Fachbereich Mathematik und Naturwissenschaften
Fachgruppe Physik
Bergische Universität Wuppertal

Topics in the Measurement of Top Quark Events with ATLAS

**Pixel Detector Optoelectronics,
Track Impact Parameter Calibration,
Acceptance Correction Methods**

**Dissertation zur Erlangung des Doktorgrades
vorgelegt von
Stephan A. Sandvoss**

Oktober 2009

To Objectivity, Reliability, Validity,
and Teamwork

Überblick

Diese Arbeit beschreibt Verfahren, die insbesondere zur Messung von Top-Quark Ereignissen mit dem ATLAS Detektor am Large Hadron Collider (LHC) des CERN angewendet werden können. Es wurden Beiträge zu drei wesentlichen Bereichen erbracht: dem Detektoraufbau und seiner Überprüfung, der Datenkalibration und der ersten physikalischen Analyse.

Der Pixel Detektor verfügt über etwa 80 Millionen Auslesekanäle und repräsentiert damit mehr als 90% aller Auslesekanäle des ATLAS Detektors. Es wurde eine Methode zur Anwendung eines optischen Reflektometers entwickelt, mit der die Unversehrtheit der Lichtwellenleiterkabel, die die auslesende Optoelektronik außerhalb des Pixel Detektors mit derjenigen innerhalb verbinden, schnell getestet werden kann. Damit wurde die Einsatzbereitschaft des Pixel Detektors nach Produktion, Test und Installation dieser Ausleseelektronik (Back of Crate Karten) sichergestellt.

Präzisionsmessungen erfordern eine sorgfältige Detektorkalibration. Hierzu kann ein Verfahren beitragen, das die Spurstoßparameter von Monte Carlo (MC) Datensätzen an die Verteilungen von Daten anpaßt. Zunächst werden die Spur- und Jetrekonstruktion sowie das b-Tagging von Jets in ATLAS vorgestellt. Die Stoßparameter von Spuren sind vor allem durch den innersten Detektor bestimmt, den Pixel Detektor. Die Genauigkeit der Spurstoßparameter hat großen Einfluß auf die Resultate des b-Taggings und damit auf die Selektion und Analyse von Top Quark Ereignissen.

Zu Anfang der Messungen wird es Abweichungen in den Stoßparameterverteilungen zwischen simulierten und realen Daten geben, z.B. weil die relative Anordnung von Detektorteilen nicht genau genug bekannt ist. Auf die Spuren mit negativen Stoßparametern gestützt, die hauptsächlich von der intrinsischen Detektorauflösung abhängen, werden die Stoßparameter der simulierten Spuren angepaßt. Eine Implementierung im Athena Framework von ATLAS wurde aufgrund fehlender Kollisionsdaten mit zwei MC Top-Quark-Paar Datensätzen verschiedener Detektorgeometrien getestet. Mit der verfügbaren Statistik lassen sich Verbesserungen bei der Übereinstimmung von simulierten mit (pseudo-) realen Daten für Spuren mit Stoßparametern kleiner als 0.4 mm bis zu einem Faktor von 4 erreichen.

Die Analyse der ersten Top-Quark Ereignisse dient sowohl der Kalibration und Leistungsverbesserung des ATLAS Detektors als auch der Validierung bisheriger Messungen und der Überprüfung theoretischer Vorhersagen bei Energien, die nie zuvor erreicht wurden. Ein weiteres Verfahren gewichtet die Vorhersagen von Monte Carlo Generatoren für Top-Quark Observablen, so daß sie mit Daten übereinstimmen. Bei Top-Quark Messungen lassen sich so systematische Unsicherheiten verringern, die durch unterschiedliche Modellierung zugrundeliegender physikalischer Prozesse verursacht werden. Zunächst werden alle Schritte der MC Simulation in ATLAS vorgestellt, von der Generierung eines Ereignisses über die Simulation im Detektor bis zu seiner Rekonstruktion.
Die Vorhersagen von fünf Monte Carlo Generatoren werden hinsichtlich der totalen Selektionseffizienz von semileptonischen Top-Quark-Paar Ereignissen verglichen. Diese Effizienzen weichen wie einige Verteilungen der Selektionsobservablen und des Transversalimpulses des hadronisch zerfallenden Top Quarks um bis zu 20% voneinander ab. Das führt zu einer systematischen Unsicherheit der Messung aufgrund der Modellierung des zugrundeliegenden physikalischen Prozesses.
Um diese Unsicherheit für die ersten Daten zu verringern, rewichtet das entwikkelte Verfahren die MC Ereignisse anhand einer einzigen gemessenen Verteilung: anhand des Transversalimpulses des hadronisch zerfallenden Top-Quarks. Eine Implementierung und Anwendung im Athena Framework zeigt Verbesserungen von bis zu einem Faktor 10, die systematische Unsicherheit in der totalen Selektionseffizienz fällt um mehr als den Faktor 4 von etwa 20% auf unter 5%.

Contents

Preface

I am strongly inspired by Richard Feynman, his "Lectures on Physics" and many other of his books. Therefore, I start off with these excerpts from his lectures, that are most important to me:
"We must, incidentally, make it clear from the beginning that if a thing is not a science, it is not necessarily bad. For example, love is not a science. So, if something is said not to be a science, it does not mean that there is something wrong with it; it just means that it is not a science." [1, I, 3-1] "Poets say science takes away from the beauty of the stars — mere globs of gas atoms. Nothing is 'mere'. I too can see the stars on a desert night, and feel them...It does not do harm to the mystery to know a little about it." [1, I, 3-4] "In order to understand physical laws, you must understand that they are all some kind of approximation. In the same way, to define the mass of a single object is impossible, because there are not any single, left-alone objects in the world—every object is a mixture of a lot of things, so we can deal with it only as a series of approximations and idealiziations. One may prefer a mathematical definition; but mathematical definitions can never work in the real world." [1, I, 12-1] "In its efforts to learn as much as possible about nature, modern physics has found that certain things can never be 'known' with certainty. Much of our knowledge must always remain uncertain. The most we can know is in terms of probabilities. This is the best description of nature that one can give." [1, I, 6-5] "The test of all knowledge is experiment. Experiment is the sole judge of scientific 'truth'." [1, I, 1-1]
"Why repeat all this? Because there are new generations born every day. Because there are great ideas developed in the history of man, and these ideas do not last unless they are passed purposely and clearly from generation to generation." [2]

For any questions the author can be reached by sending an email to Stephan.Sandvoss@cern.ch.

Acknowledgements

I would like to thank my supervisor, Prof. Dr. Peter Mättig, to give me the opportunity to do my PhD under his survey.

Many thanks to all members of the ATLAS Group of Wuppertal for their friendship and support.

Thanks to all colleagues of ATLAS and CERN who helped me according to circumstances.

Stephan Sandvoss

Chapter 1

Introduction

Currently, several thousands of scientists from all over the world are preparing for the first beam collisions at the Large Hadron Collider (LHC). This accelerator at the European Center for Nuclear Research (CERN) in Geneva/Switzerland will provide the opportunity to make precision measurements of top quark properties and eventually to discover the last fundamental particle proposed by the Standard Model (SM) of particle physics: the Higgs particle. In order to do so, the detectors, for example ATLAS, have to work reliably and be well understood, and systematic uncertainties in the event reconstruction or in the physics modelling must be reduced as much as possible.

In this thesis contributions to three fields were made: hardware production, installation and commissioning for the ATLAS pixel detector (Chapter 4), data calibration (Chapter 6), and physics analysis preparation for first data (Chapter 7). Each of these three fields is important for or strongly related to top quark physics.
This thesis is organized as follows: A theoretical overview of particle physics, especially of the top quark sector, is given in Chapter 2.

Top quarks will be produced in the four interaction points of the LHC. Centered around these points, where the two proton beams circling in opposite directions collide, large detectors are installed: ALICE, ATLAS, CMS, LHCb. They will measure the products of the collisions, and hopefully some unexpected physics will be discoverd. The main features of LHC and ATLAS are presented in Chapter 3.

A part of the work for this thesis was devoted to the production, installation and commissioning of optoelectronics for the innermost subdetector of ATLAS: the pixel detector. This detector is crucial for a precise reconstruction of the collision events, especially for the accurate measurements of electrons, vertices and for b-(jet-)tagging. The pixel detector has more than $8 \cdot 10^7$ readout channels, these represent more than 90% of all ATLAS readout channels. The steering and the data readout of the pixel detector is performed through an optical data transmission line, the optical link. Chapter 4 overviews the pixel detector and the

optical link. In particular, the work for and the results of production, installation and commissioning of the corresponding optoelectronics are presented.

After detection and recording of the collision events the raw data is analyzed. Physics analyses like top quark mass measurements need high level reconstructed objects such as jets or electrons. Therefore, a complete event reconstruction is performed and the corresponding information is stored in separate files. The reconstruction of all common high level objects is summarized in Chapter 5.

After their reconstruction, collision events are analyzed to find, for example, certain Higgs particle decays or to perform top quark precision measurements. In order to attain such goals, it is crucial to determine, whether jets originate from the fragmentation/hadronization of b-quarks. This identification method is called b-tagging, and the physics analyses mentioned above heavily rely on its performance. Track impact parameters are the leading ingredients for this procedure, which are mainly determined by the pixel detector.
Initially, there may be disagreements between track impact parameter distributions of Monte Carlo (MC) and real data samples. Such systematic uncertainties can be caused for example by misaligned detector modules. In Chapter 6, a method is presented, with which the MC track impact parameters can be tuned from data. Hence, agreement between simulation and data, and, therefore, a smaller systematic uncertainty shall be reached.

In Chapter 7, the predictions of MC generators for top quark pair production and decay are analyzed. Five common MC generators were used to simulate the semileptonic decay channel of top-antitop events. The distributions of observables, which are relevant to the selection and analysis of such events, are presented and compared with each other.
Furthermore, a method is presented that allows to reweight the Monte Carlo predictions such that they agree better with data. In addition, the method provides an estimate of the systematic uncertainty caused by imperfections of the physical model.

Finally, the results of this work are summarized, and an outlook to the forthcoming tasks is given in Chapter 8.

