of systematic uncertainty are the ones with

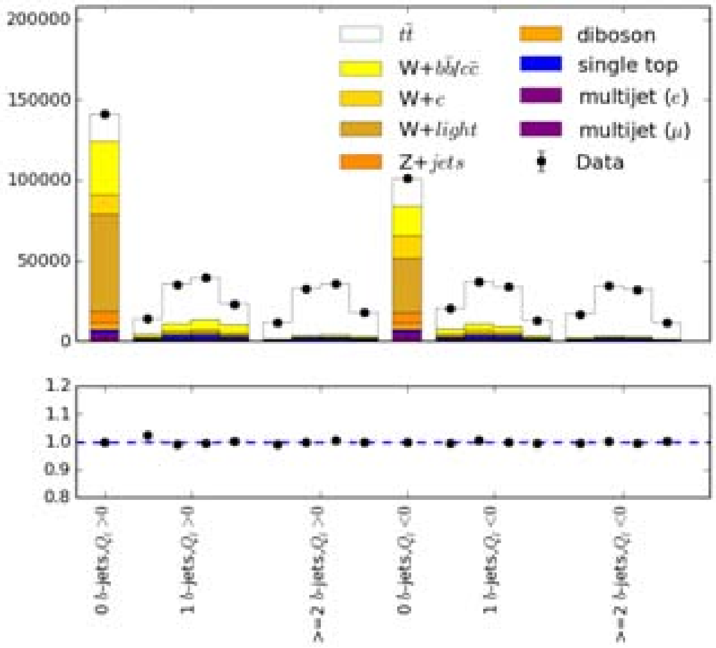


Figure 9.4

a large impact on the size of the $W+$jets background, such as the uncertainty on the energy scale and resolution of lepton and jets.

Source of systematic uncertainty	δA_C
Lepton reconstruction/identification	< 0.0005
Lepton energy scale and resolution	0.0007
Jet reconstruction efficiency	< 0.0005
Jet energy scale and resolution	0.0017
Missing transverse momentum and pile–up modeling	< 0.0005
Multi–jets background normalization	< 0.0005
Other backgrounds normalization	< 0.0005
b–tagging/mis–tag efficiency	< 0.0005
Signal modeling	0.004
Parton shower/hadronization	0.003
Monte Carlo statistics	0.0008
PDF	0.0006
Unfolding response	< 0.0005
Statistical uncertainty	0.005
Total uncertainty	0.005

Table 9.3: Expected impact of individual sources of uncertainty on the measured A_C. Variations below 10% of the statistical uncertainty are quoted as < 0.0005 and considered negligible.

The asymmetry spectrum as a function of $m_{t\bar{t}}$ is compared in Fig. 9.5 with the theoretical SM predictions and found to be compatible.

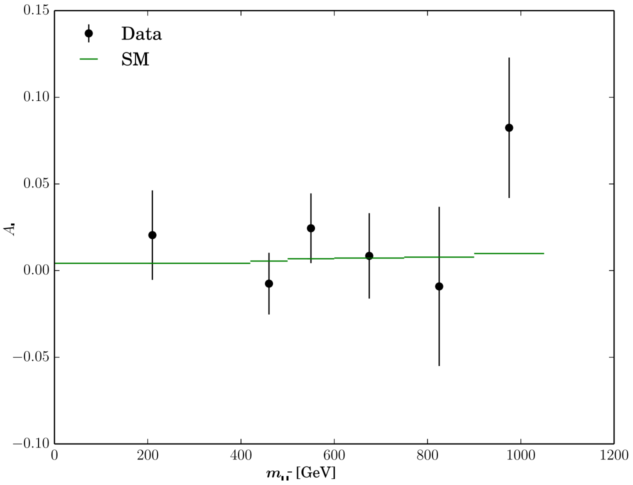


Figure 9.5: Measured A_C spectrum as a function of $m_{t\bar{t}}$ compared with SM prediction.

Chapter 10

Conclusion

Measurements of the asymmetry in the production of top quark pairs at the ATLAS experiment have been performed, using a dataset corresponding to an integrated luminosity of 4 7 fb^{-1} taken over the course of 2011 at a centre-of-mass energy $\overline{s} = 7$ TeV and a dataset corresponding to an integrated luminosity of 20 fb^{-1} taken over the course of 2012 at a centre-of-mass energy $\overline{s} = 8$ TeV.

The lepton+jets signature of the semileptonic $t\bar{t}$ decay has been exploited to select samples enriched in $t\bar{t}$ events by identifying and reconstructing one isolated lepton with large transverse momentum, at least four jets and a large missing transverse momentum. The background contamination of the sample has been studied and data-driven techniques have been developed to precisely determine its composition. In particular, the normalization of the different W+heavy flavor processes has been determined in situ exploiting the intrinsic charge asymmetry in W production at LHC. Other background processes, such as single top, Z+jets and diboson production have been estimated with Monte Carlo simulations, while QCD multijet events have been modeled with well established data-driven techniques. The four momenta of the $t\bar{t}$ pair has been reconstructed event by event with a kinematic fit of the measured objects. The inclusive and differential distribution of the difference of absolute rapidities of the reconstructed top and antitop quarks, $\Delta\ y$, has then been measured in data and estimated for the backgrounds. A novel unfolding procedure, based on a bayesian inference approach, has been performed to estimate the inclusive and differential $\Delta\ y$ distribution, and the corresponding asymmetries, at the parton level, accounting for the distortions induced by acceptance and resolution effects. The procedure has been calibrated with the goal of obtaining the most precise and accurate achievable asymmetry measurement, by studying the response in distributions for which the corresponding parton level asymmetry is known.

In the $\overline{s} = 7$ TeV dataset, the asymmetry A_C has been measured inclusively, as a function of $m_{t\bar{t}}$, $p_{t\bar{t}}$ and $y_{t\bar{t}}$, and has been found to be compatible with the SM predictions in all cases. The inclusive measurement and the one as a function of $m_{t\bar{t}}$ have also been performed in a $q\bar{q} \quad t\bar{t}$ enriched sample ($\beta_{z,t\bar{t}} > 0\ 6$), yielding SM compatible results as well. The precision of both inclusive and differential measurements is limited by the statistical uncertainty, while the largest systematic uncertainties are the ones with a large impact on the acceptance for asymmetric backgrounds ($W+$jets and single top), such as the uncertainty on the energy scale and resolution of jets and leptons. In the $\overline{s} = 8$ TeV dataset, the asymmetry A_C has been measured inclusively and as a function of $m_{t\bar{t}}$; in both case the results are compatible with the SM predictions. The set of measurements will be completed to match the $\overline{s} = 7$ TeV set, with the differential measurement as a function of $p_{t\bar{t}}$ and measurements in the $q\bar{q}$ enriched sample with $\beta_{z,t\bar{t}} > 0\ 6$.

At this stage the asymmetry measurements at LHC do not show significant deviations from the SM predictions. Due to the different initial state and subsequently different asymmetry definitions, a direct comparison with the anomaly reported at Tevatron cannot be performed. One important step would be to simultaneously measure the dependence of the asymmetries on the $t\bar{t}$ invariant mass $m_{t\bar{t}}$ and velocity $\beta_{z,t\bar{t}}$, in order to extrapolate the collider-independent asymmetries A_u and A_d [86]. This measurement is quite demanding from the experimental side, since it requires a 3 dimensional unfolding in $m_{t\bar{t}}$, $\beta_{z,t\bar{t}}$ and $\Delta\ y$. But it offers a unique possibility of testing at LHC the same quantities that are at the origin of the Tevatron A_{FB}.

Measuring the asymmetry A_C as an effective probe for new physics in future LHC datasets will become more challenging. Proton collisions will be delivered at higher center-of-mass energy 13 TeV in 2015 thus making the fraction of $q\bar{q}$ $t\bar{t}$ events even smaller. Beside enhancing the $q\bar{q}$ fraction by requiring a large z component of the $t\bar{t}$ velocity, a promising approach is to study the asymmetry in $t\bar{t}$ events produced in association with photons [87].

Bibliography

[1] O. U. Press, Oxford Dictionaries, **2014**.

[2] R. Elia et al., *Mod.Phys.Lett.* **1993**, *A8*, 2237 2248.

[3] W. Braunschweig et al., *Z.Phys.* **1988**, *C40*, 163 170.

[4] G. Arnison et al., *Phys.Lett.* **1983**, *B122*, 103 116.

[5] M. Banner et al., *Phys.Lett.* **1983**, *B122*, 476 485.

[6] V. M. Abazov et al., **2014**.

[7] T. Aaltonen et al., *Phys.Rev.* **2013**, *D87*, 092002.

[8] F. Abe et al., *Phys.Rev.Lett.* **1995**, *74*, 2626 2631.

[9] S. Abachi et al., *Phys.Rev.Lett.* **1995**, *74*, 2422 2426.

[10] L. Fiorini, **2012**.

[11] J. C. Collins, D. E. Soper, G. F. Sterman, *Adv.Ser.Direct.High Energy Phys.* **1988**, *5*, 1 91.

[12] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Phys.Rev.* **2008**, *D78*, 013004.

[13] A. Sherstnev, R. Thorne, *Eur.Phys.J.* **2008**, *C55*, 553 575.

[14] A. Martin, W. Stirling, R. Thorne, G. Watt, *Eur.Phys.J.* **2009**, *C64*, 653 680.

[15] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *Phys.Rev.* **2010**, *D82*, 074024.

[16] R. D. Ball et al., *Nucl.Phys.* **2009**, *B823*, 195 233.

[17] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, P. Nason, *Phys.Lett.* **2012**, *B710*, 612 622.

[18] M. Beneke, P. Falgari, S. Klein, C. Schwinn, *Nucl.Phys.* **2012**, *B855*, 695 741.

[19] P. B arnreuther, M. Czakon, A. Mitov, *Phys.Rev.Lett.* **2012**, *109*, 132001.

[20] M. Czakon, A. Mitov, *JHEP* **2012**, *1212*, 054.

[21] M. Czakon, A. Mitov, *JHEP* **2013**, *1301*, 080.

[22] M. Czakon, P. Fiedler, A. Mitov, *Phys.Rev.Lett.* **2013**, *110*, 252004.

[23] M. Czakon, A. Mitov, **2011**.

[24] Combination of ATLAS and CMS top-quark pair cross section measurements using up to 1.1 fb-1 of data at 7 TeV, tech. rep. ATLAS-CONF-2012-134, CERN, Geneva, **Sept. 2012**.

[25] Combination of ATLAS and CMS top-quark pair cross section measurements using proton-proton collisions at sqrt(s) = 7 TeV, tech. rep. CMS-PAS-TOP-12-003, CERN, Geneva, **2013**.

[26] G. Aad et al., **2014**.

[27] S. Chatrchyan et al., *JHEP* **2014**, *1402*, 024.

[28] **2014**.

[29] V. Abazov et al., *Phys.Rev.Lett.* **2007**, *98*, 041801.

[30] T. Aaltonen et al., *Phys.Rev.Lett.* **2010**, *105*, 101801.

[31] G. Aad et al., *JHEP* **2013**, *1311*, 031.

[32] G. Aad et al., *Phys.Rev.Lett.* **2013**, *111*, 232002.

[33] G. Aad et al., **2014**.

[34] S. Chatrchyan et al., *Phys.Rev.Lett.* **2014**, *112*, 182001.

[35] J. H. Kuhn, G. Rodrigo, *Phys.Rev.* **1999**, *D59*, 054017.

[36] J. H. Kuhn, G. Rodrigo, *JHEP* **2012**, *1201*, 063.

[37] W. Bernreuther, Z.-G. Si, *Phys.Rev.* **2012**, *D86*, 034026.

[38] J. Aguilar-Saavedra, A. Juste, F. Rubbo, *Phys.Lett.* **2012**, *B707*, 92 98.

[39] J. Aguilar-Saavedra, D. Amidei, A. Juste, M. Perez-Victoria, **2014**.

[40] J. Aguilar-Saavedra, M. Perez-Victoria, *Phys.Rev.* **2011**, *D84*, 115013.

[41] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, P. Proudlock, *LHC Design Report*, CERN, Geneva, **2004**.

[42] S. Chatrchyan et al., *JINST* **2008**, *3*, S08004.

[43] J. Alves, A. Augusto et al., *JINST* **2008**, *3*, S08005.

[44] K. Aamodt et al., *JINST* **2008**, *3*, S08002.

[45] G. Aad et al., *JINST* **2008**, *3*, S08003.

[46] G. Aad et al., *Eur.Phys.J.* **2012**, *C72*, 1849.

[47] M. Mangano, T. Stelzer, *Ann.Rev.Nucl.Part.Sci.* **2005**, *55*, 555 588.

[48] T. Sjostrand, S. Mrenna, P. Z. Skands, *JHEP* **2006**, *0605*, 026.

[49] T. Sjostrand, S. Mrenna, P. Z. Skands, *Comput.Phys.Commun.* **2008**, *178*, 852 867.

[50] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *JHEP* **2001**, *0101*, 010.

[51] J. Butterworth, J. R. Forshaw, M. Seymour, *Z.Phys.* **1996**, *C72*, 637 646.

[52] B. P. Kersevan, E. Richter-Was, *Comput.Phys.Commun.* **2013**, *184*, 919 985.

[53] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. D. Polosa, *JHEP* **2003**, *0307*, 001.

[54] S. Frixione, B. R. Webber, *JHEP* **2002**, *0206*, 029.

[55] P. Nason, *JHEP* **2004**, *0411*, 040.

[56] G. Aad et al., *Eur.Phys.J.* **2010**, *C70*, 823 874.

[57] S. Agostinelli et al., *Nucl.Instrum.Meth.* **2003**, *A506*, 250 303.

[58] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, A. Salzburger, Concepts, Design and Implementation of the ATLAS New Tracking (NEWT), tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, **Mar. 2007**.

[59] R. Fruhwirth, *Nuclear Instruments and Methods in Physics Research* **1987**, *262*, 444 450.

[60] ATLAS Collaboration, Performance of primary vertex reconstruction in proton-proton collisions at $\overline{s}$ =7 TeV in the ATLAS experiment, tech. rep. ATLAS-CONF-2010-069, CERN, Geneva, **July 2010**.

[61] ATLAS Collaboration, *Eur.Phys.J.* **2012**, *C72*, 1909.

[62] E. Abat, J. Abdallah, T. Addy, P. Adragna, M. Aharrouche, et al., *JINST* **2010**, *5*, P11006.

[63] ATLAS Collaboration, Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector, tech. rep. ATLAS-CONF-2011-063, CERN, Geneva, **Apr. 2011**.

[64] B. Acharya *et al.*, Object selection and calibration, background estimations and MC samples for the Winter 2013 Top Quark analyses with 2012 data, tech. rep. ATL-COM-PHYS-2013-088, CERN, Geneva, **Jan. 2013**.