Chapter 2

Theoretical Overview

Top quark pair events of MC datasets are used in Chapter 6 and 7 to test the methods developed there and make some estimations for first data. Therefore, this chapter overviews the theory related to the top quark and MC generators. The top quark represents a fundamental particle of the Standard Model (SM) of particle physics. This model describes the behavior of matter at dimensions of about $10^{-15} - 10^{-18}$ m and is currently most supported by experiments.

2.1 The Standard Model of Particle Physics

At dimensions like or below the size of a molecule particles behave quantum-mechanically. In high energy physics, subatomic particles are used for scientific research, which usually travel at relativistic velocities near the speed of light. As a consequence, theoretical calculations for and interpretations of measurements of quantum particles have to be according to special relativity (Lorentz transformations). The theory that is supported by most experiments of particle physics and combines quantum effects and special relativity is a relativistic quantum field theory. It assumes the conservation of quantities like energy, momentum, electric charge and angular momentum for quantum particles.

Particles are characterized by their properties, among these are mass, electrical charge and intrinsic angular momentum, which is called spin. Hence, they are important observables in high energy physics, for example in precision measurements of top quark characteristics. Some particle properties, or quantum numbers, are additive, for example the momentum, other are multiplicative, for example parity, which is linked to space reflection. There is for each particle an antiparticle, which has the opposite additive quantum numbers and the same multiplicative quantum numbers. For example, the antitop quark $\bar{t}$ is the antiparticle of the top quark t.
Particles are divided into two groups by regarding their intrinsic property spin: the ones that obey Fermi-Dirac statistics are called fermions, to which the top

quark belongs. Fermions have half-integral spin (in units of $\hbar$), whereas bosons, which represent the second group, have integral spin and obey Bose-Einstein statistics.

Currently, four fundamental forces are known to act between particles: the strong, the electromagnetic, the weak and the gravitational force. The action between particles can be described by the exchange of virtual particles, which cannot be measured. These are called gauge bosons, and most of them can interact directly with each other.
As illustrated in figure 2.1, the gravitation is at dimensions, which are typical

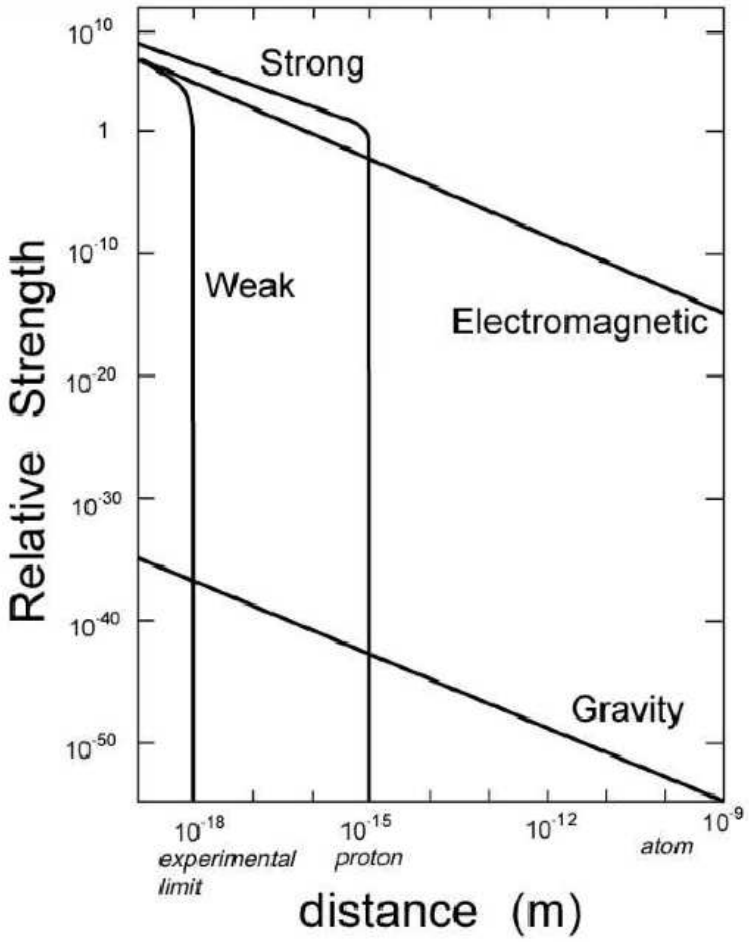

Figure 2.1: The relative strenghts of the four fundamental forces in dependence of the distance. The current experimental limit is noted as well. [3]

of particle physics ($< 10^{-15}$ m), by many orders of magnitudes weaker than the first three interactions. Therefore, gravitation is omitted in the Standard Model of particle physics and is irrelevant at the LHC, except if extra dimensions exist. Up to now, there is neither any direct experimental evidence of a gauge boson of

gravitation (graviton) nor a satisfactory quantum mechanical theory of gravity. Hence, the classical theory of general relativity is still the most accurate description of this force. Table 2.1 overviews the three interactions of the Standard Model and their gauge bosons, and table 2.2 summarizes the known fermions.

interaction/force	strong	electromagnetic	weak
coupling	color charge	electrical charge	weak charge
approximate range [m]	10^{-15}	∞	10^{-17}
gauge bosons	8 gluons (g)	photon (γ)	W^+, W^-, Z^0
their mass [GeV/c^2]	0	0	80 , 80 , 91

Table 2.1: The three fundamental forces of the Standard Model and their basic properties. The corresponding force carrier particles and their masses (in units of GeV/c^2 according to Einstein's equivalence of mass and energy: $E = m \cdot c^2$) are noted as well.

Since all three forces of the Standard Model act on the top quark, they are shortly presented. Quantum Chromo-Dynamics (QCD), the theory of the strong force, describes the interactions between so called color charges. There are three colors, or quantum states, which are usually called "blue", "green" and "red", but they have nothing to do with hue or frequency of light. Free observable particles are color singlets ("white"), in other words, they have zero net color. Directly observable particles that are acted upon by the strong force are called hadrons. They have a substructure and consist of bound colored particles called partons, which are not directly observable.
There are two types of partons: quarks (q) and gluons (g). Gluons are the gauge bosons of the strong force, and they have spin 1, no electric charge and no rest mass. Gluons can interact directly with each other via the strong force, because they carry color charge. A color-neutral gluon was never observed, hence there are only 8 different gluons, although 3 colors times 3 anticolors would result in 9 independent combinations or states.
Quarks are massive fermions with spin $\frac{1}{2}$ and carry a color charge of unit one. An antiquark $\bar{q}$ carries an anticolor of unit one ("antiblue", "antigreen" or "antired") and has the opposite electric charge as the quark. Six types or flavors of quarks are known: d(own)- and u(p)-, s(trange)- and c(harm)-, b(ottom)- and t(op)-quark. Their masses can be never measured directly in a well defined way, because free quarks have never been observed. 3 of the quark flavors — d, s and b — have an electric charge of $-\frac{1}{3}$ each. Whereas, a u-, c- and t-quark has an electric charge of $+\frac{2}{3}$ each. Except quarks all observable particles have integral charges, that are multiples of the elementary charge, as we will see below.

Leptons	1. Generation		2. Generation		3. Generation	
flavor	ν_e	e	ν_μ	μ	ν_τ	τ
electric charge [e]	0	-1	0	-1	0	-1
approximate mass [MeV/c^2]	$< 3 \cdot 10^{-6}$	0.511	< 0.19	105.7	< 18	1777
forces acted upon		el.mag.		el.mag.		el.mag.
	weak					

Quarks	1. Generation		2. Generation		3. Generation	
flavor	d	u	s	c	b	t
electric charge [e]	$-1/3$	$+2/3$	$-1/3$	$+2/3$	$-1/3$	$+2/3$
approximate mass [MeV/c^2]	5	3	120	1250	4200	175000
forces acted upon	strong, electromagnetic, weak					

Table 2.2: The basic properties and the interactions of the elementary fermions (spin $\frac{1}{2}$) known from the Standard Model. Leptons and quarks are sorted into three generations. The gravitation acts upon all particles.

Apart from their mass and life time, the d-, s- and b-quark have identical properties, the same is true for the u-, $c-$ and t-quark. Therefore, the quarks are grouped into three pairs, ordered by mass: d- and u-quark are called first generation quarks. s- and $c-$ quark represent the second generation, b- and t-quark the third. However, only the d- and u-quarks — together with gluons — may form stable hadrons.

Up to now, two ways were observed, in which quarks are combined to form a hadron: mesons and baryons, both together represent the hadrons, which are the strongly interacting particles. A combination of a quark q and an antiquark $\bar{q}$ is called meson, which is not stable and decays after some time. Its quark and antiquark have corresponding color and anticolor, so that the meson is color-neutral. Three quarks, which are bound together by gluons, constitute a baryon. Each of these quarks has a different color, so that the baryon is color-neutral. The best known baryons are the proton, which consists of 2 u- and 1 d-quark, and the neutron, which consists of 1 u- and 2 d-quarks. Protons and neutrons are the constituents (nucleons) of atomic nuclei. Usually, protons and neutrons are bound together in the nucleus by the strong force. However, not all combinations of protons and neutrons represent stable nuclei. Unstable nuclei decay or radiate particles, hence they are called radioactive. For example, nuclei decay, if the electrical repulsion between the protons, which carry an electrical charge of unit 1, is stronger than the strong attraction between the nucleons.

The electromagnetic force explains the interactions between electric charges and their interactions with light. According to Millikan's experiment, electric charges are positive (+) or negative (−) and an integral multiple of the elementary charge e. A proton carries an electric charge of $+e$, an electron $-e$. Two charges attract each other, if their signs are different, otherwise they repel. The theory of classical electrodynamics using static and dynamic fields was generalized to the Quantum Electro-Dynamics (QED). This theory describes electromagnetic effects by probabilities for the emission, flow and absorption of photons (γ). These are the gauge bosons of the electromagnetic force, which have spin 1 and no rest mass. They cannot interact directly with each other via the electromagnetic force, because they have no electric charge. Furthermore, the photon is its own antiparticle. The agreement between QED and experiments is very precise. For example, the value of the anomalous magnetic moment of the electron agrees between theory and experiment better than a part in a billion (10^9).

Finally, the weak force, which is the reason for the decay of the top quark, has 3 gauge bosons: W^+, W^-, Z^0. They have spin 1 and rest masses of about 80, 80 and 91 GeV/c^2. The weak force acts upon hadrons and leptons, which are fermions not affected by the strong force, because they have no substructure of colored particles.
There are three types or flavors of electrically charged leptons, which differ only in their rest mass and life time: $e^{\pm1}$ (electron/positron), $\mu^{\pm1}$ (muon), $\tau^{\pm1}$ (tau). Only the lightest, the electron, is stable. Moreover, there are three flavors of electrically neutral leptons, which are called neutrinos: ν_e, ν_μ, ν_τ. According to recent measurements, there are neutrino oscillations: a neutrino can make a transition from one flavor to any another, e.g. $\nu_e \rightarrow \nu_\mu$. This implies that at least 2 neutrinos have a nonzero rest mass. Like quarks, leptons can be divided into 3 generations, which have the same properties apart from rest mass and life time: e and ν_e represent the first, μ and ν_μ the second, τ and ν_τ the third generation. Since hadrons and leptons are affected by the weak force, there are 9 types of weak transition processes of particles. The initial state can be hadronic, leptonic or mixed (semileptonic), that is, there are only hadrons, leptons or both. The same is true for the final state. However, the total lepton and total baryon number are conserved under the weak interaction as well as under the other interactions. Moreover, the weak force is responsible for flavor changes of leptons or quarks. All possible transitions from one quark flavor to another quark flavor can be described by a corresponding combination of 9 probabilities, which represent the Cabibbo-Kobayashi-Maskawa (CKM) matrix.
The weak force explains also several effects as the beta decay or the proton-proton fusion in stars. Moreover, the weak force is the only interaction under which parity and CP-symmetry (symmetry of charge conjugation and parity) are violated.

As it can be seen in figure 2.1, the electromagnetic and weak force have about the same relative strength at dimensions below 10^{-18} m. This dimension corre-

sponds to an energy of $\Lambda \approx 246\,\mathrm{GeV}$, the so called electroweak scale. Therefore, the theories of the electromagnetic and weak force are combined to the electroweak theory, which suggests one further boson: the (Standard Model) Higgs particle. It is most likely that this particle completes the fundamental particles of the Standard Model. The Higgs particle is predicted to be massive, electrically neutral and to have spin 0. Though experimentalists search the Higgs boson already for several decades, it has not been found yet. A mass below $114.4\,\mathrm{GeV}/c^2$ and between 160 and $170\,\mathrm{GeV}/c^2$ is excluded at 95% confidence level by direct experimental searches [4, 5]. The most recent particle accelerator, the Large Hadron Collider (LHC), will provide by proton-proton collisions searches in mass ranges of up to $14\,\mathrm{TeV}/c^2$. The main Higgs production mechanism at the LHC is the gluon fusion process for all possible Higgs masses. The cross section of this process is determined by the Higgs-fermion coupling. Since the top quark mass is much larger than that of all other known fermions, the Higgs production is dominated by the Higgs-top coupling. Therefore, a detailed understanding of the top quark is important for the Higgs sector. The LHC and one of its four experiments, ATLAS, are presented in Chapter 3.

It seems that the Higgs boson is the only missing part for a full validation of the Standard Model. However, there are several other puzzles and inconsistencies, which arise from different experiments and measurements. For example, due to the observed neutrino oscillations mentioned above either a neutrinoless double beta decay or right-handed neutrinos should be observable. Moreover, neither observations of glueballs, which are hadrons containing no valence quark, nor of pentaquarks, which are baryons containing four quarks and one antiquark, have been confirmed, though QCD allows such objects. Furthermore, neither the wide mass range of the fermions (the so called hierarchy problem) nor the net excess of matter compared to antimatter by about 10 orders of magnitude (probably caused by CP violation) are satisfactorily explained.
Measurements in the top quark sector may play a major role in solving the mentioned inconsistencies. Since the top quark is by far the heaviest known fundamental fermion and thus quite near to the electroweak scale $\Lambda \approx 246\,\mathrm{GeV}$, its properties may indicate new physics. The following section introduces the top quark sector.

2.2 Top Quark Physics

The experiments of the Tevatron [6] accelerator, CDF [7] and DØ [8], discovered the top quark. Many top quark properties were measured there, for instance its mass ($172.4 \pm 1.2\,\mathrm{GeV}/c^2$ [9]). At the LHC, precision measurements with top quarks will be performed with high statistics: The cross section of a $t\bar{t}$-production in proton-proton collisions with design center of mass energy (14 TeV) will be about $\sigma(t\bar{t}) \approx 800\,\mathrm{pb}$ [10], which corresponds to 10 $t\bar{t}$-pairs per second

at design luminosity ($10^{34}\,\frac{1}{\mathrm{cm}^2\cdot\mathrm{s}}$). Such a production happens mainly by a gluon-gluon fusion, which is illustrated by the three leading Feynman diagrams in figure 2.2.

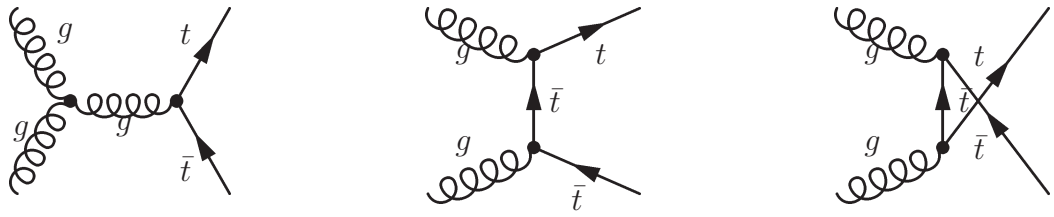

Figure 2.2: Feynman diagrams of a top/antitop quark pair production by gluon-gluon fusion. [11]

The life time of the top quark ($\approx 10^{-24}\,\mathrm{s}$) is shorter than the typical time, in which bound states of hadrons are formed by the strong interaction ($\approx 10^{-23}\,\mathrm{s}$). Therefore, this process called hadronization is very unprobable for top quarks, and hadrons containing top quarks will most likely never be observed. Since all other quarks are much lighter than the top, its decay is the only observable decay of a quark in an unbound state.
The decay of the top quark is reconstructed from detectable decay products. In more than 99% of the cases the top quark decays into a W boson and a bottom quark, so the other two possible transitions (to a strange or down quark) are negligible. Since a W can decay into a charged lepton and its corresponding neutrino ($\approx 33\%$) or into a quark-antiquark pair ($\approx 67\%$) [12] the final state of a $t\bar{t}$-decay can be leptonic, hadronic or mixed (semileptonic), as illustrated in figure 2.3. Since free quarks do not exist, they form color-neutral, often unstable hadrons, which later decay into several particles. Usually, these are also unstable,

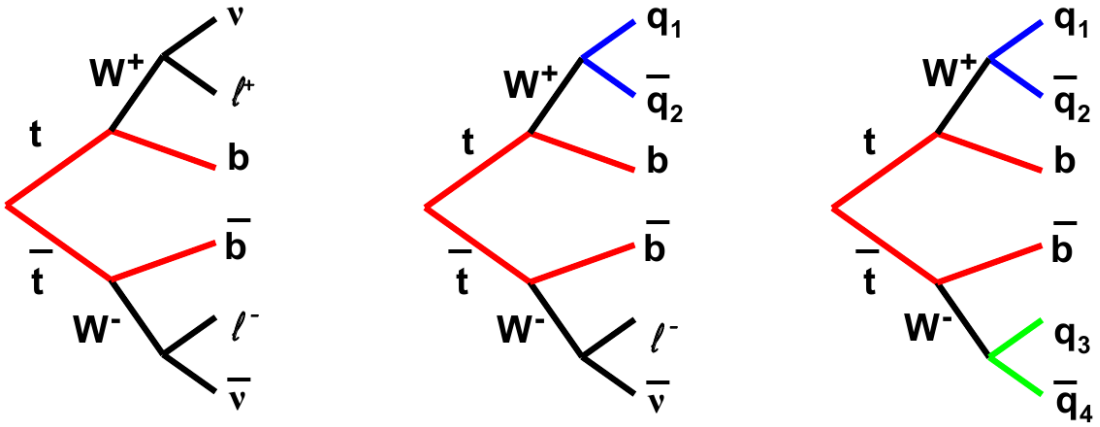

Figure 2.3: The final states of the decay of a top/antitop quark pair. [13]

what results in a long decay chain, until stable particles are left. Due to the strong production process at the parton level, most particles originating from a certain quark are concentrated in a small solid angle and measured as a single object called jet. Hence, a hadronic final state is also called a jet state; vice versa the semileptonic state is called lepton+jets state. Since τ-leptons can decay hadronically, the branching ratios observed in a detector may be different from the theoretical values provided by table 2.3, if the τ is not reconstructed.

final state	leptonic						lepton+jets			jets
channel	ee	$\mu\mu$	$\tau\tau$	$e\mu$	$e\tau$	$\mu\tau$	e+jets	μ+jets	τ+jets	jets
BR [%]	1.2	1.2	1.2	2.4	2.4	2.4	14.7	14.7	14.7	45.1

Table 2.3: Branching ratios of the decay of a top-antitop quark pair in percent. [12]

In order to measure precisely top quark properties, an accurate event reconstruction is required. Ideally, the energies and momenta of all particles and jets in the event are well measured and can be used for that purpose. However, neutrinos cannot be detected, because they interact neither strongly nor electromagnetically. As a consequence, the momentum of the collision event in the plane transverse to the colliding particles seems to be nonzero. That is why one or several neutrinos are also referred as missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$ or $\not\!\!E_{\mathrm{T}}$). In order to understand such effects, to estimate the influence of background processes and to compare experimental data with theoretical predictions, it is necessary to simulate particle collision events. This is described in the next section.