[65] T. Barillari et al., *ATL-LARG-PUB-2009-001* **2009**, https://cds.cern.ch/record/1112035.

[66] G. P. Salam, *Eur.Phys.J.* **2010**, *C67*, 637 686.

[67] ATLAS Collaboration, *ATLAS-CONF-2012-043* **2012**, https://cdsweb.cern.ch/record/1435197.

[68] ATLAS Collaboration, *ATLAS-CONF-2012-039* **2012**, https://cdsweb.cern.ch/record/1435193.

[69] ATLAS Collaboration, *ATLAS-CONF-2012-040* **2012**, https://cdsweb.cern.ch/record/1435194.

[70] K. Melnikov, F. Petriello, *Phys.Rev.* **2006**, *D74*, 114017.

[71] N. Kidonakis, *Phys.Rev.* **2010**, *D81*, 054028.

[72] N. Kidonakis, *Phys.Rev.* **2011**, *D83*, 091503.

[73] N. Kidonakis, *Phys.Rev.* **2010**, *D82*, 054018.

[74] J. M. Campbell, R. K. Ellis, *Phys.Rev.* **1999**, *D60*, 113006.

[75] C.-H. Kom, W. J. Stirling, *Eur.Phys.J.* **2010**, *C69*, 67 73.

[76] G. Aad et al., *Eur.Phys.J.* **2011**, *C71*, 1577.

[77] G. Aad et al., *Eur.Phys.J.* **2013**, *C73*, 2518.

[78] G. Aad et al., **2014**.

[79] M. Bosman, L. Fiorini, C. Helsens, A. Juste, L. Mir, J. Nadal, V. Vorwerk, Weighting method to propagate heavy-flavor tagging calibrations and related uncertainties, tech. rep. ATL-COM-PHYS-2010-331, CERN, Geneva, **June 2010**.

[80] B. Acharya, J.-F. Arguin, M. Barisonzi, I. Besana, M. Bosman, I. Brock, R. Calkins, D. Charkaborty, M. Cobal, M. Cristinziani, J. Ferrando, L. Fiorini, A. Henrichs, B. Heinemann, C. Helsens, G. Khoriauli, T. Kuhl, A. Juste, T. Lari, F. Meloni, L. Mir, J. Nadal, E. Nuncio-Quiroz, M. Pinamonti, B. Radics, U. de Sanctis, E. Shabalina, L.-Y. Shan, K. Suruliz, D. Ta, C. Troncon, V. Vorwerk, Estimation of the W+Jets Background for Top Quark Re-Discovery in the Single Lepton+Jets Channel, tech. rep. ATL-COM-PHYS-2010-834, Supporting document for top paper, CERN, Geneva, **Oct. 2010**.

[81] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al., **2011**.

[82] J. Erdmann, S. Guindon, K. Kroeninger, B. Lemmer, O. Nackenhorst, et al., *Nucl.Instrum.Meth.* **2014**, *A748*, 18 25.

[83] G. Choudalakis, **2012**.

[84] G. Aad et al., *JHEP* **2014**, *1402*, 107.

[85] J. Aguilar-Saavedra, M. Perez-Victoria, *Phys.Lett.* **2011**, *B705*, 228 234.

[86] J. Aguilar-Saavedra, A. Juste, *Phys.Rev.Lett.* **2012**, *109*, 211804.

[87] J. Aguilar-Saavedra, E. Álvarez, A. Juste, F. Rubbo, *JHEP* **2014**, *1404*, 188.

Appendix A

Tag Rate Function

To overcome the loss of simulated events due to the b tag requirement, the tag rate function (TRF) method is introduced. By using the TRF method, no event is discarded based on its b tagging count, but instead all the events are weighted. This weight can be interpreted as the probability of the given event to contain the desired number of b jets. The tagging efficiency, parametrized as a function of η, p_{T} and true jet flavor, is used to compute the event weight.

Given a jet with η, p_{T} and flavor f, its tagging probability can be expressed as:

$$(f \quad \eta \quad p_{\mathrm{T}})$$

For a given event with N jets, its probability of containing exactly one b-tag jet is computed as:

$$P_{=1} = \sum_{i=1}^{N} \left({}_{i} \prod_{i=j} (1 - {}_{j}) \right)$$

In the same way, it can be used to compute the probability for inclusive b-tag selections:

$$P_{=0} = \prod_{i=1}^{N} (1 - {}_{j})$$
$$P_{\geq 1} = 1 - P_{=0}$$

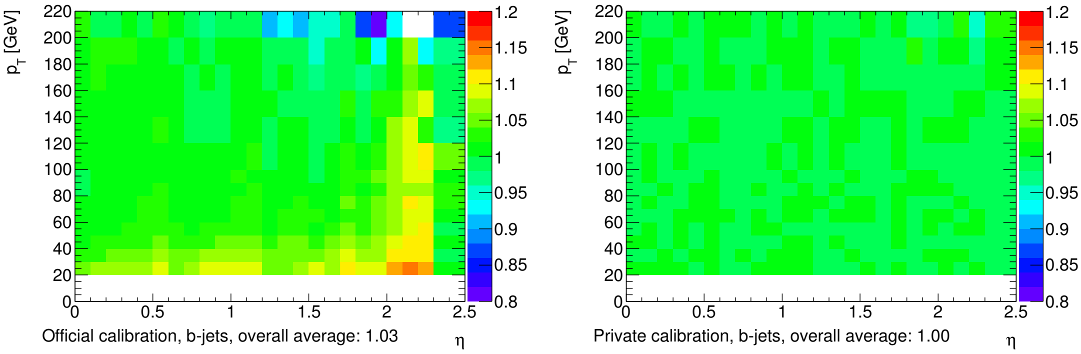
A.1 Validation

This method relies on the correct calibration of the tagging efficiency in MC samples. However, closure tests performed with the official calibration files have shown that the efficiency parametrization is not as accurate as expected. Assuming a correct calibration, the average of the histogram of 1 vs η, p_{T} and true jet flavour should be flat and with mean equal to one. Figure A.1 shows the result of this test.

As it can be observed, there is a departure from the expected behavior in the edges of the distribution that amounts to up to 20% in some regions of the light and c flavour. Using the McAtNlo sample, an efficiency map has been derived in order to overcome this problem.

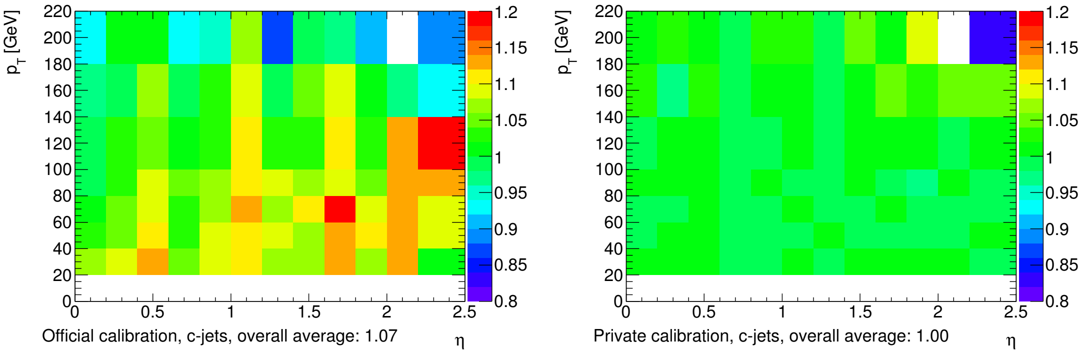
Figure A.1 shows the result of the closure test on the derived efficiency map. The derived efficiency map will therefore be used for the probability computations in the TRF method.

The validation of the method is further done by comparing in `ALPGEN` the normalisation and shape of the relevant distributions shown in Fig. A.2. As it is seen in the plots, the prediction of the TRF method is accurate up to the statistical error.



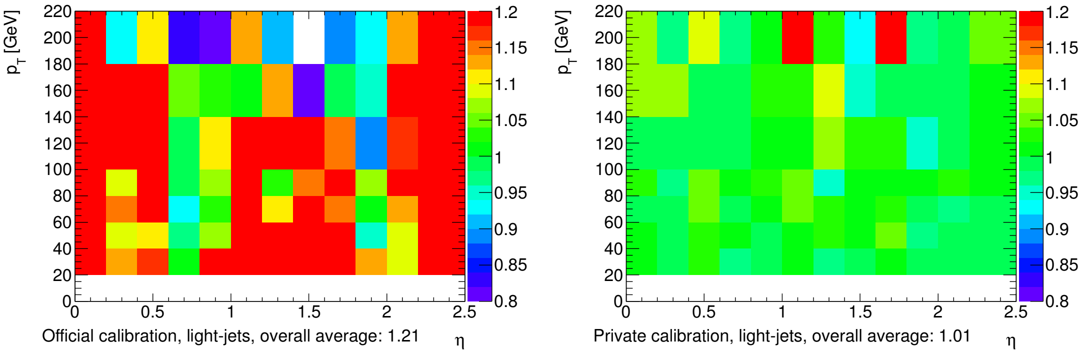
Official calibration, b-jets, overall average: 1.03

Private calibration, b-jets, overall average: 1.00



Official calibration, c-jets, overall average: 1.07

Private calibration, c-jets, overall average: 1.00



Official calibration, light-jets, overall average: 1.21

Private calibration, light-jets, overall average: 1.01

Figure A.1: Results of the closure test using efficiency from the official calibration file (left column) and the private efficiency map (right column). The test is split in the different jet flavours: b jets (top), c jets (middle) and light jets (bottom)

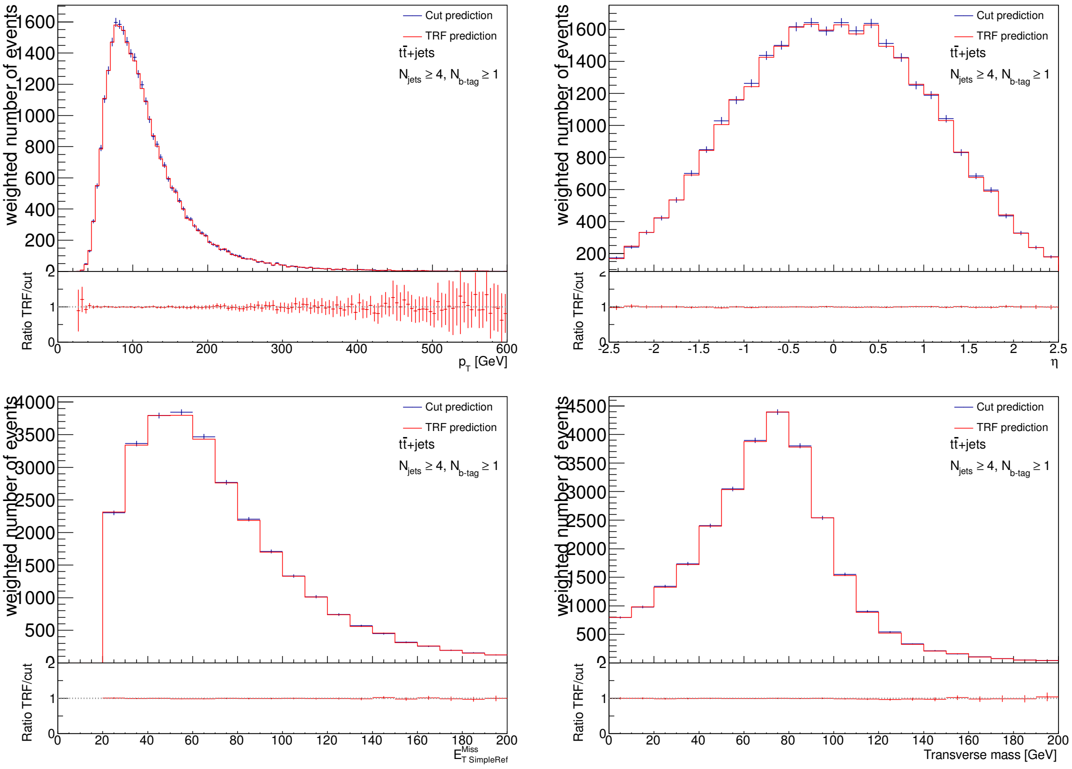


Figure A.2: Comparison between the TRF prediction and the cut based b tag prediction in `ALPGEN`. The plots show, for the $t\bar{t}$ μ+jets sample, the p_{T} (top left) and η (top right) distributions for the highest p_{T} jet, the E_{T} (bottom left) and the m_{T}^{W} (bottom right) distributions.

Appendix B

Control plots

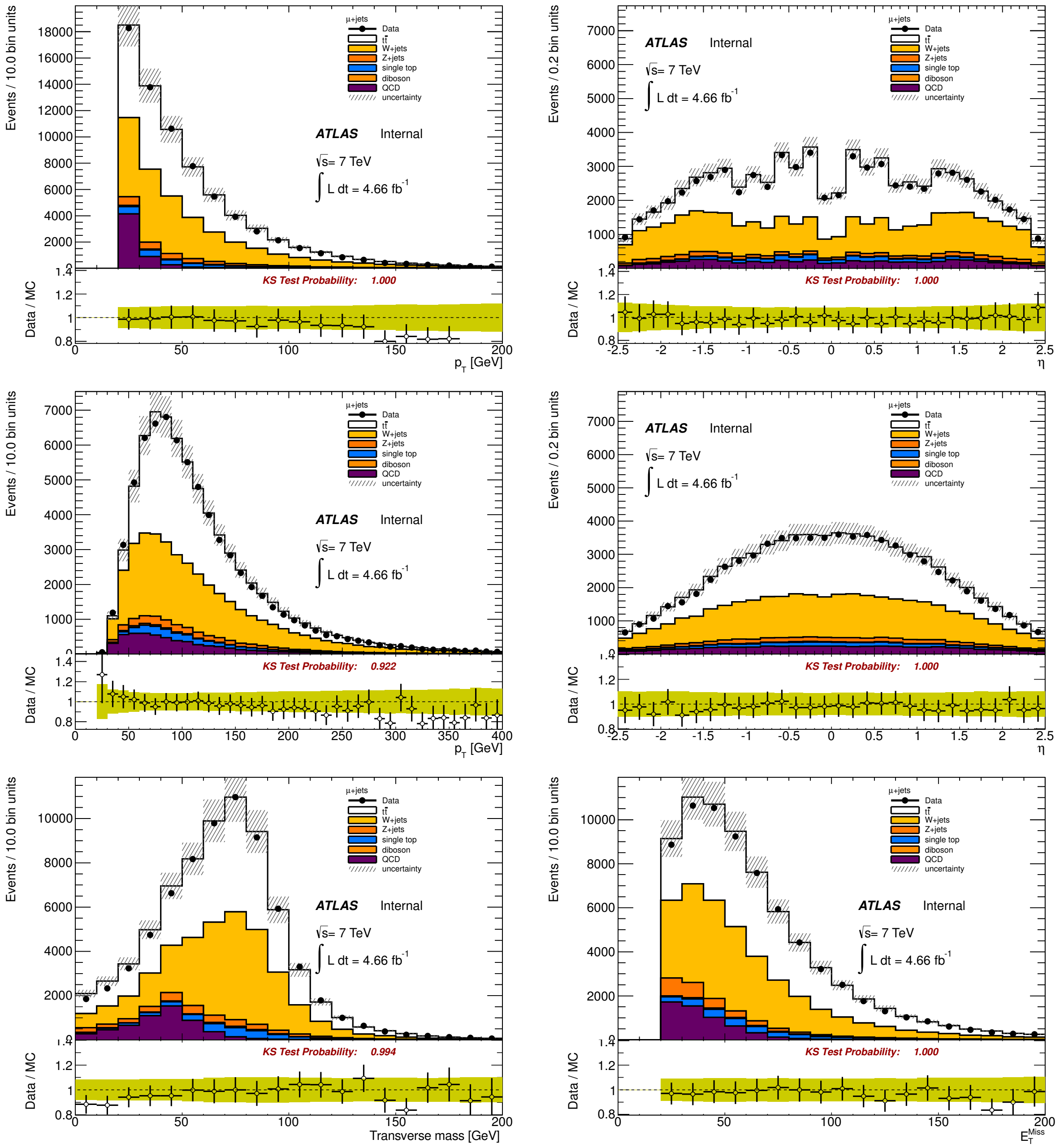


Figure B.1: Control plots for the $\mu +$ jets channel in the pretag sample at $\sqrt{s} = 7$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The distributions for the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T} are shown in the bottom row. Uncertainty on background normalization is shown.