2.3 Monte Carlo Generators

The experimental measurements of top quark properties at Tevatron were included into the theory, in order to make further predictions. But since no analytical expressions are available for some quantities of experimental interest and since the kinematics of single collision events shall be studied, they have to be predicted by Monte Carlo generators. These are programs which use theoretical or emperical models and to which data from experiments is given. The generators simulate collisions of high energetic particles, in which many new particles are produced like illustrated in figure 2.4. As stated in [14], all Monte Carlo event generators adhere to the paradigm to divide a collision event into stages following the diagram of figure 2.4 inside out:

- The hard (high energetic) scattering process of two partons, when two hadrons collide. The probability to resolve a parton i with the longitudinal momentum fraction x_i inside a hadron by a momentum transfer of

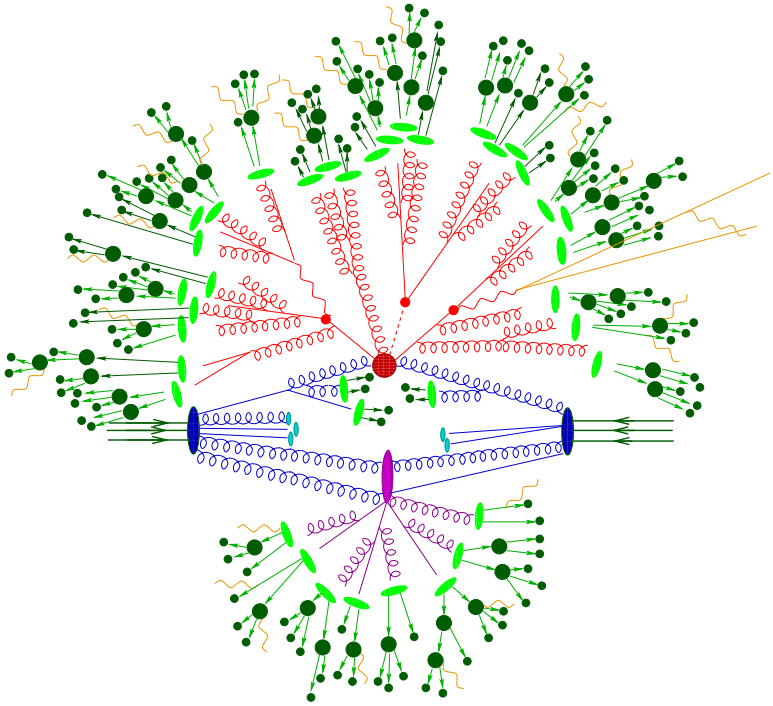

Figure 2.4: Sketch of a proton-proton collision event simulated by Monte Carlo. The incoming protons are drawn in dark blue. The dark red blob represents the hard scattering process of two colliding hadrons. QCD-Bremsstrahlung is drawn for the initial state (blue) and final state (red). The modelling of multiple parton interactions (underlying event) is shown in purple. The hadronization of partons is represented by light green ellipses. The decays of hadrons are drawn in dark green. Finally, QED-Bremsstrahlung is sketched in yellow. [15]

Q_i^2 between the incoming particle and the resolved parton is encoded in the parton distribution function $f(x_i, Q_i^2)$ (PDF). The scattering process of two partons can be calculated by integrating the matrix element (probability distribution) of the perturbation theory over a large phase space (momentum, spin). Usually, the calculation can be done within reasonable time only for matrix elements that represent the leading order (LO) or next-to-leading order (NLO) of perturbation theory. The leading order does not contain any loop and is also called tree level. The next-to-leading order involves one-loop and one-leg amplitudes.

- The QCD-Bremsstrahlung (gluons), which is emitted by accelerated colored partons. This leads to a cascade of emissions, which can be calculated by a parton shower approach. It uses perturbation theory (usually to leading logarithmic accuracy LLA) and stops the cascade by a cut off at a low energy scale ($\approx 1\,\mathrm{GeV} \gtrsim \Lambda_{\mathrm{QCD}}$), where perturbation theory is not reliable.

- The hadronization, which confines colored partons into uncolored ("white")

hadrons. This is not predictive from first principles, so phenomenological models are used, for example the string [16] or the cluster [17] model.

- The QED-Bremsstrahlung (photons), which is emitted by charged accelerated particles and mathematically identical with QCD-Bremsstrahlung.
- The underlying event, which is not well known, even often not clearly defined. It can include initial/final state showers, the hadronization of the remnants of the two collided particles, interactions between both remnants (multiple parton interactions), etc.
- The pile-up, which means that there is more than one interaction in each beam crossing. The number of soft or hard interactions per bunch crossing depends on the luminosity.

There exist many different approaches for every stage, and each MC generator is a certain combination of these approaches. Table 2.4 gives an overview of the main properties of five common MC generators (Pythia [18], AcerMC [19], Sherpa [20], Herwig [17], MC@NLO [21]). More information about MC generators can be found in [14]. Due to the various used approaches, MC generators may provide

Model	Matrix Element	Parton Shower	Hadronization	Underlying Event
Pythia	LO	p_T ordered	string model	multiple interactions, beam remnants
Herwig	LO	angular ordered	cluster model	minimum-bias like UA5 [22] or via JIMMY [23]
Sherpa	LO	p_T ordered	cluster or via Pythia	multiple interactions
AcerMC	LO	p_T ordered	via Pythia or Herwig	via Pythia or Herwig
MC@NLO	NLO	p_T ordered	via Herwig	via Herwig

Table 2.4: Basic properties of five common Monte Carlo generators.

different predictions for the same physical process. But experiments can decide, which predictions are more accurate for a certain physical process and a given MC generator.

If the MC predictions differ from data and systematic errors are excluded, there are two alternatives to provide an agreement between theory (MC generator) and experiment: First, the MC generators can be tuned, until simulation

and measurement agree within statistical and systematic uncertainty. This can be an elaborate and time consuming procedure. The alternative is to reweight the MC distributions, so that they agree with data afterwards. Such an approach for an analysis of the first data is presented in this thesis using the example of observables of top quark pair events. But first, we have a look on how particles are accelerated, lead to collision and how such a collision is detected.

Chapter 3

LHC and ATLAS

3.1 The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) [24] started beam commissioning in September 2008 at the laboratory of the European Organization for Nuclear Research (CERN) in Geneva. The LHC is located about 100 m underground in a large ring tunnel of about 27 Km circumference. It is partly in France and partly in Switzerland, between the Jura mountains and the airport of Geneva. The LHC can accelerate protons or ions.

At four points of the ring the two LHC vacuum beam pipes intersect. Here, the charged particles of the two beams travelling in opposite directions collide. In such collisions lots of new particles are produced, which fly from the collision point into all directions. By measuring these particles with a large detector, that is arranged around the interaction point, it can be concluded how matter behaves at the highest energies or the smallest dimensions. Figure 3.1 shows a schematic view of the LHC and its four experiments ALICE (A Lhc Ion Collider Experiment) [25], ATLAS (A Toroidal Lhc ApparatuS) [26], CMS (Compact Muon Solenoid) [27] and LHCb (LHC beauty experiment) [28].

A collider system of several linear accelerators and two ring accelerators (Proton Synchrotron (PS) [29] and Super Proton Synchrotron (SPS) [30]) provides the LHC with the beam particles. If these are protons, they are delivered with an energy of 450 GeV. The LHC can accelerate them up to an energy of 7 TeV. The charged particles get this energy by radio frequency radiation from superconducting cavities operating at a temperature of 4.5 K ($= -268.7\,^{\circ}$C) and at a frequency of 400 MHz. For each beam there are eight cavities, each delivering an accelerating field of 5 MV/m. To keep the particles circulating around the ring and inside the beam pipe (vacuum pressure of 10^{-13} atm), a magnetic field of 8.3 T is used. This field is generated by superconducting dipole electromagnets operating at a current of 11.7 KA and at a temperature of 1.9 K ($= -271.3\,^{\circ}$C). Superfluid helium acts as cooling fluid. It has a very high thermal conductivity,

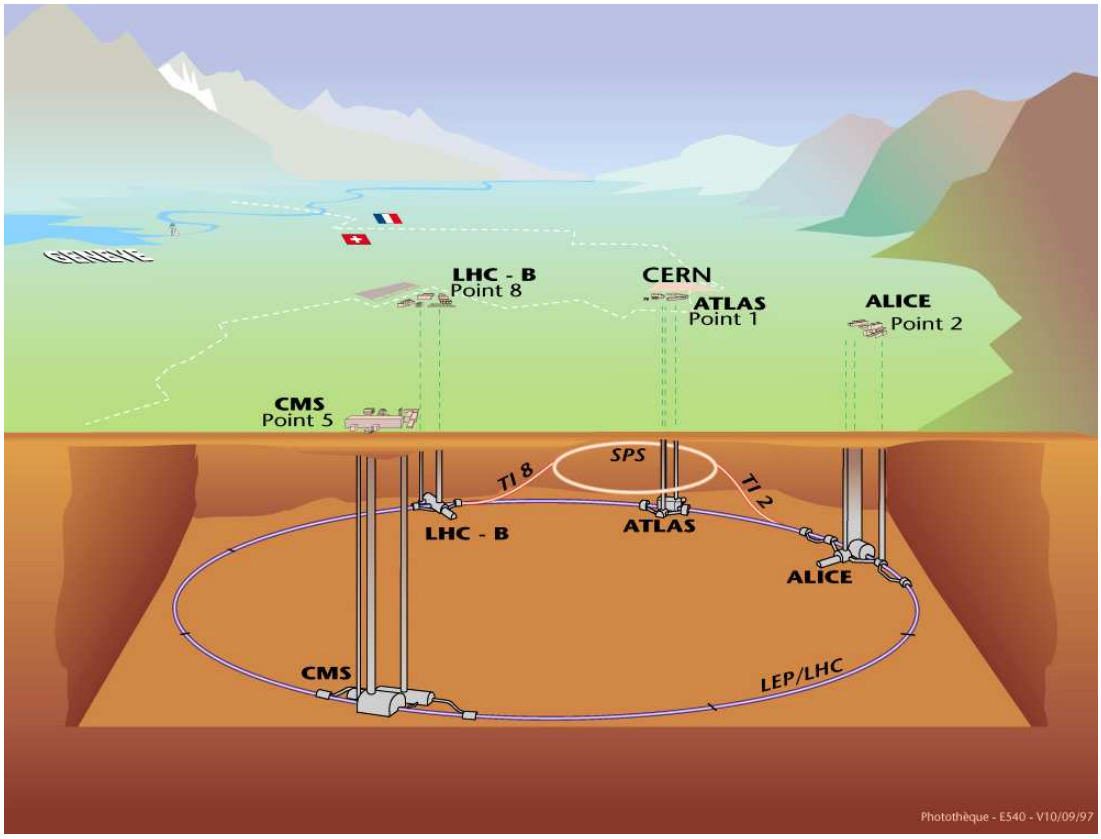

Figure 3.1: Schematic view of the LHC-accelerator and its four experiments. [31]

which is intended to stabilize the large superconducting system. Figure 3.2 shows the cross section of a LHC dipole. In total, there are 1232 dipoles of 15 m length and 35 t weight in the LHC ring cooled by 120 t of helium.