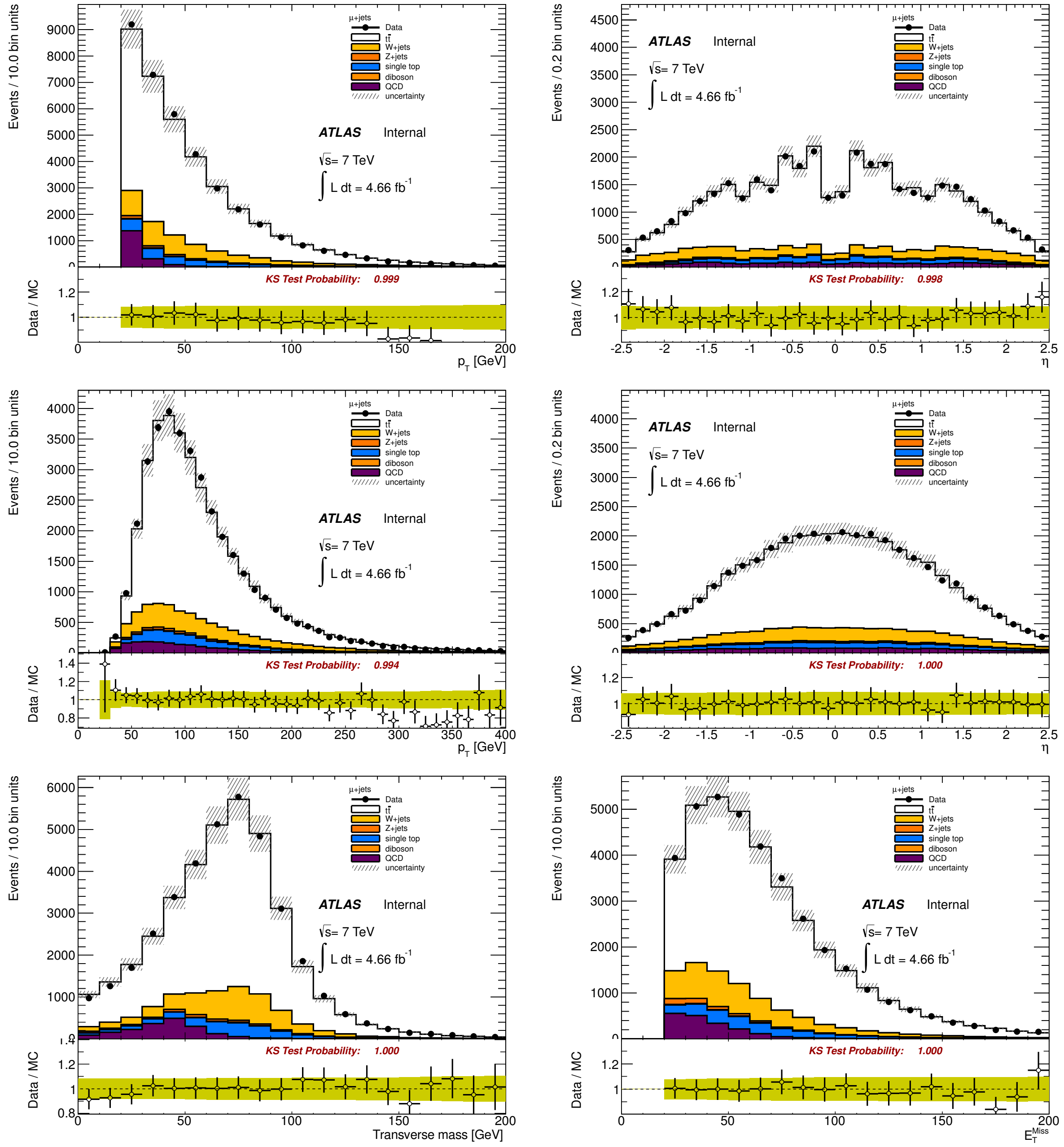


Figure B.2: Control plots for the $\mu+$jets channel in the tagged sample at $\sqrt{s}=7$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}. Uncertainty on background normalization is shown.

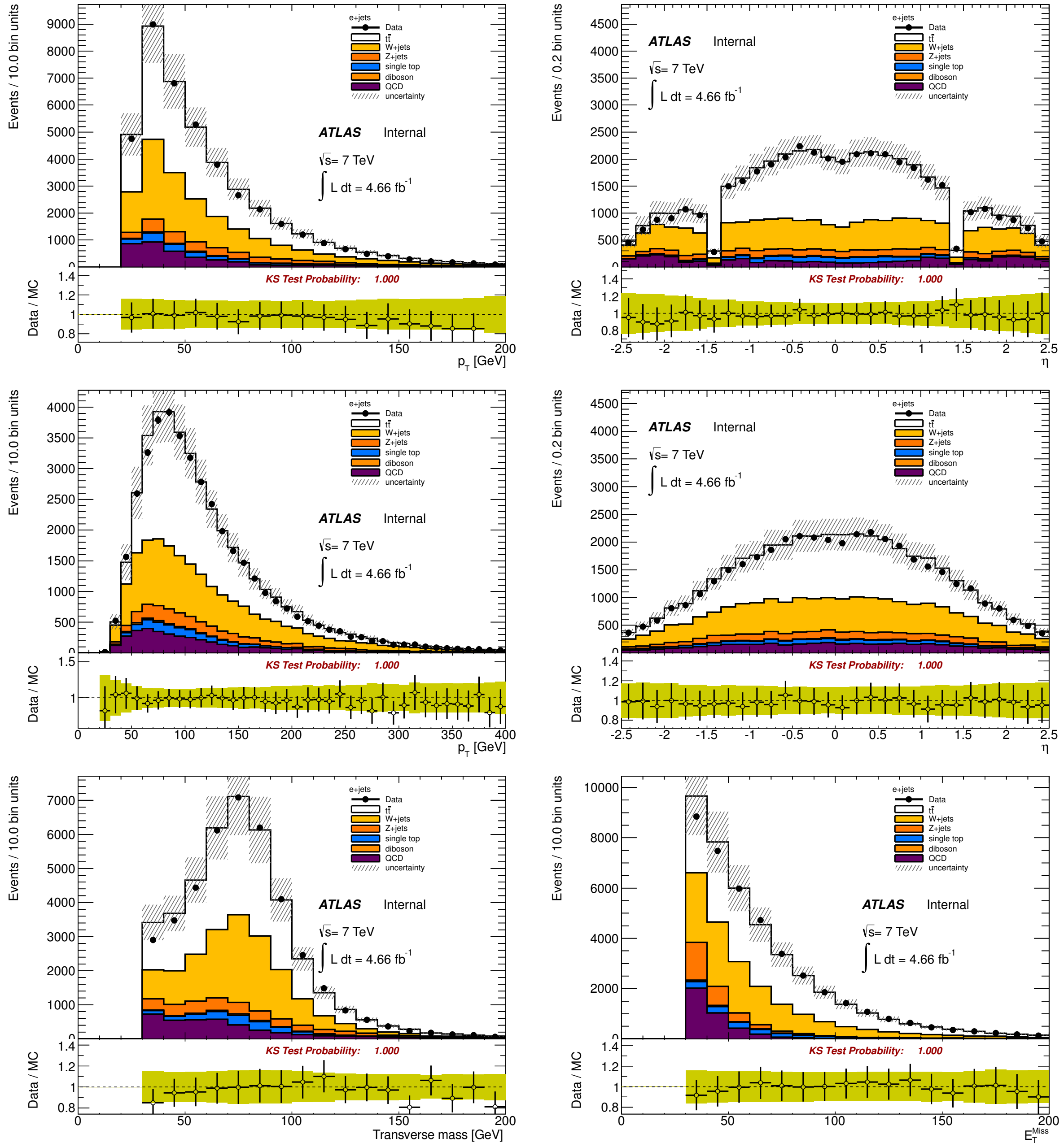


Figure B.3: Control plots for the $e+$jets channel in the pretag sample at $\sqrt{s} = 7$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}. Uncertainty on background normalization is shown.

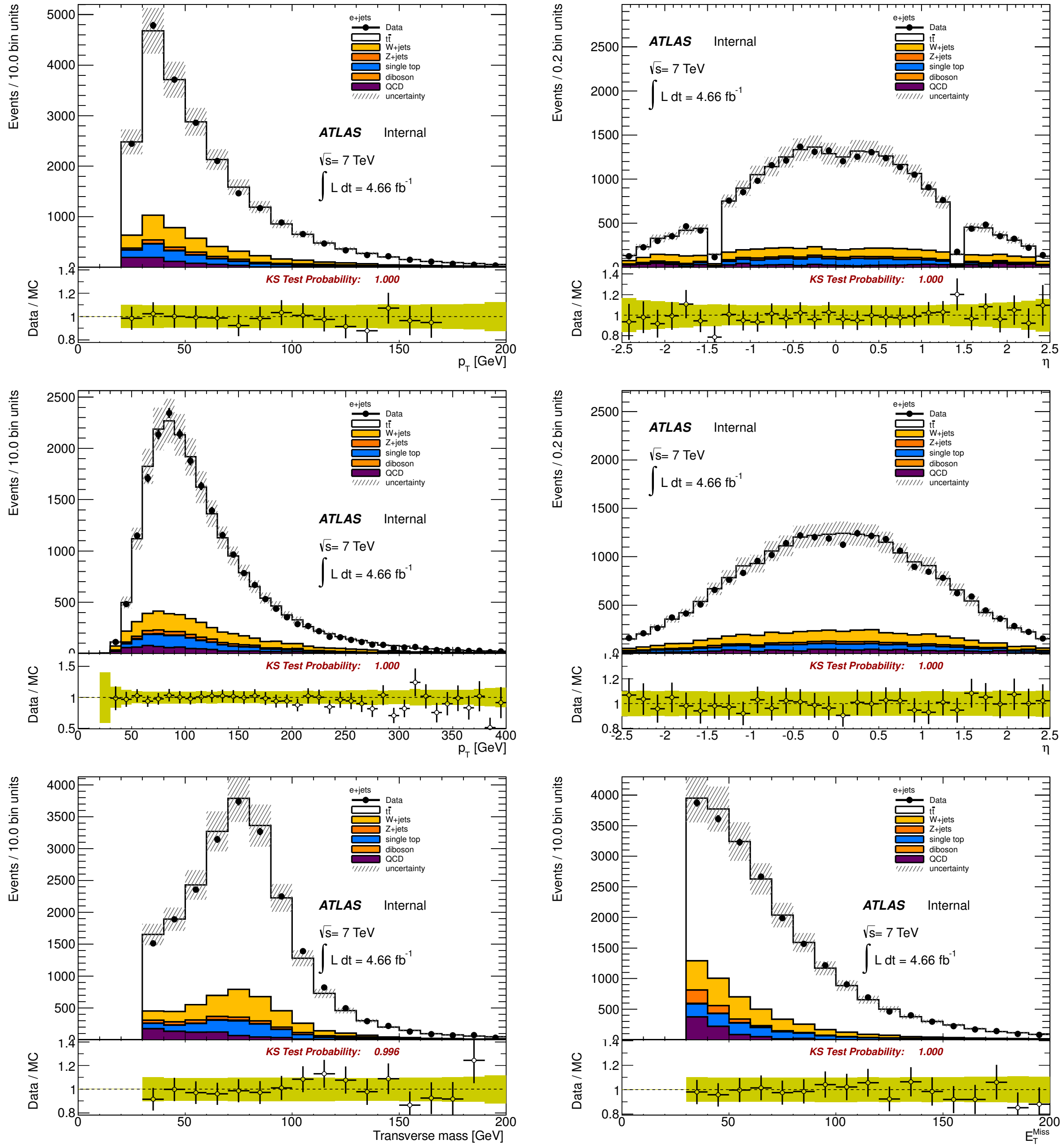


Figure B.4: Control plots for the $e+$jets channel in the tagged sample at $\sqrt{s} = 7$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}. Uncertainty on background normalization is shown.

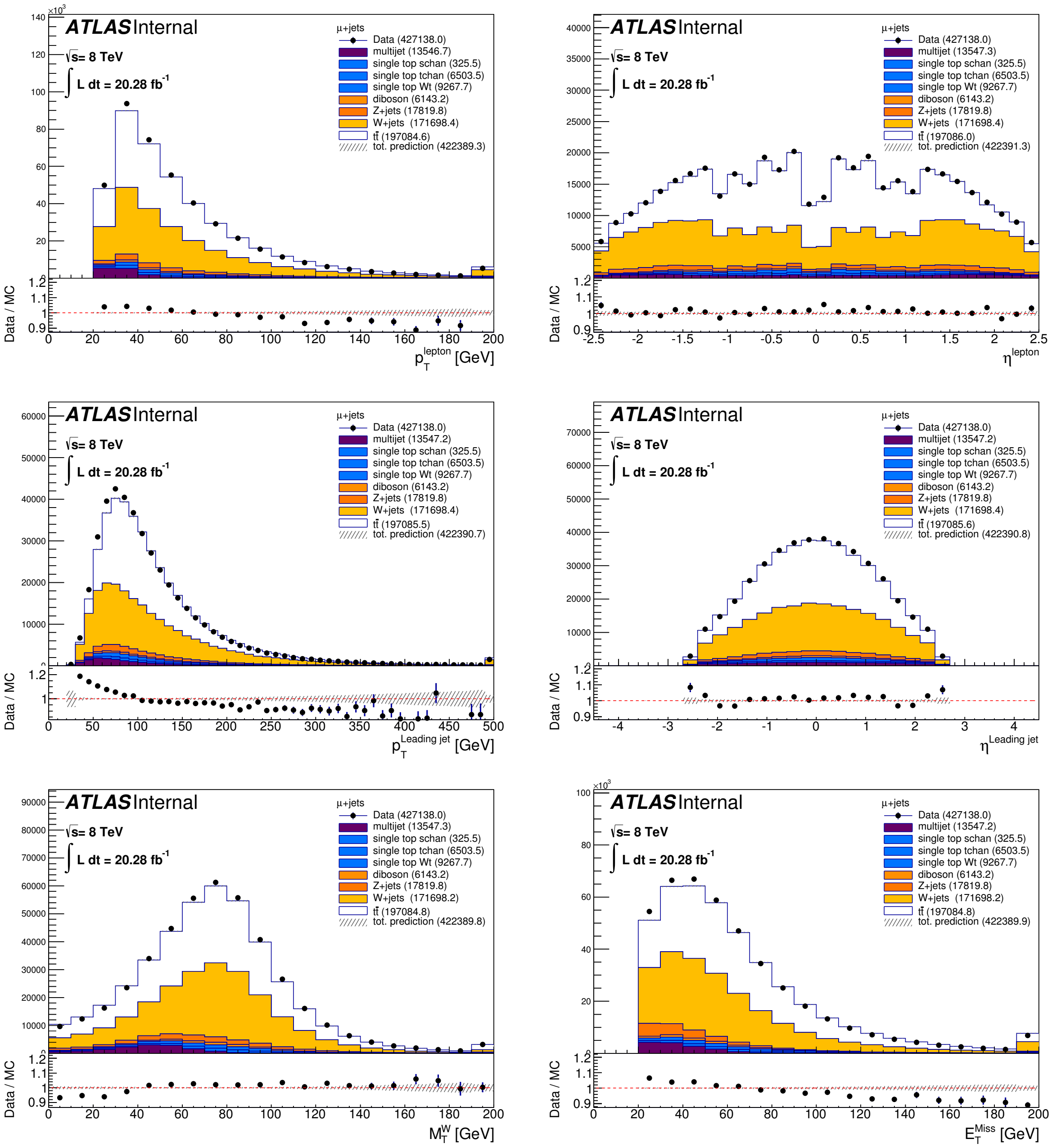


Figure B.5: Control plots for the $\mu + $jets channel in the pretag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The distributions for the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T} are shown in the bottom row.

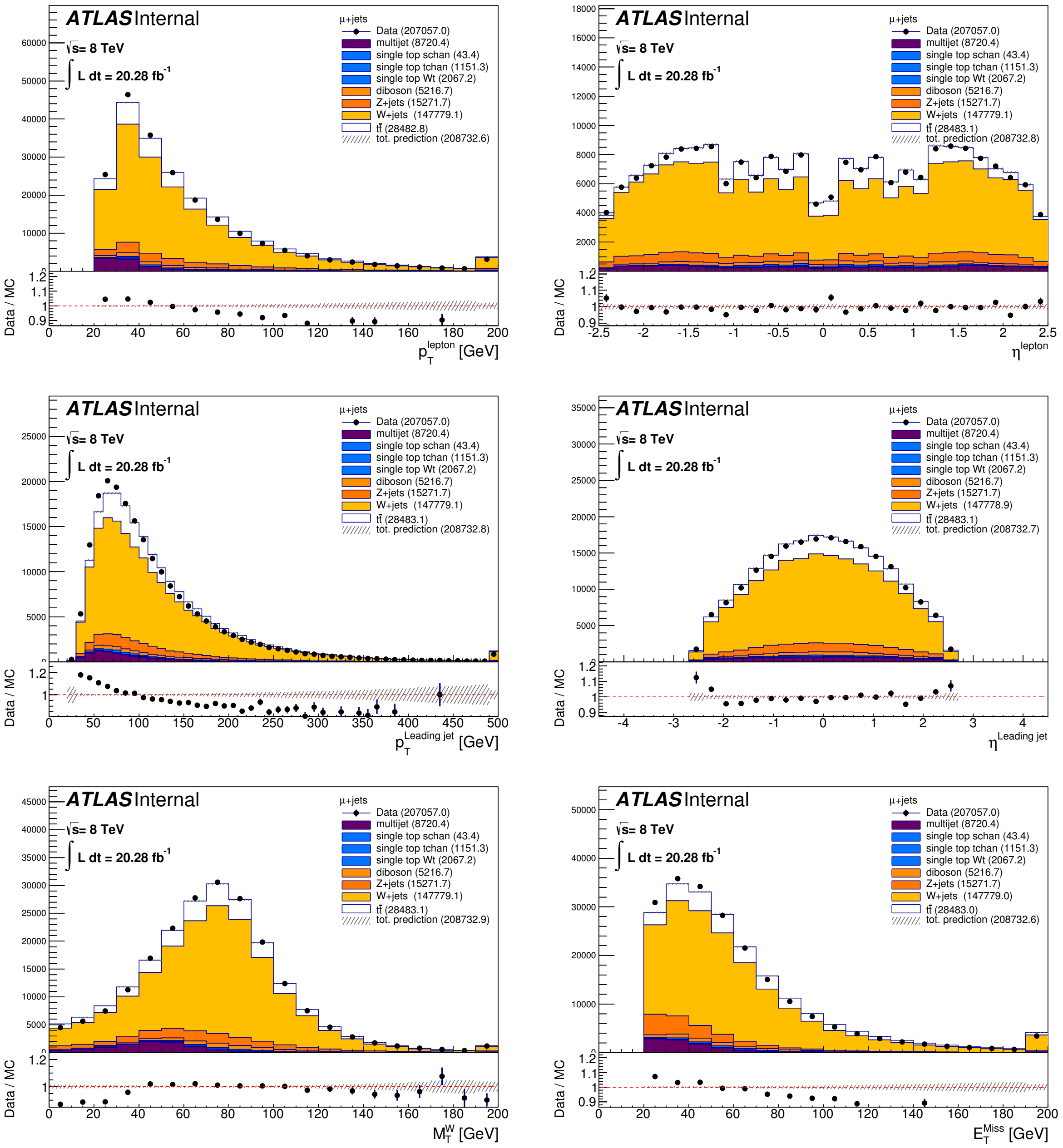


Figure B.6: Control plots for the $\mu +$jets channel in the 0-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_T, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_T^W and the missing transverse momentum E_T.

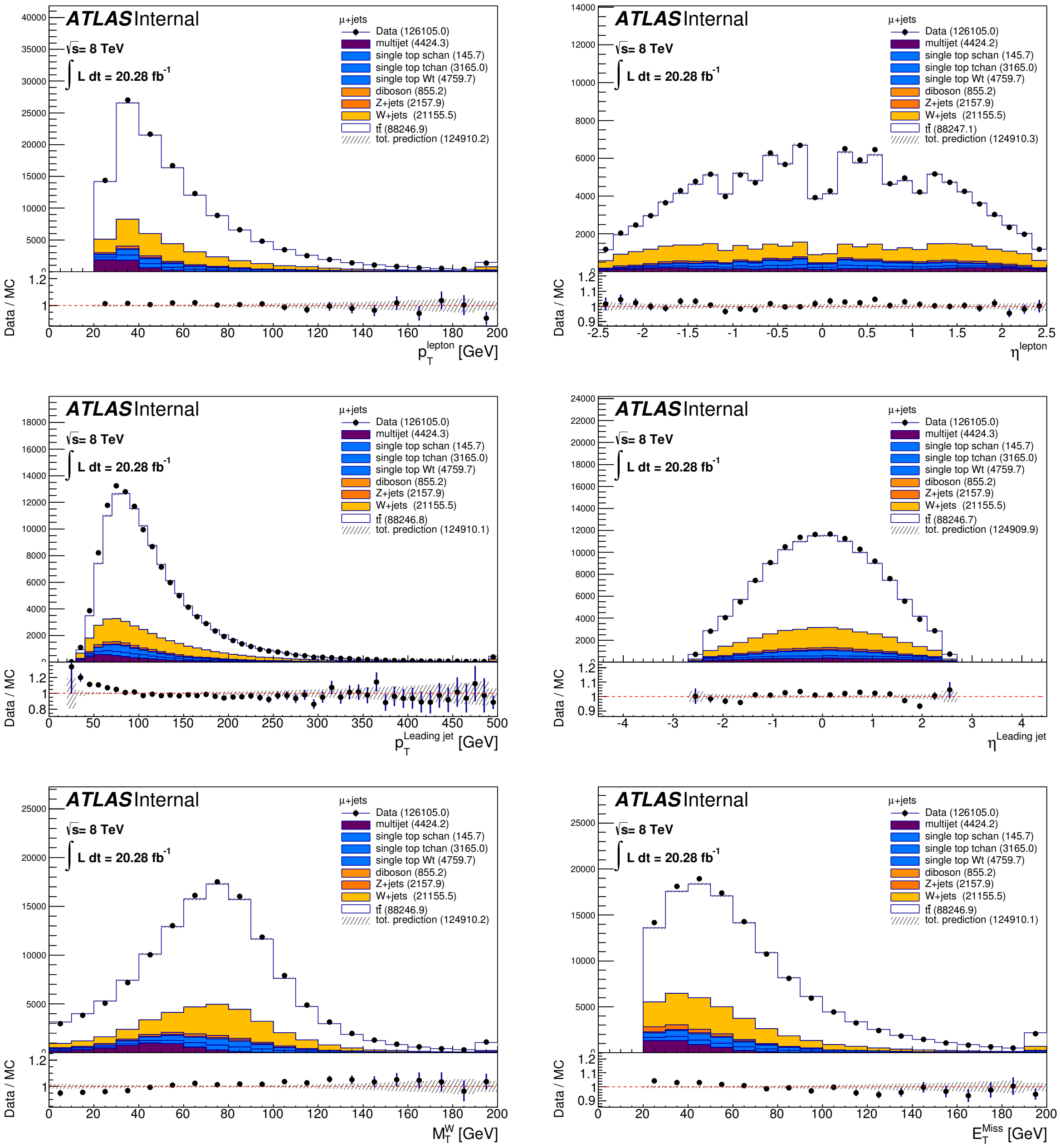


Figure B.7: Control plots for the $\mu + \text{jets}$ channel in the 1-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}.

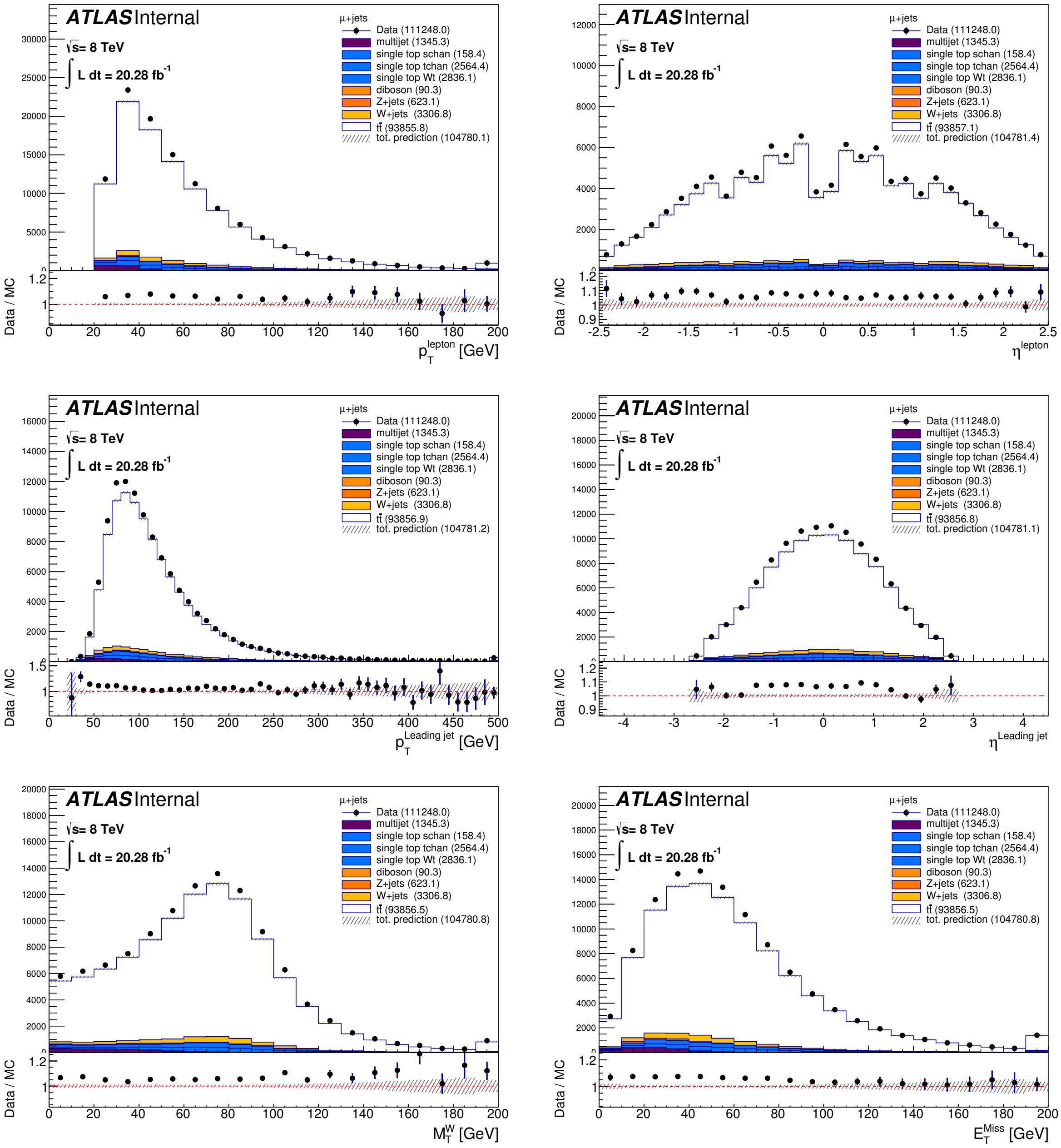


Figure B.8: Control plots for the $\mu +$ jets channel in the 2-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the muon (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}.