Without further magnetic fields the charged particles repelling themselves would deviate from the ideal trajectory, diverge in the plane transverse to the beam and be absorbed by the beam pipe vacuum vessel. To keep the particles densely together, many other focussing magnets, including 392 quadrupoles, are used. Such magnets are also used around the interaction points, in order to squeeze the particles as much as possible to increase the probability of a collision.

Due to the radio frequency acceleration scheme, the protons circulating inside one beam pipe ring are divided into bunches of about 10^{11} particles. 2808 of these bunches circumnavigate the ring at nearly the speed of light ($v/c = 0.999\,999\,991$). Thus, they traverse the entire ring about 10 000 times per second. The bunches follow each other at separation of about 7.5 m in length or 25 ns in time. This corresponds to a collision rate of ≈ 40 MHz at the interaction points.

The particle current density in the LHC is specified by the machine parameter L, named luminosity. As defined in equation 3.1, the (instantaneous) luminosity

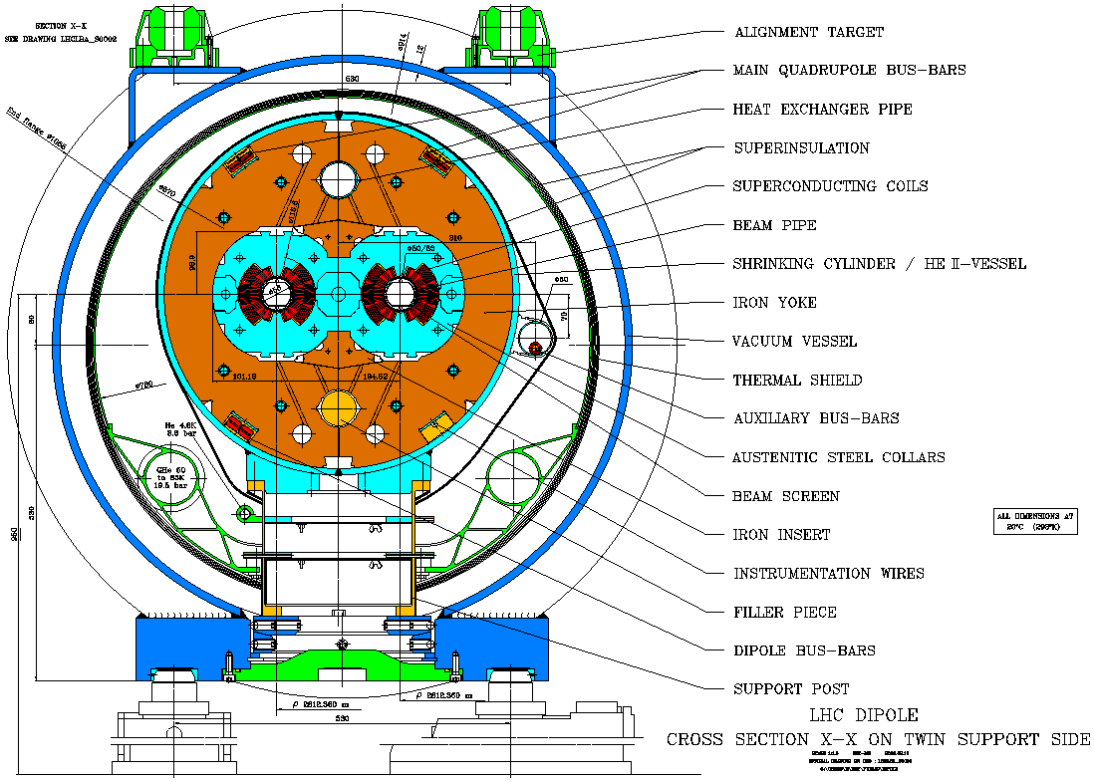

Figure 3.2: Cross section of the vacuum pipes inside a dipole magnet. [32]

L is the product of the numbers of particles n_1, n_2 in both crossing bunches and the frequency f of bunch crossings, divided by the cross sectional area A of a bunch:

$$L = \frac{n_1 \cdot n_2 \cdot f}{A} \tag{3.1}$$

The full formula according to the LHC design (beam profile) can be found in [33].

The time between beam commissioning at injection energy of 450 GeV per beam and the first physics run at design center of mass energy of 14 TeV and design luminosity of $L = 10^{34} \frac{1}{cm^2 \cdot s}$ will be probably more than one year. At this luminosity, there are about 23 inelastic scatterings per bunch crossing, in which nearly 1000 new particles are produced. The rate R of the proton-proton reactions inside the LHC machine is given by the product of the proton-proton cross section σ and the luminosity: $R = \sigma \cdot L$.

In the next section, we present the LHC experiment ATLAS in detail, where such particle collision events are detected and analyzed.

3.2 The ATLAS Experiment

ATLAS is a multi purpose detector [34], that is designed to observe particle collision events within the complete LHC parameter range. ATLAS is about 44 m long, has a diameter of approximately 25 m and a weight of about 7000 t. Due to the design of the LHC as a ring collider, ATLAS has a nearly cylindrically symmetric structure (barrel with 2 endcaps). Figure 3.3 shows a detailed view of ATLAS.

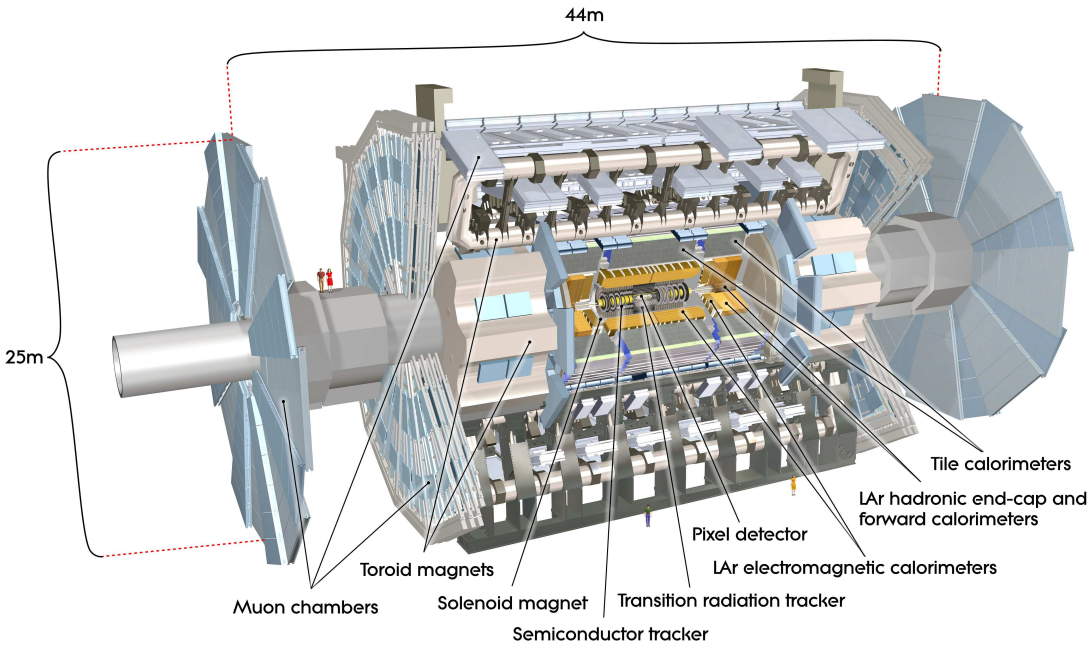

Figure 3.3: Detailed view [35] of the ATLAS detector and its subdetectors

To identify positions within ATLAS, the following coordinate system is used: The beam direction defines the z-axis, and the x-y plane is the plane transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the center of the LHC ring, and the positive y-axis is pointing upwards. [26] The ATLAS side that is in positive z-direction is called A-side, the other side is called C-side. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$. The distance ΔR in the pseudorapitidy-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta^2\eta + \Delta^2\phi}$. [26]
In a LHC collision event many particles are produced. Some types of particles can be directly detected. But since particles differ in their properties and interactions, ATLAS is a combination of different specialized particle detectors: The inner

tracking detector [36], the electromagnetic and hadronic calorimeters [37], and the muon spectrometer [38], that is the outermost detector. The spatial arrangement of the subdetectors and an event cross section is sketched in figure 3.4. The

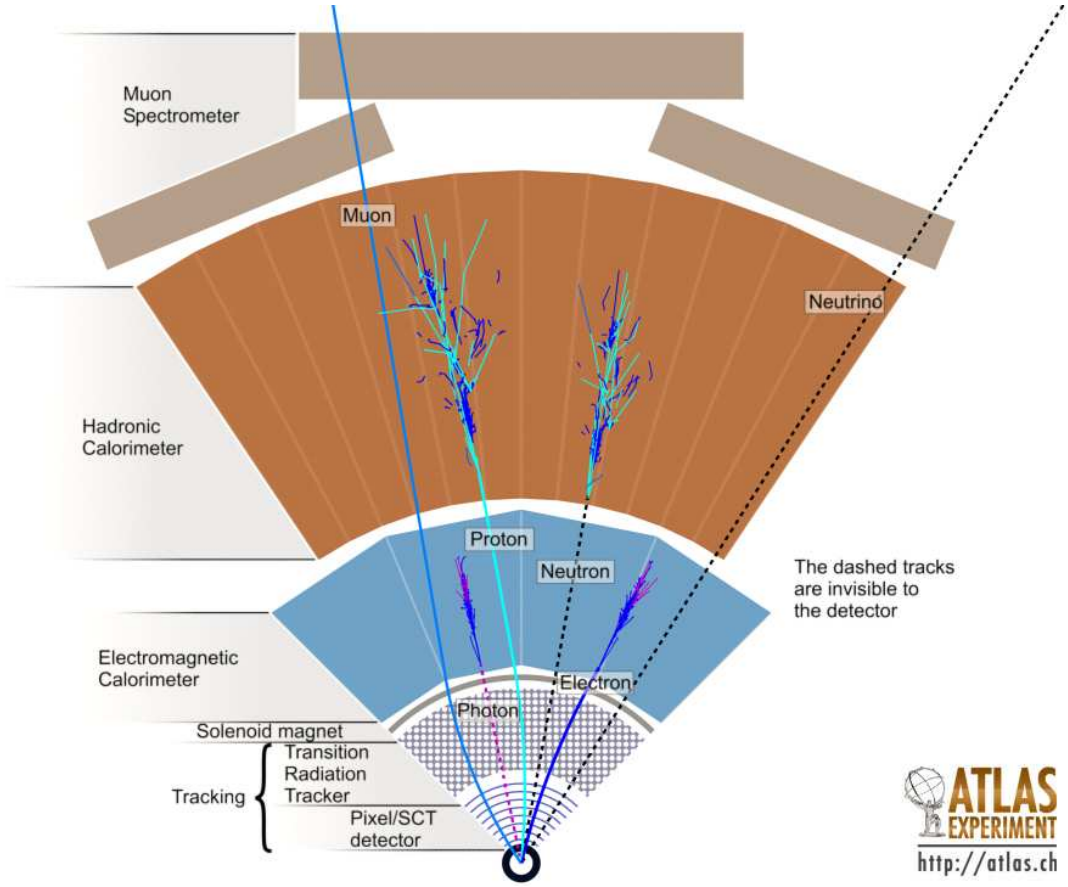

Figure 3.4: Schematical view of the specialized particle detectors of ATLAS and sketch of an event cross section. Not drawn: the toroid magnets between the hadronic calorimeter and the muon spectrometer; the radiation shields. [39]