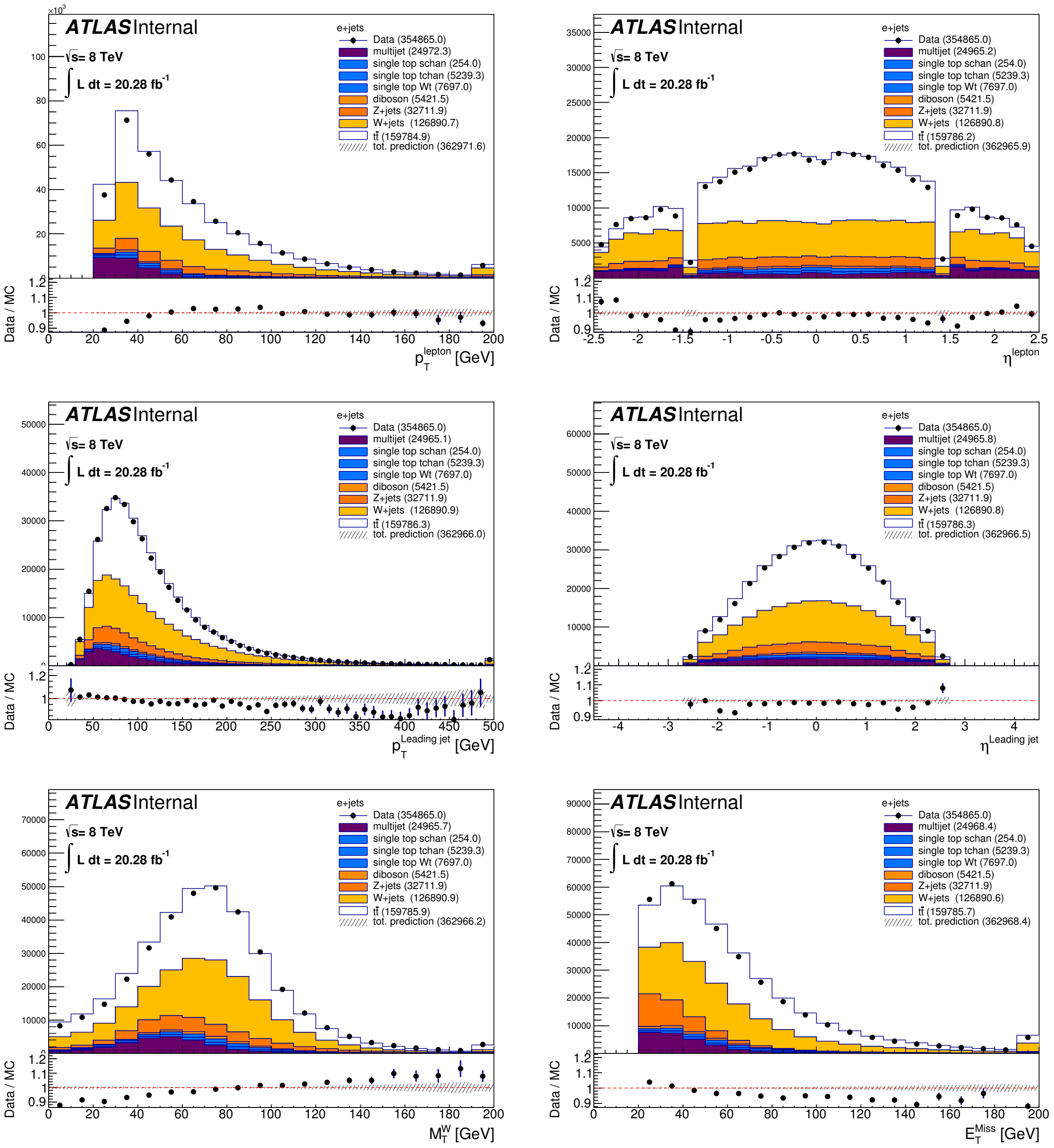


Figure B.9: Control plots for the $e +$ jets channel in the pretag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}.

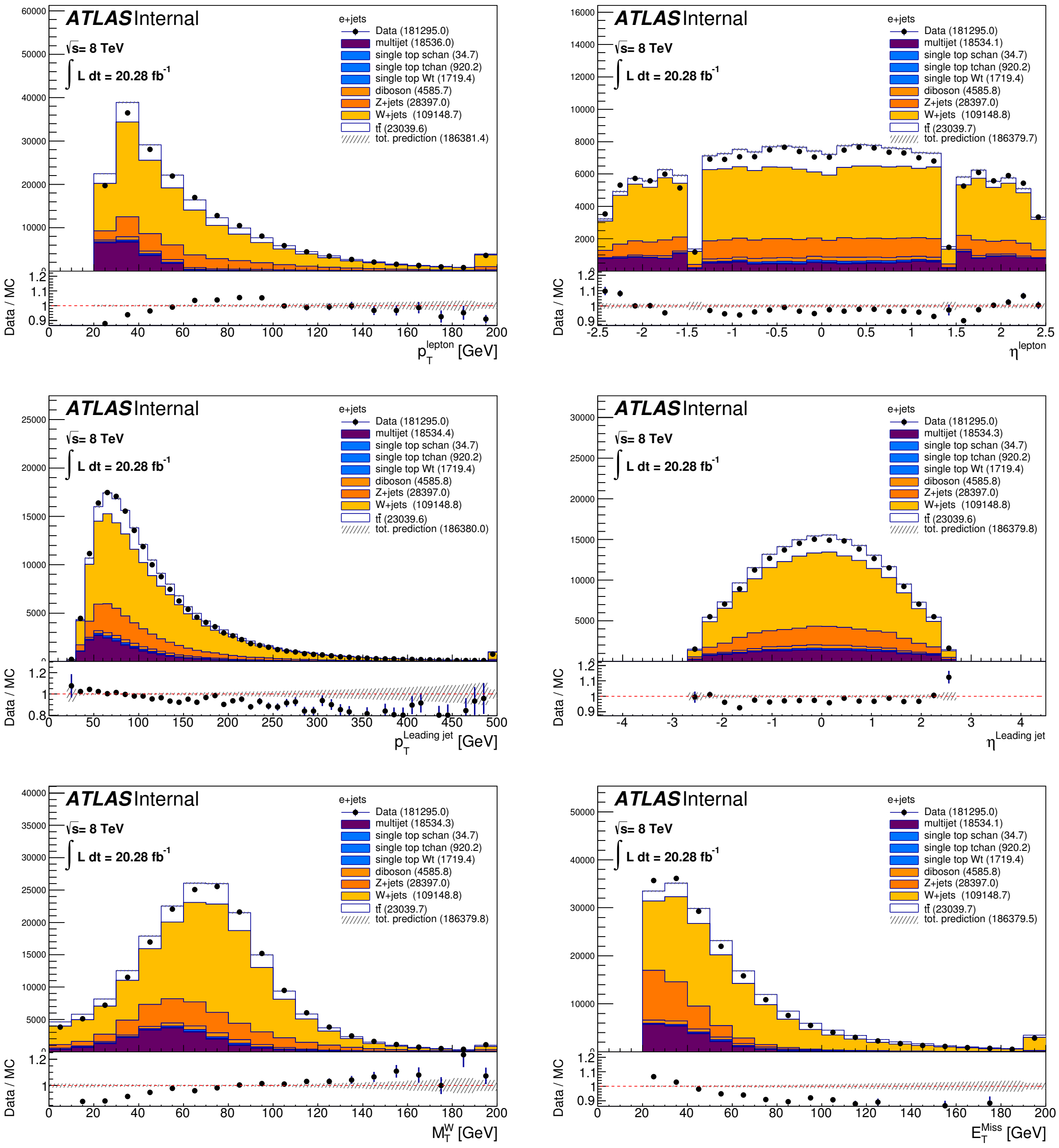


Figure B.10: Control plots for the $e+$jets channel in the 0-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}.

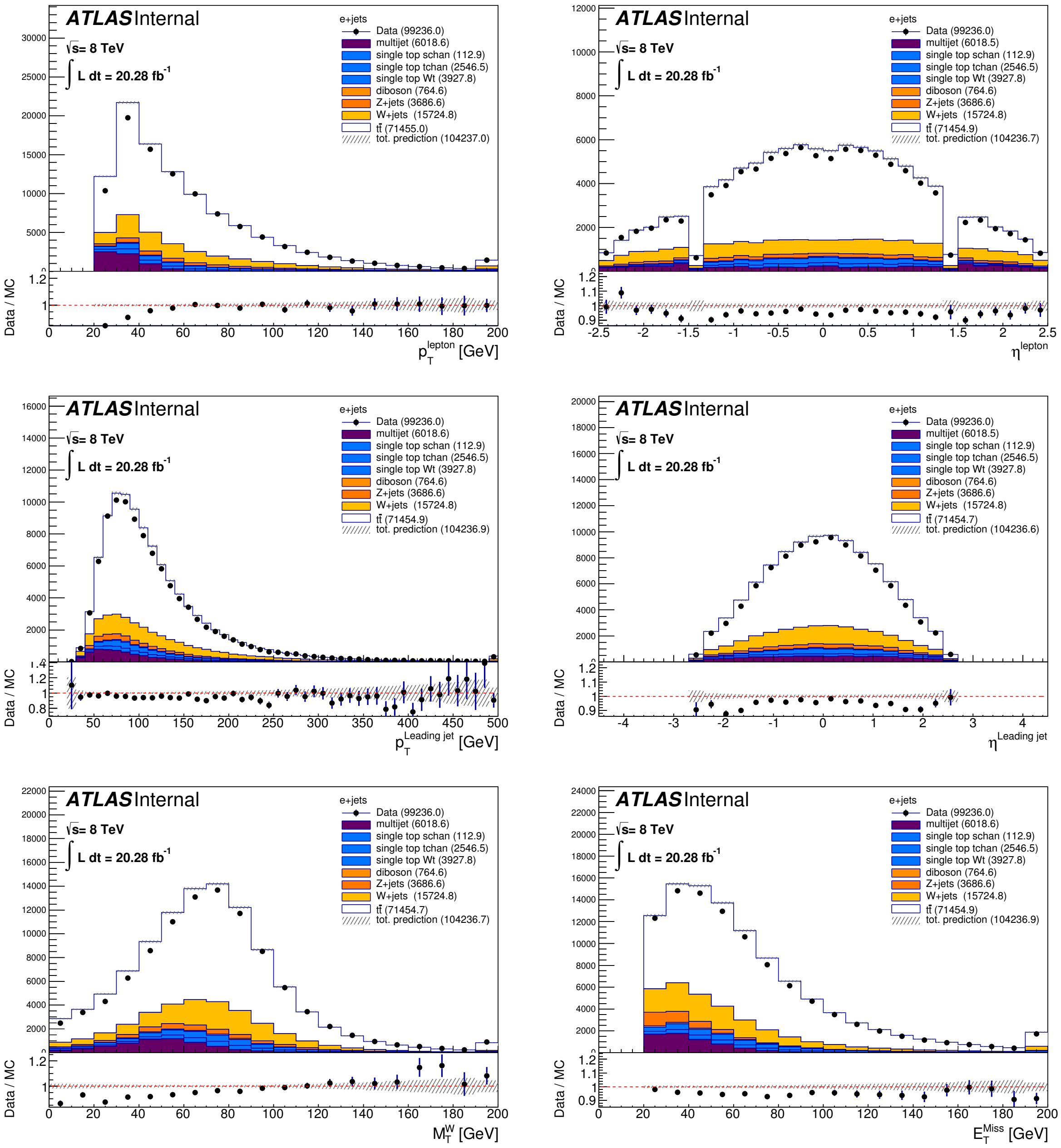


Figure B.11: Control plots for the $e +$jets channel in the 1-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_{T}, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_{T}^{W} and the missing transverse momentum E_{T}.

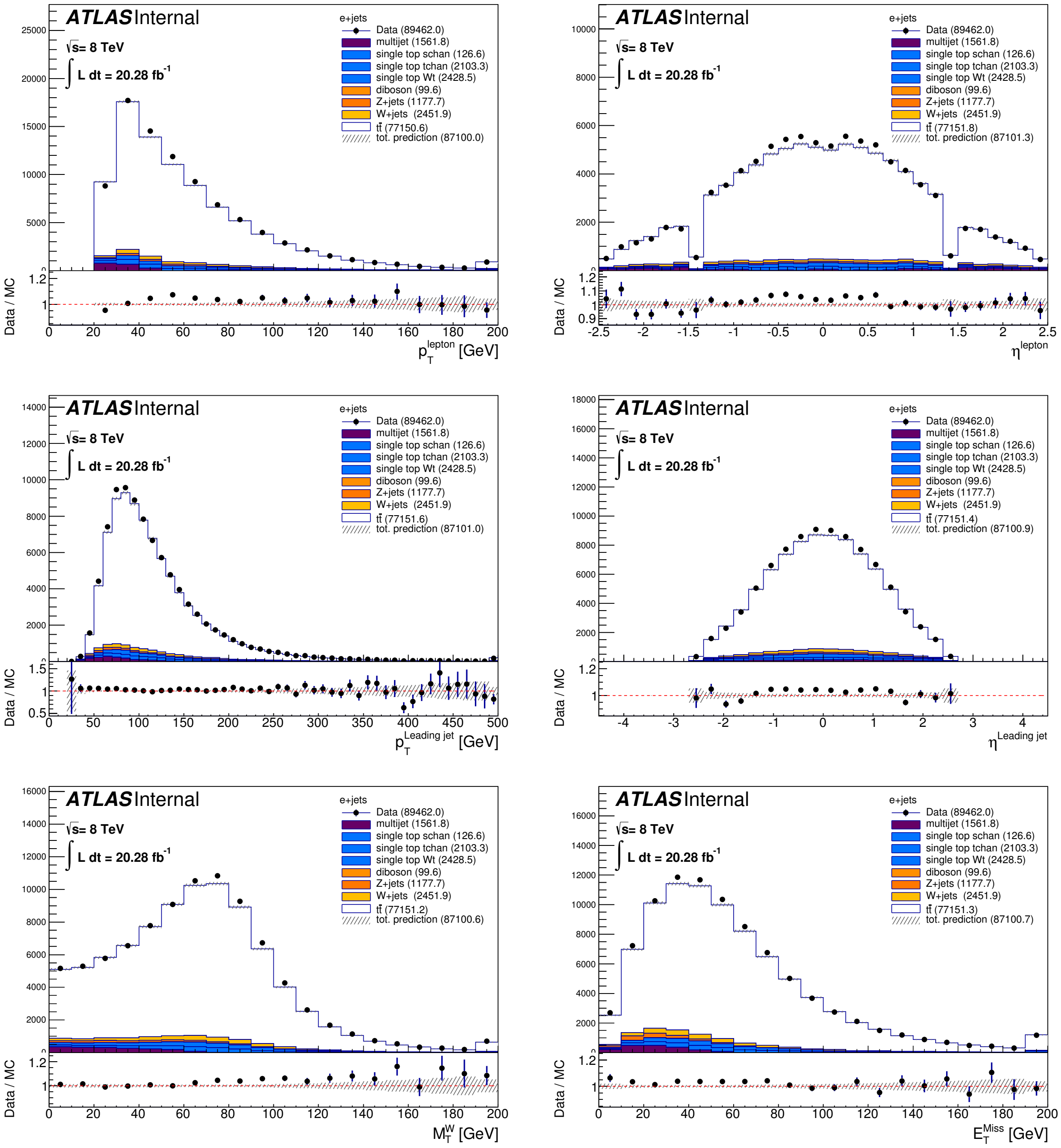


Figure B.12: Control plots for the $e +$ jets channel in the 2-tag sample at $\sqrt{s} = 8$ TeV. From the top left to the bottom right, the transverse momentum p_T, and the pseudorapidity η are shown for the electron (top) and for the leading jet (middle). The bottom plots show the W transverse mass m_T^W and the missing transverse momentum E_T.

Appendix C

Unfolding checks

In this appendix we illustrate the details of the FBU implementation the unfolding cross-checks.

C.0.1 Regularization

In this section we briefly summarize the details of the regularization used for the inclusive unfolding and for the differential one as a function of $p_{t\bar{t}}$.

We use the definition of curvature

$$C(\boldsymbol{T}) = \sum_{t=2}^{N_b-1} \left[(T_{t+1} - T_t) - (T_t - T_{t-1}) \right]^2 \tag{C.1}$$

where N_b is the number of bins. We compute the regularization function as the absolute value of the curvature difference between the truth spectrum and the $\boldsymbol{T}$ spectrum of the simulated events that are used to fill the transfer matrix:

$$S(\boldsymbol{T}) = \ C(\boldsymbol{T}) - C(\boldsymbol{T}) \tag{C.2}$$

The main idea behind this function is that it decreases the weight of generated points $\boldsymbol{T}$ with curvature that is very different from the one of the truth spectrum.

For the unfolding of two-dimensional spectra, the curvature is computed separately in each differential subrange and the regularization function is the sum of the regularization functions computed in each subrange. We choose to compute the curvature separately within each subrange for two reasons: first, we do not want to include in this calculation the bins that mark the transition between one subrange and the next; second, the numerical value of the curvature can be significantly different in different bins of the differential variable. The numerical values of the curvature are reported in table C.1; the numerical value of the parameter α is chosen so that $\alpha \cdot S$ is not far from unity.

Fig. C.1 shows the FBU posterior for the inclusive measurement corresponding to four different choices of the regularization parameter α. While the width of the posterior is sensibly reduced for larger values of α, the mean value is stable.

	Curvature : $\frac{\text{truth}}{\text{reco}}$				
inclusive	$\frac{6.1e+05}{9.7e+05}$				
$m_{t\bar{t}}$	$\frac{1.4e+09}{1.0e+07}$	$\frac{1.2e+08}{1.4e+05}$	$\frac{1.8e+08}{5.5e+05}$	$\frac{7.4e+07}{3.1e+05}$	$\frac{1.7e+07}{7.6e+04}$
$p_{t\bar{t}}$	$\frac{1.4e+06}{8.4e+04}$	$\frac{7.0e+05}{1.8e+05}$	$\frac{2.6e+06}{2.7e+05}$		
$y_{t\bar{t}}$	$\frac{1.8e+09}{1.5e+07}$	$\frac{3.2e+09}{2.3e+07}$	$\frac{5.2e+07}{2.8e+06}$	$\frac{4.9e+08}{1.5e+06}$	

Table C.1: Curvature numerical values of the spectra being unfolded, at the truth level and at the reconstruction level. For the differential spectra, the curvature is computed separately in each subrange corresponding to one bin of the differential variable.

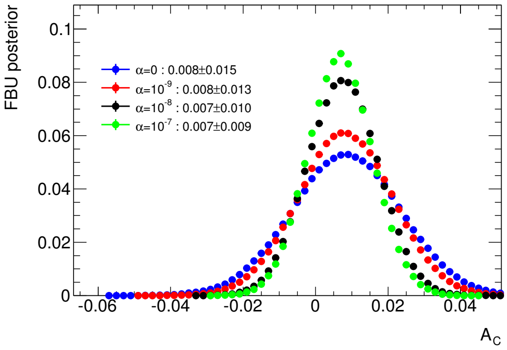


Figure C.1: FBU posterior for the inclusive measurement corresponding to four values of α.

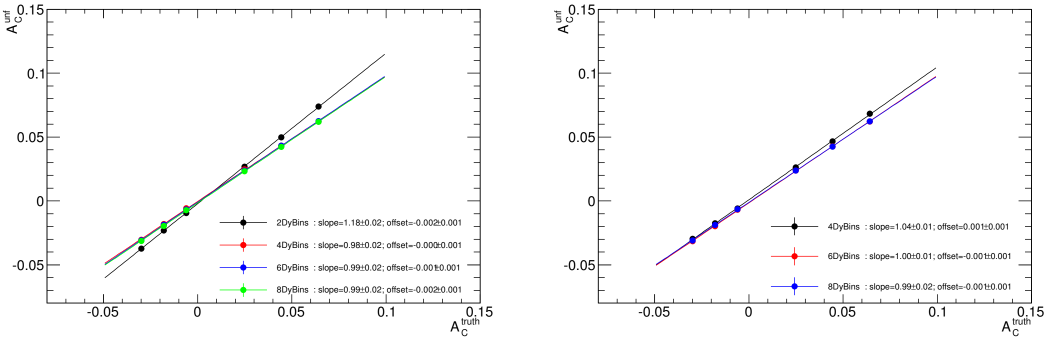


Figure C.2: Unfolding with FBU for $A_{C\text{incl}}$: linearity for several $\Delta\ y$ binnings using a flat prior (left) and a regularization prior based on the curvature of the $\Delta\ y$ distribution (right).

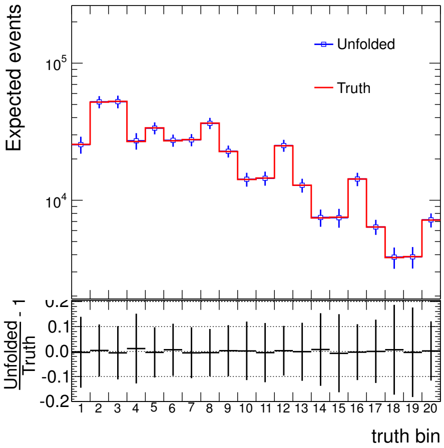


Figure C.3: Unfolded $\Delta\ y$ distribution for axigluon sample with 2% asymmetry in the five $t\bar{t}$ mass bins using a flat prior.

C.0.2 Linearity

The following figures show the results of the tests that we perform to verify the linearity of the FBU method. These tests are performed for the one-dimensional unfolding used in the inclusive measurement (figure C.2), and for the two-dimensional unfolding used in the differential measurements (figures C.7 C.8. The linearity is tested by measuring the A_C value after unfolding for several `PROTOS` axigluon samples that have a non-zero asymmetry. Figure C.12 illustrates the reduction in the unfolding statistical error that can be obtained by requiring that the kinematic fit likelihood is greater than -50.

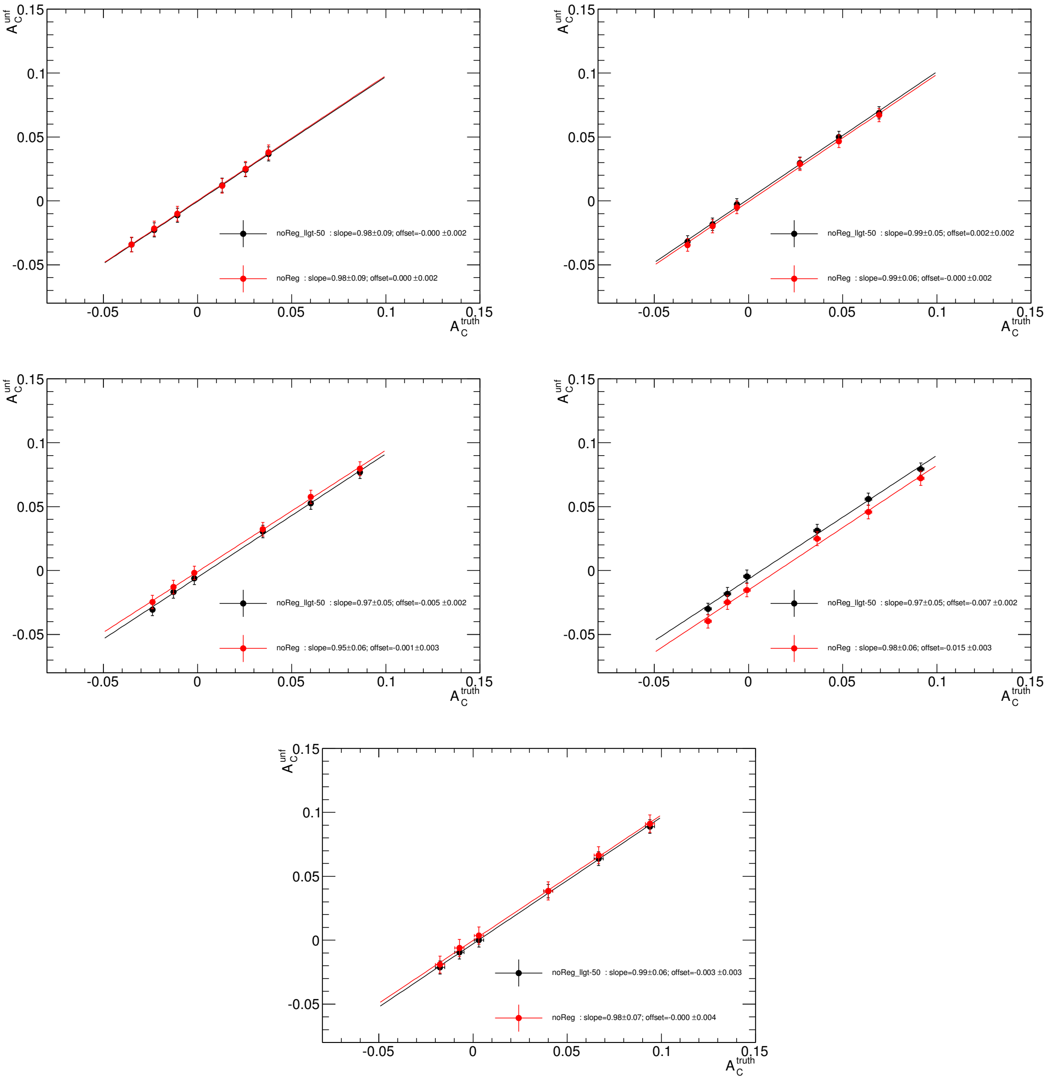


Figure C.4: Unfolding with FBU for $A_{C m_{t\bar{t}}}$: linearity in the five $t\bar{t}$ mass bins. Red curves are obtained without the cut on the kinematic fit likelihood; black curves are obtained applying the cut.

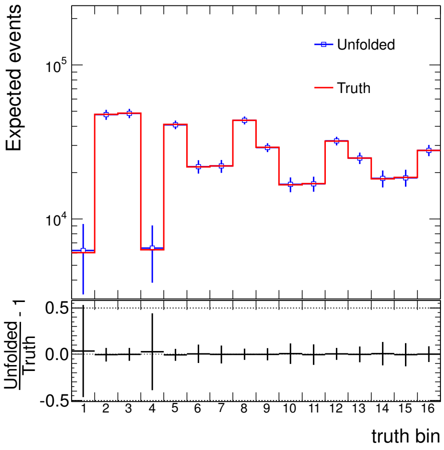


Figure C.5: Unfolded $\Delta\ y$ distribution for axigluon sample with 2% asymmetry in the three $t\bar{t}$ rapidity bins using a flat prior.

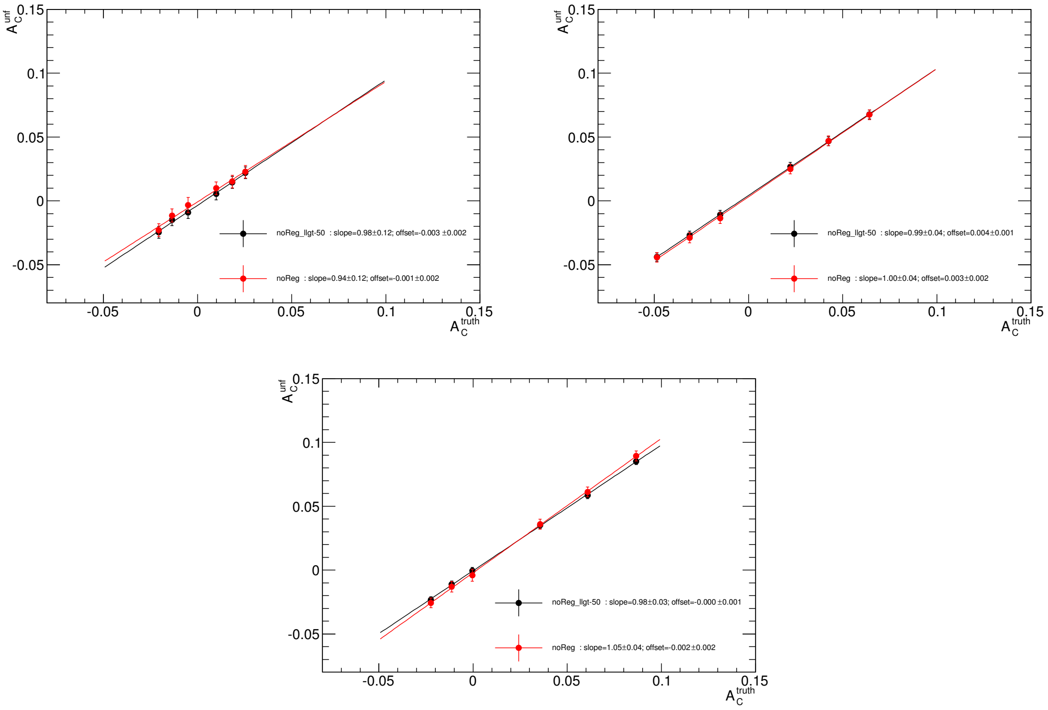


Figure C.6: Unfolding with FBU for $A_{Cy_{t\bar{t}}}$: linearity in the three $t\bar{t}$ rapidity bins. Red curves are obtained without the cut on the kinematic fit likelihood; black curves are obtained applying the cut.