subdetectors are explained in detail in the following sections. By measuring precisely the directly detectable particles, other particles can be reconstructed: For example, some particles exist after production at the interaction point for such a short time that they do not leave the beam pipe before they decay.

The momenta of charged particles can be measured by applying a magnetic field and reconstructing the trajectories of the particles. In ATLAS, there are four superconducting magnets [40]: one solenoid [41] with a nominal strength of 2 T for the inner detector, and three air-core toroids (one barrel [42], two endcaps [43]) with a peak strength of about 4 T for the muon spectrometer.

Due to the high design luminosity and energy of the LHC beams, a lot of secondary radiation is generated in the detector. To reduce this background

radiation, nearly 3000 t of shielding material [44] has been added to the detector. For the operation and control of the ATLAS detector following systems are needed [45, 46]:

- CERN infrastructure systems including safety, cryogenics, cooling and ventilation, gas, electricity, vacuum
- CERN magnet system
- LHC conditions and beam interlock system (BIS) [47], LHC fast signals [48] and timing, trigger, control system (TTC) [49]
- Finding Persons Inside ATLAS Area system (FPIAA) [50]
- Detector Safety System (DSS) [51]
- Detector Control System (DCS) [52]
- TDAQ: Trigger and Data AcQuisition (DAQ) [53, 54]

The control and data flow of the ATLAS experiment is illustrated in figure 3.5. All communications between the CERN infrastructure, the magnets, the LHC

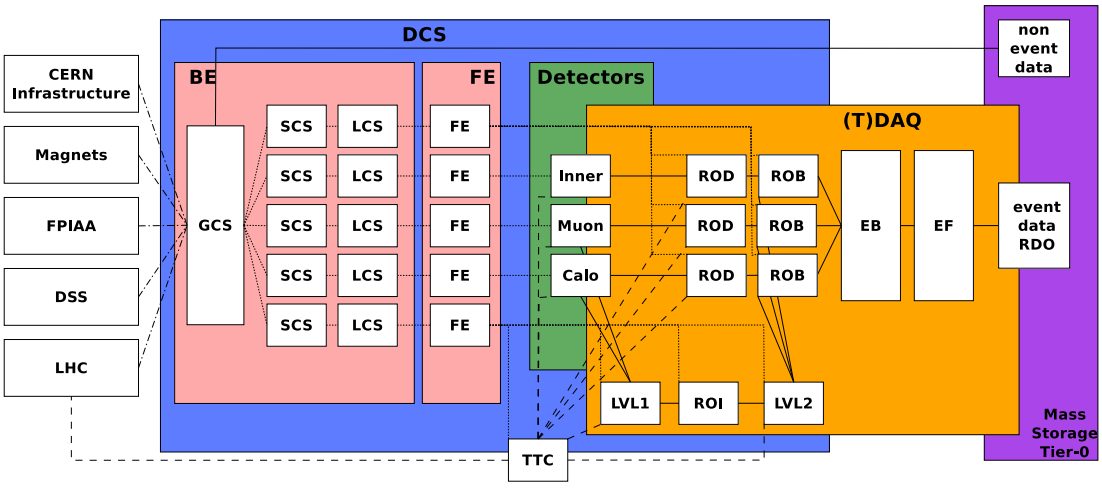

Figure 3.5: ATLAS control and data flow: DCS, (T)DAQ and external services. Dotted lines indicate DCS connections, dashed lines represent timing signals, dashed-dotted lines indicate information from or to external services, solid lines represent trigger or data information. All acronyms are explained in the text.

machine, the FPIAA, the DSS and the TDAQ are done via DCS, with the exception of fast signals (such as the LHC 40.079 MHz bunch crossing clock and 11.246 KHz orbit signal) that are distributed by the TTC system and handled in the level-1 (LVL1) trigger [55]. The DCS has to work continuously to ensure the

safe operation of the detector. The DCS guarantees also a reliable coordination with the LHC control system and with other external services. Even when TDAQ is not available, DCS controls hierarchically all subdetectors, common services and the environment in the cavern.

The DCS is divided into the Front-End (FE) and the Back-End (BE). The BE is organized in three layers, the Global Control Stations (GCS) with server applications and human interfaces in the ATLAS control room for the overall operation, the Sub-detector Control Stations (SCS) for high-level control of a subdetector or of common services (e.g., cooling or TDAQ devices), and the Local Control Stations (LCS) for process control of subsystems. The LCS controls all devices in the cavern via the FE. The FE is based on Controller Area Network (CAN) field buses deployed over the whole experiment and uses Embedded Local Monitor Boards (ELMB), that are insensitive to the magnetic fields and tolerant to radiation. The DCS stores all ATLAS detector conditions data in a relational database, which is important for the offline reconstruction and analysis.
As illustrated in figure 3.6, the LHC beam has to be stable and the DCS must

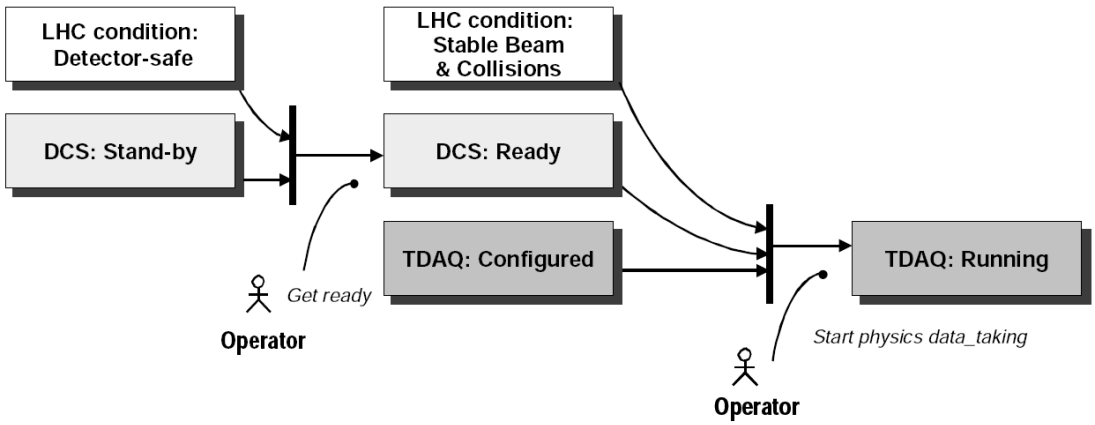

Figure 3.6: ATLAS can be operated, if LHC, DCS and DAQ have status safe and ready. [53]

have enabled the detector and the TDAQ, so that data can be taken.

The TTC system synchronizes ATLAS and LHC, so that the hit data of all subdetectors can be assigned to the correct bunch crossing. The default bunch crossing rate is $\approx 40\,\text{MHz}$, and at design luminosity every candidate event for new physics will be accompanied by 23 inelastic events on average. However, the event data recording is limited to about 200 Hz because of technology and resource limitations. Therefore, an overall rejection factor of $5 \cdot 10^6$ against minimum-bias processes is required while maintaining maximum efficiency for the new physics. The selection (trigger) of events and their transfer to the storage (DAQ) is done by the TDAQ.

The trigger system consists of three levels of event selection: level-1 (LVL1) trigger, level-2 (LVL2) trigger and event filter (EF). The LVL2 and EF together form the High-Level Trigger (HLT). The LVL1 trigger is implemented using custom-made electronics, while the HLT is almost entirely based on commercially available computers and networking hardware.
The LVL1 trigger uses a subset of the total detector information (parts of the calorimeter and of the muon spectrometer) to make a decision on whether or not to continue processing an event. The LVL1 trigger reduces the data rate to approximately 75 KHz (limited by the bandwidth of the readout system, which is upgradeable to 100 KHz). If an event is accepted, a LVL1 accept signal (L1A) is fed into the TTC system. When the L1A signal reaches the Read Out Drivers (ROD) of the subdetectors, all hit data of the requested event are read out and transferred to the Read Out Buffers (ROB). Furthermore, the LVL1 trigger identifies regions in the detector where it has found interesting features, so-called Regions of Interest (RoI). The RoI information from the various parts of the LVL1 trigger are combined by the RoI Builder (RoIB) and then passed to the LVL2 trigger.
For those events fulfilling the LVL2 selection criteria (maximal acceptance rate: 3.5 KHz), event-building (EB) is performed. The assembled events are then moved by the DAQ to the EF, and the events selected there are moved to permanent event storage (200 Hz) at CERN. Each event has a total size of about 1.5 Mbytes and is stored in a Raw Data Object file (RDO). Hence, several petabytes of raw data will be stored at CERN per year.

This data is analyzed [56, 57] via the LHC Computing Grid (LCG) [58, 59].