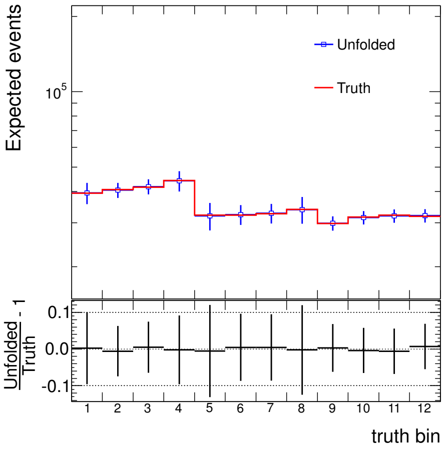


Figure C.7: Unfolded $\Delta\, y$ distribution for axigluon sample with 2% asymmetry in the three $t\bar{t}$ transverse momentum bins.

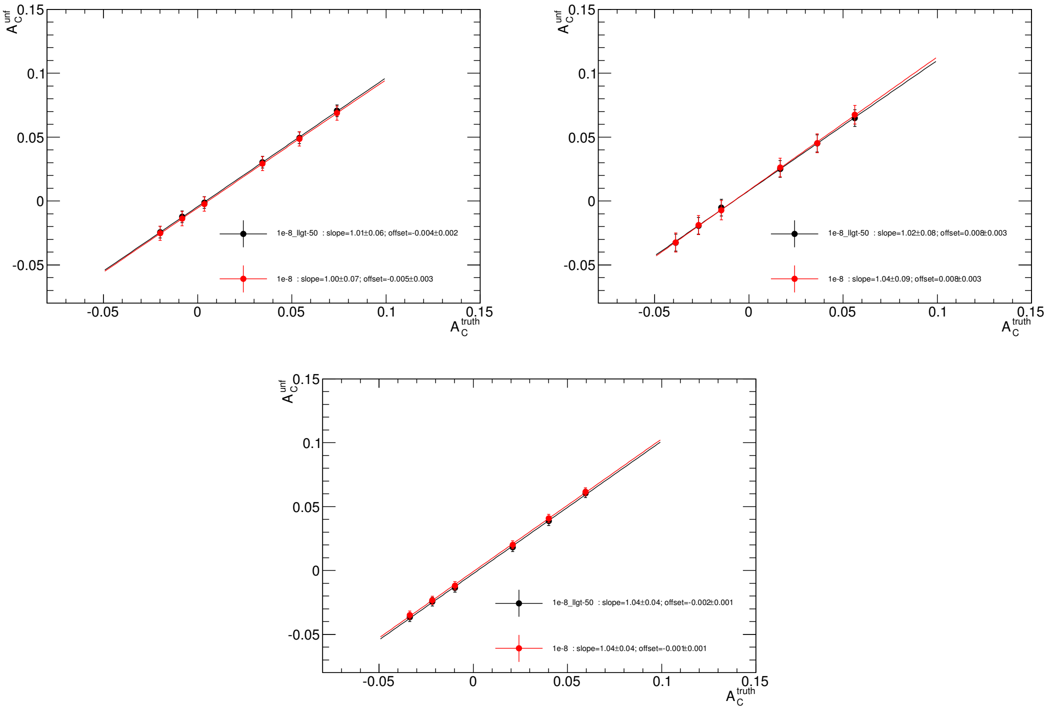


Figure C.8: Unfolding with FBU for $A_{C y_{t\bar{t}}}$: linearity in the three $t\bar{t}$ transverse momentum bins. Red curves are obtained without the cut on the kinematic fit likelihood; black curves are obtained applying the cut.

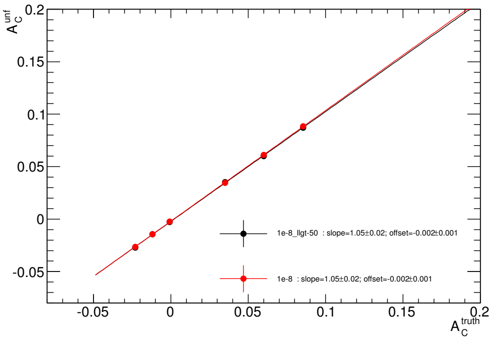


Figure C.9: Unfolding with FBU for $A_{C\text{incl}}$ with $\beta_{z,t\bar{t}}$>0.6: linearity. Red curves are obtained without the cut on the kinematic fit likelihood; black curves are obtained applying the cut.

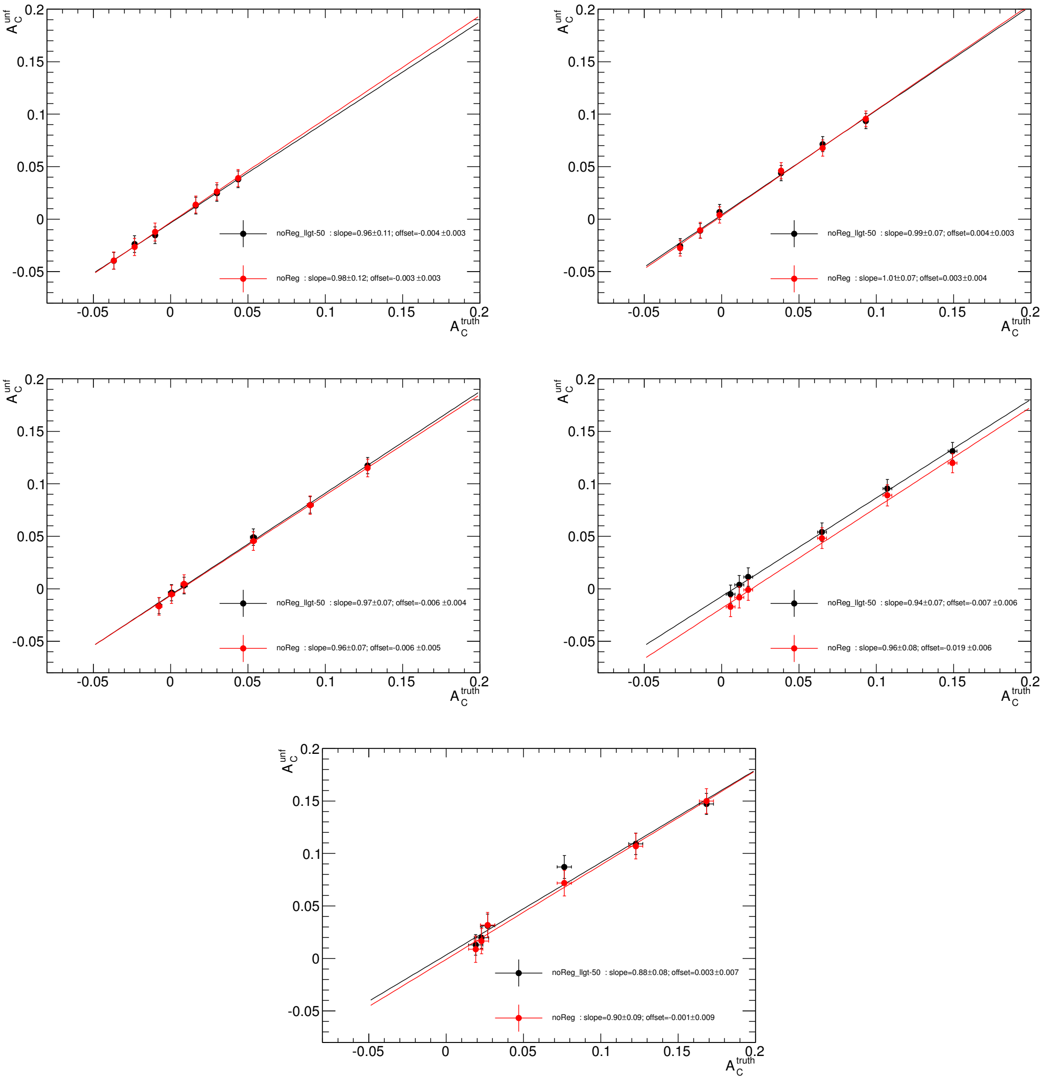


Figure C.10: Unfolding with FBU for $A_{Cm_{t\bar{t}}}$ with $\beta_{z,t\bar{t}}$>0.6: linearity in the five $t\bar{t}$ mass bins. Red curves are obtained without the cut on the kinematic fit likelihood; black curves are obtained applying the cut.

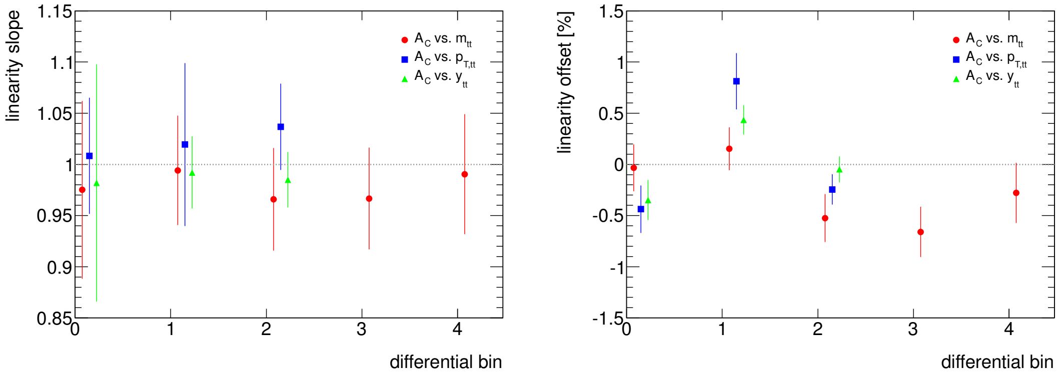


Figure C.11: Unfolding with FBU: slope (left) and offset (right) from the linearity fit for $A_{C m_{t\bar{t}}}$, $A_{C p_{t\bar{t}}}$, and $A_{C y_{t\bar{t}}}$.

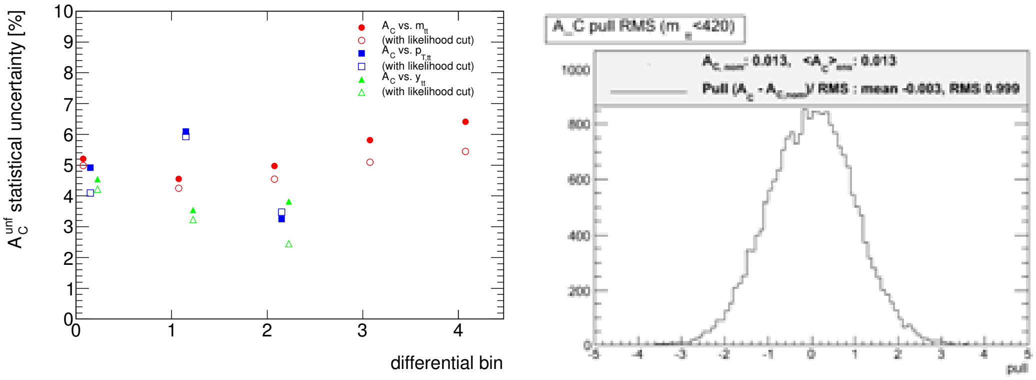


Figure C.12: Two-dimensional unfolding with FBU: A_C statistical error as a function of the differential bin (left) and example pull distribution (right) for the first $m_{t\bar{t}}$ bin.

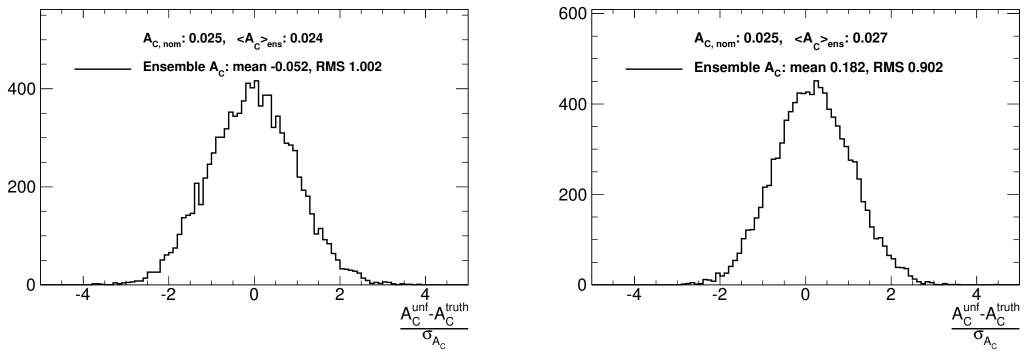


Figure C.13: Pull distributions for inclusive A_C with flat prior (left) and with regularization (right).

C.0.3 Pulls

We checked that the mean and RMS of the A_C posteriors are good estimators for the A_C interval by looking at pull distributions produced from ten thousand pseudo-experiments. Figures C.13,C.14,C.16 and C.15 show the pull distributions for inclusive and differential measurements. The RMS close to unity indicates the intervals are correctly estimated.

C.0.4 Acceptance efficiency in asymmetric samples

Fig. C.17 shows that the acceptance in $\Delta\, y$ bins does not change significantly in the reweighted `ALPGEN` samples used for the linearity test.