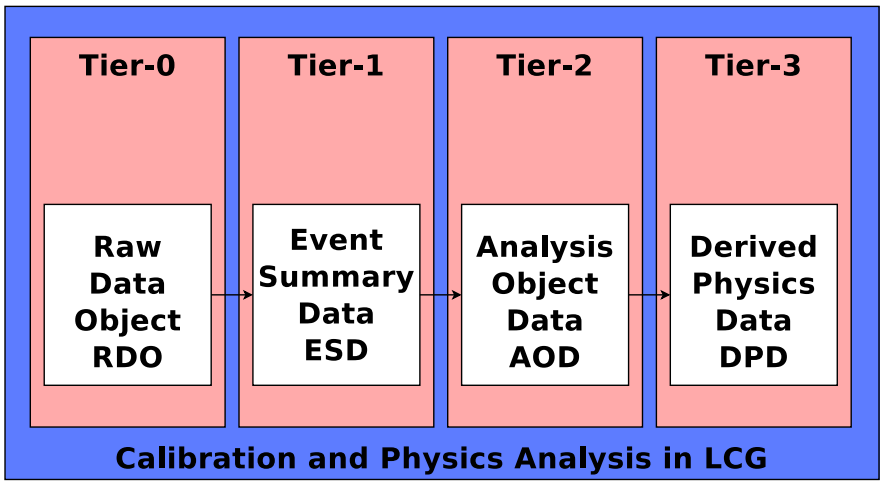

Figure 3.7: Calibration and physics analysis in the LCG-GRID.

According to a four-tiered model [60] illustrated in figure 3.7, the data is distributed around the globe. A primary backup will be recorded on tape at CERN,

the Tier-0 center of the LCG. After primary event processing at the Tier-0, the raw and the primary processed data is copied to a series of Tier-1 centers. These large computer centers have sufficient storage and reprocessing capacity with round-the-clock support for the Grid. At the Tier-1 centers, derived datasets are produced as Event Summary Data files (ESD) or Analysis Object Data files (AOD). For analysis, this data is distributed to the Tier-2 facilities, that are accessed by individual scientists, who use Tier-3 facilities, such as a workstation. For quick event access and histogramming, Derived Physics Data files (DPD) can be produced out of the AODs at the Tier-2 centers and stored at the Tier-3 facilities.

For calibration and analysis an object-oriented approach to software is used, based on the programming language C++ [61, 62]. Other languages are also used, including FORTRAN [63], Java [64] and Python [65]. The Athena framework [66], that has a component-based architecture, allows one to perform all needed activities such as: calibration and alignment, simulation (event generation, detector simulation, digitization, pile-up) and analysis (event selection, reconstruction, physics analysis).

3.2.1 The Muon System

Final state muons are very important in many analyses and provide robust signatures of physics at the LHC, for example the Higgs boson decay $H \rightarrow ZZ^* \rightarrow 4\mu$ or a semileptonic channel of a decaying top/antitop quark pair $t\bar{t} \rightarrow 4j + \not{E}_T + \mu$. Therefore, the muon spectrometer has a high resolution and a stand-alone triggering momentum measurement capability. Since muons penetrate the large amount of material where all other particles are absorbed, in order to measure their energy, the muon spectrometer [38] is the outermost subdetector of ATLAS. As explained above, the muon spectrometer measures the charge and momentum of muons by reconstructing their trajectories in the magnetic field of the air-core toroids. The barrel chambers form three cylinders concentric with the beam axis and cover the pseudorapidity range $|\eta| < 1$. The outer barrel chambers are 10 m away from the beam pipe, their half length is 12.5 m. The forward chambers cover the range $1 < |\eta| < 2.7$ and the outermost are about 21.5 m from the interaction point. In order to provide a resolution of $\sigma(p_T)/p_T = 10\%$ at $p_T = 1\,\text{TeV}$, the space resolution must be at least 50 μm in z and 0.5 mrad in ϕ.

Several types of muon tracking chambers exist: Monitored Drift-Tube chambers (MDT) and Cathode Strip Chambers (CSC) are precision chambers, Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) are fast LVL1 trigger chambers. The principle of measurement is the same for all four types: Muons, that cross a gas gap between an anode and a cathode (for example a wire inside a tube or two parallel plates), cause a local gas discharge and hence an electric signal, that can be read out. In total, there are 5376 chambers or $1.0757 \cdot 10^6$ readout channels.

The MDTs provide a precision measurement (mechanical accuracy of 30 μm) of the track coordinates (80 μm single wire resolution) over most of the η-range. At large pseudorapidities and close to the interaction point, CSCs are used, that have a higher granularity and withstand the demanding rate and background conditions.
The trigger chambers have a time resolution better than the LHC bunch spacing of 25 ns to identify the bunch crossing. The requirements on the time and space resolution are met by following specialization: The RPCs have the better time resolution (1.5 ns), whereas the TGCs have the better spatial resolution (anode wire pitch 1.8 mm).

Over the large global dimensions of the muon spectrometer, it is not possible to stabilize the dimensions and positions of the chambers at the required 30 μm level. Therefore, chamber deformations and positions are constantly monitored by means of optical alignment systems. As a result, displacements between $\approx 30\,\mu$m and ≈ 1 cm can be corrected for in the offline analysis [26].

3.2.2 The Calorimeters

The ATLAS calorimeters [37] are crucial in many measurements, for example for the Higgs boson decay $H \rightarrow \gamma\gamma$ or the semileptonic electron channel of a decaying top/antitop quark pair $t\bar{t} \rightarrow e + \not{E}_{\mathrm{T}} + 4j$. In order to identify events, the calorimeters must separate well electrons/positrons, photons and jets. When electrons/positrons or photons enter the calorimeter, an electromagnetic shower is initiated. The shower breaks down and is absorbed, when secondary photons fall below the pair production threshold due to energy losses. Whereas hadrons, which are the main constituents of jets, rather initiate hadronic showers than electromagnetic ones. Since both types of showers have different characteristics, for example electromagnetic showers begin a shorter distance into the calorimeter than hadron showers, the origin of a shower may be determined.
Hence, the calorimeters consist of an outer hadronic calorimeter (HCAL) and an inner electromagnetic calorimeter (ECAL), both covering the pseudorapidity region $|\eta| < 3.2$. A dedicated dense liquid argon forward calorimeter (FCAL) with rod shaped electrodes in a tungsten matrix covers the range $3.1 < |\eta| < 4.9$, where the level of radiation is very high. The total weight of the calorimeter system including the solenoid flux return iron yoke is nearly 4000 t.

Since the calorimeters contribute to the selection of events at the LVL1 trigger, their corresponding parts have a fast response time of few nanoseconds. The ECAL is dedicated to the measurement of energy and position of electrons and photons (and separation of γ/π^0, e/π, etc.). Whereas, the HCAL measures the energy and position of jets and separates hadronic decays of τ-leptons from jets. Both calorimeters, especially the HCAL, contribute to the reconstruction of the missing transverse momentum of the event. The region $1.37 < |\eta| < 1.52$ is not used for precision physics measurements involving photons because of the large

amount of material (about 7 radiation lengths) situated in front of the ECAL (cryostats).

The HCAL is a scintillation calorimeter. Its barrel consists of plastic scintillator plates (tiles), in which light flashes are produced, when high energetic charged particles (or photons) traverse the plates. These are embedded in an iron absorber and read out by photomultipliers. This tile calorimeter system [67] has an outer radius of 4.25 m and a total length spanning of ±6.10 m. At larger rapidities, where higher radiation resistence is needed, the intrinsically radiation hard liquid argon technology is used. Due to the high granularity ($\Delta\eta \cdot \Delta\phi$) between $0.1 \cdot 0.1$ and $0.2 \cdot 0.2$ (representing about 10^4 readout channels) the HCAL provides a precise measurement of the three-dimensional jet direction (from the direction of the shower) and a good jet resolution. By also measuring the shower energy with the calibrated photomultipliers the four-momentum of a jet may be obtained. In the range $|\eta| < 3$, the energy resolution is $\frac{\Delta E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$. Moreover, the jet energy scale is linear within 2% up to an energy of 4 TeV. In order to absorb all particles except muons in front of the muon spectrometer, the HCAL has a thickness of about 10 interaction lengths. This leads to the so called shower containment inside the HCAL and to a reduction of background in the muon chambers.

The ECAL is a lead/liquid argon (LAr) detector with accordion geometry. Together with the hadronic endcap calorimeter (HEC), and the FCAL it forms the Liquid Argon Calorimeter [68], which are types of ionization calorimeters. The whole LAr calorimeter has a length of ±6.65 m along beam axis and an outer radius of 2.25 m, giving a total thickness of about 25 radiation lengths. When high energetic charged particles (or photons) traverse the liquid argon, argon atoms are ionized and free electrons are produced. In an electrical field, that is generated by an applied high voltage, the electrons drift to the cathode and induce a current there, which is read out. By an appropriate calibration the shower energy and hence the energy of the high energetic particle is measured.
In order to identify (electron/photon/tau – jet separation) and resolve particles well, the granularity ($\Delta\eta \cdot \Delta\phi$) of the ECAL is between $0.003 \cdot 0.1$ and $0.05 \cdot 0.025$, resulting in a total number of readout channels of about $1.9 \cdot 10^5$. The ECAL has an energy resolution of about $10\,\%/\sqrt{E[\mathrm{GeV}]}$ and can reconstruct electrons from 1 GeV to 5 TeV. The energy scale precision is better than 0.1%, and up to 300 GeV the linearity of response is better than 0.5%. Within 10 years of operation the ECAL has to withstand neutron fluences of up to $10^{15}\,\mathrm{n/cm^2}$ and radiation doses of up to $2 \cdot 10^5$ Gy.

The amount of material in front of the muon system, that includes the support structure of the Tile calorimeter, is about 11 interaction lengths. Whereas, the total material (Inner Detector and services) seen by an incident particle up to the ECAL front face is circa 2 radiation lengths.