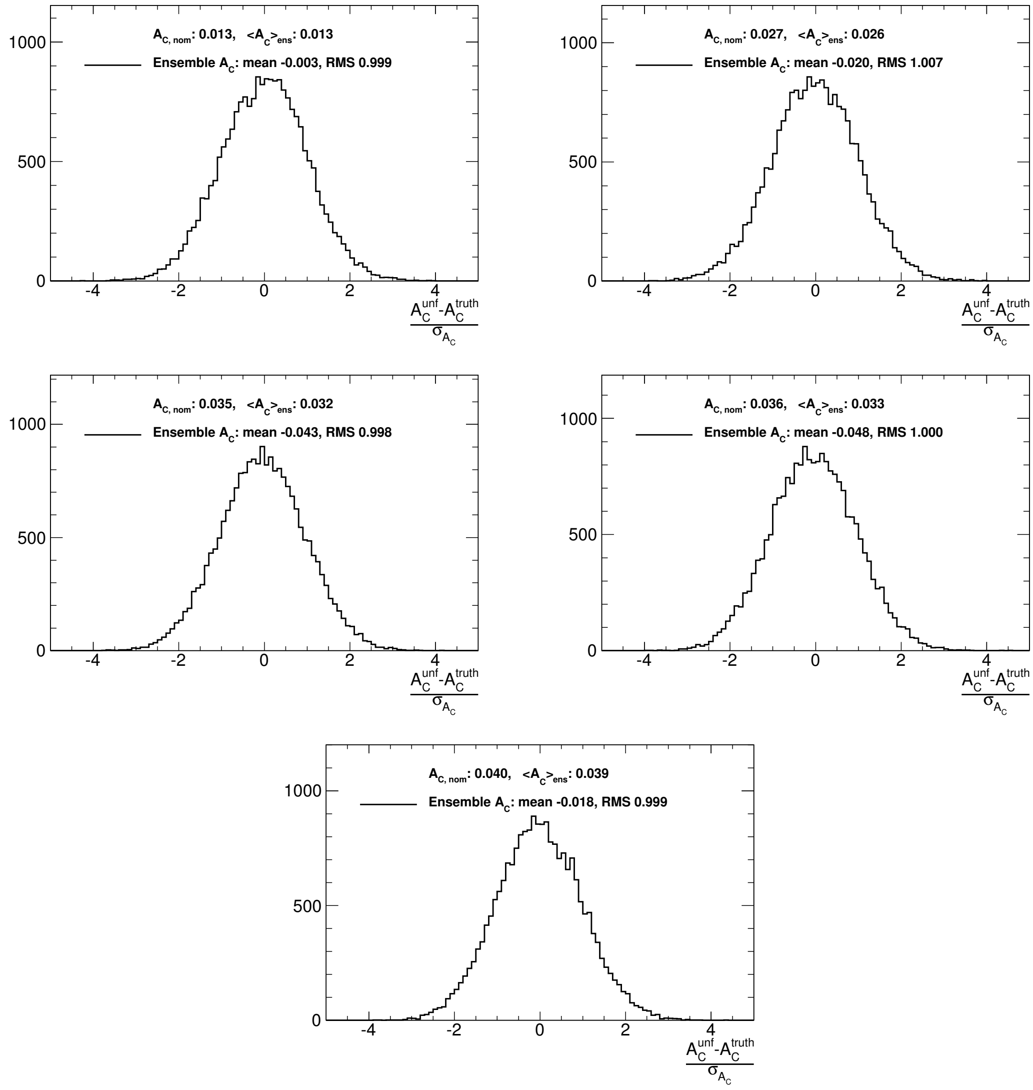


Figure C.14: Pull distributions for A_C in the five $m_{t\bar{t}}$ bins.

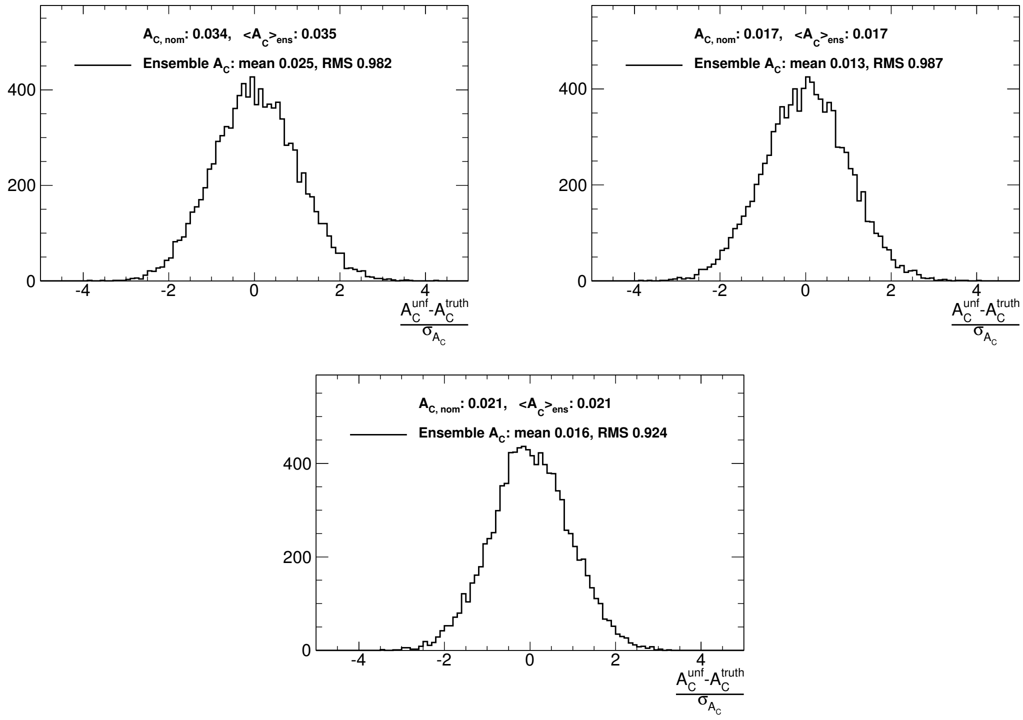


Figure C.15: Pull distributions for A_C in the three $p_{t\bar{t}}$ bins.

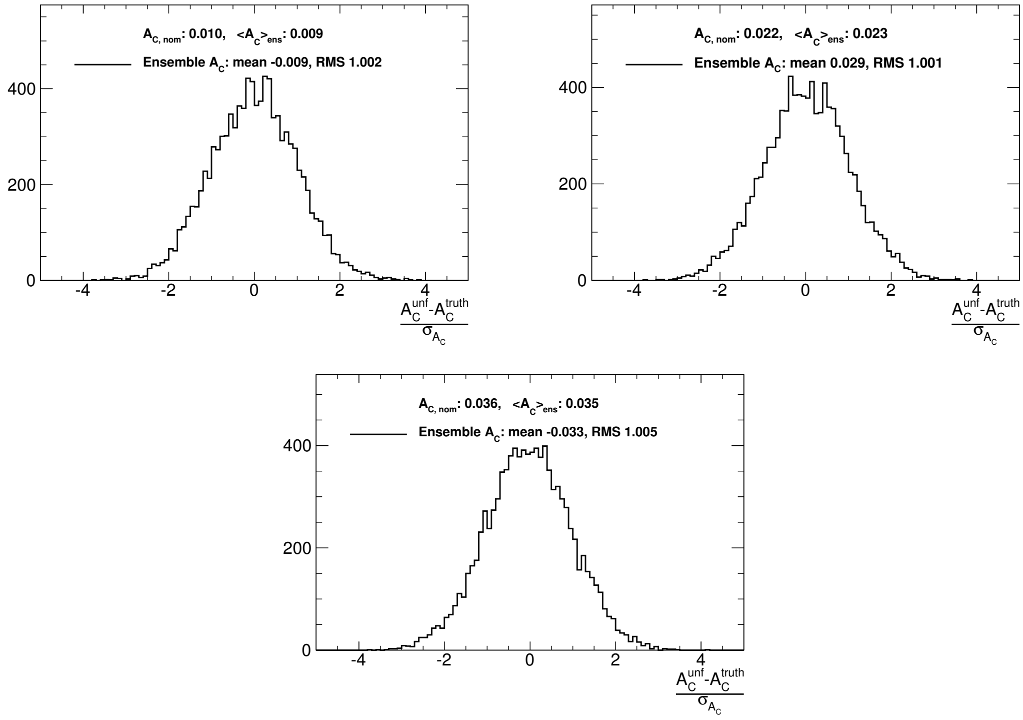


Figure C.16: Pull distributions for A_C in the three $y_{t\bar{t}}$ bins.

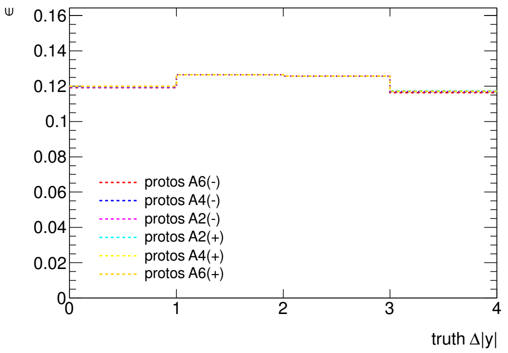


Figure C.17: Acceptance efficiencies for the six asymmetric `ALPGEN` samples obtained by reweighting with truth $\Delta\, y$ distributions from protos axigluon samples.

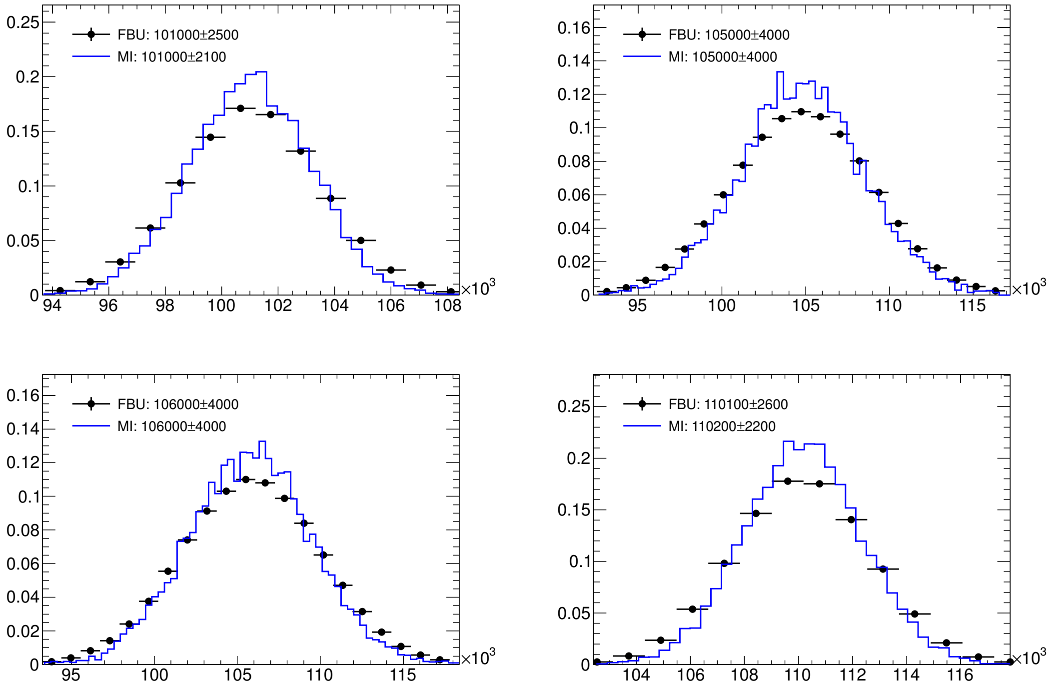


Figure C.18: FBU posterior distributions for the four unfolded Δy bins (black points) with overlayed the distributions of corresponding bin content from ensembles of unfolded Δy distributions obtained with matrix inversion (blue line).

C.0.5 Matrix Inversion

The FBU unfolding with a flat prior corresponds to the plain matrix inversion. In this section we compare the unfolded Δy distribution and asymmetry obtained with FBU with the ones obtained with the matrix inversion method.

Fig. C.18 shows the FBU posteriors for Δy bins compared with ensembles of the unfolded bin content obtained with matrix inversion. The mean and RMS values are compatible in the two cases. fig. C.19 shows the FBU posterior and the matrix inversion ensemble distribution for A_C. The distributions obtained with the two methods are cmpatible.

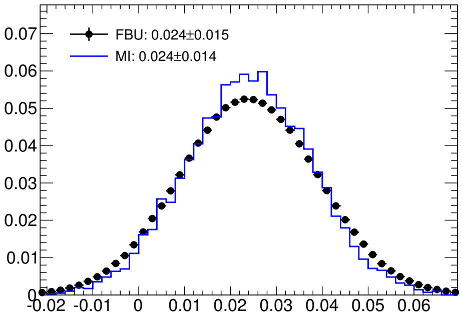


Figure C.19: FBU posterior distributions for A_C (black points) with overlayed the A_C distributions from ensembles of unfolded $\Delta\ y$ distributions obtained with matrix inversion (blue line).

Appendix D

Differential results

			$m_{t\bar{t}}$ [GeV]		
A_C	0 420	420 500	500 600	600 750	> 750
Unfolded	0 036 ± 0 055	0 003 ± 0 044	−0 039 ± 0 047	0 044 ± 0 054	0 011 ± 0 054
Theory	$0\ 0103^{+0.0003}_{-0.0004}$	$0\ 0123^{+0.0006}_{-0.0003}$	0 0125 ± 0 0002	$0\ 0156^{+0.0007}_{-0.0009}$	$0\ 0276^{+0.0004}_{-0.0008}$

Table D.1

Table D.2

		$p_{t\bar{t}}$ [GeV]	
A_C	0 25	25 60	> 60
Unfolded	−0 032 ± 0 052	0 067 ± 0 057	−0 034 ± 0 034
Theory	$0\ 0160^{+0.0007}_{-0.0009}$	$-0\ 0058^{+0.0004}_{-0.0004}$	$-0\ 0032^{+0.0002}_{-0.0002}$

Table D.3

		$y_{t\bar{t}}$	
A_C	0 0 3	0 3 0 7	> 0 7
Unfolded	−0 010 ± 0 043	0 006 ± 0 031	0 015 ± 0 025
Theory	$0\ 0026^{+0.0008}_{-0.0001}$	$0\ 0066^{+0.0001}_{-0.0003}$	$0\ 0202^{+0.0006}_{-0.0007}$

Table D.4

	$m_{t\bar{t}}$ [GeV] for $\beta_{z,t\bar{t}} > 0\ 6$				
A_C	0 420	420 500	500 600	600 750	$>$ 750
Unfolded	$0\ 054 \pm 0\ 079$	$0\ 008 \pm 0\ 072$	$-0\ 022 \pm 0\ 075$	$-0\ 019 \pm 0\ 102$	$0\ 205 \pm 0\ 135$
Theory	$0\ 0145^{+0.0005}_{-0.0003}$	$0\ 0213^{+0.0006}_{-0.0005}$	$0\ 0240^{+0.0003}_{-0.0009}$	$0\ 0280^{+0.0012}_{-0.0007}$	$0\ 0607 \pm 0\ 0002$

Table D.5

	$m_{t\bar{t}}$ [GeV]				
$\rho_{i,j}$	0 420	420 500	500 600	600 750	$>$ 750
0 420	1	-0 38	0 13	-0 05	0 01
420 500		1	-0 53	0 17	-0 03
500 600			1	-0 54	0 14
600 750				1	-0 43
$>$ 750					1

Table D.6

	$p_{t\bar{t}}$ [GeV]		
$\rho_{i,j}$	0 25	25 60	$>$ 60
0 25	1	-0 79	0 36
25 60		1	-0 60
$>$ 60			1

Table D.7

	$y_{t\bar{t}}$		
$\rho_{i,j}$	0 0 3	0 3 0 7	$>$ 0 7
0 0 3	1	-0 33	0 05
0 3 0 7		1	-0 21
$>$ 0 7			1

Table D.8

	$m_{t\bar{t}}$ [GeV] for $\beta_{z,t\bar{t}} > 0\ 6$				
$\rho_{i,j}$	0 420	420 500	500 600	600 750	> 750
0 420	1	−0 36	0 08	−0 01	0 01
420 500		1	−0 57	0 19	−0 04
500 600			1	−0 59	0 16
600 750				1	−0 50
> 750					1

List of Figures

List of Tables