3.2.3 The Inner Detector

The Inner Detector (ID) [36] is very important for physics analyses, because it is the main responsible for reconstructing tracks, (primary) vertices and track impact parameters as well as for electron identification (between 0.5 GeV and 150 GeV). Especially its contribution to τ- and heavy flavor (c-, b-) tagging is essential for precision measurements such as of the top quark mass, as we will see in Chapter 6. The inner detector is a tracking detector with a length of 7 m and a radius of 1.15 m, a quarter section of it is illustrated in figure 3.8. The ID has

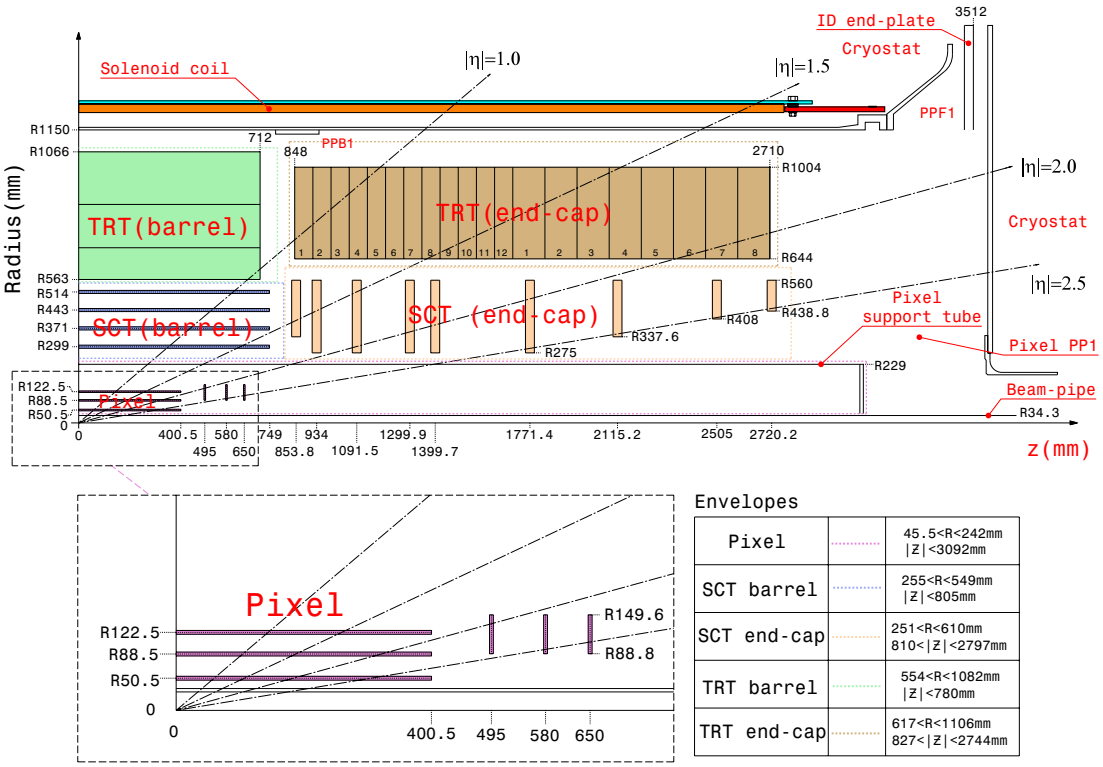

Figure 3.8: Plan view of a quarter-section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes. The labels PP1, PPB1 and PPF1 indicate the patch-panels for the ID services. [34]

a barrel-disk structure like the other detectors and it measures the momentum of tracks in the pseudorapidity range $|\eta| < 2.5$.
The ID consists of three subdetectors: the Transition Radiation Tracker (TRT) [69], the SemiConductor Tracker (SCT) [70] and the Pixel detector [71]. Due to the size and arrangement of the barrel layers and disks of the pixel and SCT detector, typically seven space points are provided by the semiconductor detectors. The TRT measures usually 36 tracking points. The resolutions of the subdetectors are given in table 3.1.

Similar to the muon chambers the TRT uses straw tubes (aluminum coated polyamid) with a length of 1.44 m (barrel tubes) and a diameter of 4 mm. Due to this small diameter and the isolation of the sense wires (30 μm diameter, gold-plated tungsten-rhenium) within the individual gas volumes, the TRT can operate at the very high rates expected at the LHC. By employing xenon gas it is also possible to detect transition radiation photons created in a radiator (polypropylene-polyethylene) between the straws. This technique is intrinsically radiation hard and allows the identification of electrons as well as a large number of measurements, typically 36, on every track. The resolution per straw is $\approx$ 130 μm. The TRT consists of 96 modules in the barrel and 20 modules per end cap, comprising about $3.51 \cdot 10^5$ readout channels in total.

Subdetector	$\sigma_{R-\phi}$ [μm]	σ_z [μm]	readout channels [10^6]
Pixel	10	115	80.4
SCT	17	580	6.3
TRT	130 per straw		0.351

Table 3.1: Key numbers of the ATLAS inner detector.

The SCT detector is made up of 4 barrel layers and 9 disk layers per end cap. In total, there are 4088 modules consisting of about $6.3 \cdot 10^6$ readout channels. The SCT detector contains about 63 m^2 of small ($6.36 \cdot 6.40$ cm^2) silicon microstrip detectors, that consist of 768 readout strips of $\approx$ 80 μm pitch each. When a charged particle traverses a strip, free electrons and holes in a low doped fully depleted silicon wafer are produced. Some electrons and holes are collected at electrodes and induce there the readout current.
Since always two small microstrip detectors are arranged back-to-back, the SCT provides on average 8 precision points in the $R\phi$ and z coordinates (4 space points). In order to obtain the z measurement, small angle (40 mrad) stereo between both small detecors is used. Tracks can be distinguished, if they are separated by more than 200 μm. The resolution in $R\phi$ is 17 μm, in z 580 μm.
Like the pixel detector, the SCT requires very high dimensional stability and cold operation. By cooling, the radiation damages are reduced, that will alter the fundamental characteristics of the silicon wafers.

The Pixel detector is the innermost subdetector of ATLAS and therefore determines mainly the impact parameters of tracks. This detector consists of 3 barrel layers and 3 disk layers per end cap, comprising 1744 modules or about $8 \cdot 10^7$ readout channels in total. The $R\phi$ resolution is 10 μm, the z resolution is 115 μm.
Since a part of this thesis was the participation in and support of production, installation and commissioning of optoelectronics for this detector, more details about it and about that work are provided in the next chapter.

Chapter 4

Optoelectronics of the Pixel Detector System

4.1 The Pixel Detector

As already mentioned in Section 3.2.3, the pixel detector is the innermost subdetector of ATLAS. The innermost layer of the pixel detector is 5 cm away from the beam pipe. Due to this proximity and its very good point resolution of $\sigma_{r-\phi} < 15\,\mu$m and $\sigma_z < 120\,\mu$m, mainly the pixel detector determines the track impact parameter. Furthermore, the pixel detector is the main responsible for the resolution of primary and secondary vertices. The determination of track impact parameters and vertices is very important for many physics analyses. This is especially true for top quark pair events, if b-tagging is used. More details about this and a procedure for the calibration of track impact parameters will be presented in Chapter 6.

The pixel detector consists of 3 barrel layers and 3 disks on each forward side, as illustrated in figure 4.1. By this geometry, the detector acceptance is within the pseudorapidity range of $|\eta| < 2.5$. As a consequence of the proximity to the beam pipe, the pixel detector must be so radiation hard that even at the end of its lifetime it delivers 3 spacepoints for each traversed charged particle. Moreover, there are three further requirements of the pixel detector: First, the readout has to be at least as fast as the level-1 trigger frequency of 75 KHz. Second, pixel hits must be correctly assigned to the corresponding bunch crossing (bunch crossing rate: ≈ 40 MHz). Third, in order to minimize the error of measurement of the calorimeters, radiation and interaction length of the pixel detector have to be minimal.

The barrel layers consist of staves that have a width of 16 mm and a length of 830 mm. In order to minimize the interaction and radiation length of the pixel detector, the support structures of the staves are made of carbon fiber. On the far side (seen from the beam pipe) of a stave, there is a cooling pipe so that the

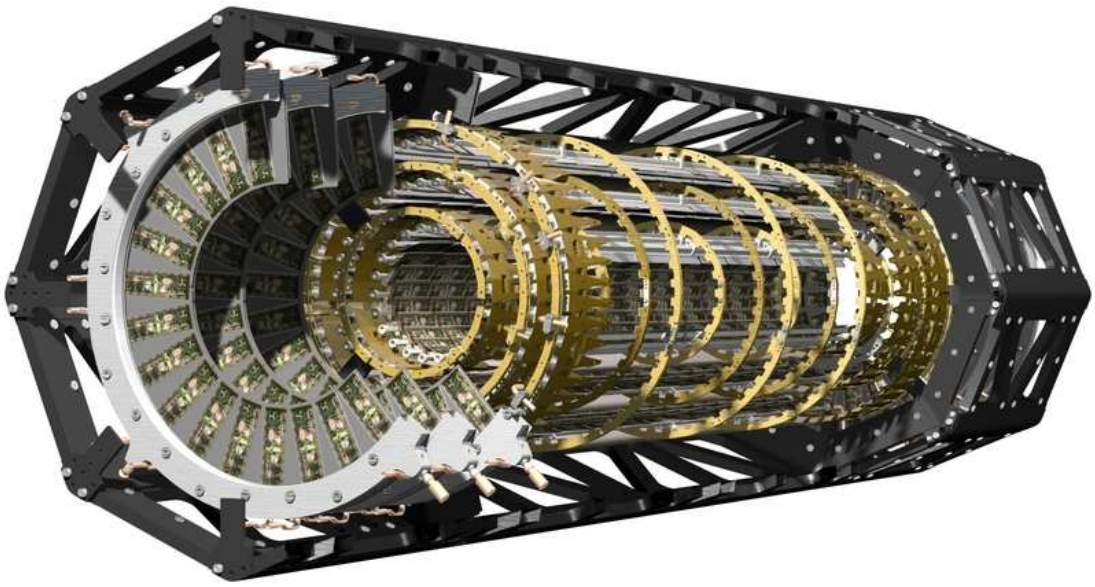

Figure 4.1: A schematic view of the pixel detector in its mounting structure. [72]

electronics is kept at the operation temperature of $-10\,^{\circ}\mathrm{C}$ and radiation damage is minimized. On the near side, there are the active elements, the pixel modules. On each stave, there are 13 modules, each with 47232 pixels as sensors [73, 74]. For reasons of space, there are 4 ganged pixels in each column of the front-end chip, thus leading to a total of 46080 readout channels. About 90% of all pixels have an area of $50\,\mu\mathrm{m} \times 400\,\mu\mathrm{m}$ each, the other have an area of $50\,\mu\mathrm{m} \times 600\,\mu\mathrm{m}$ each.

On the other hand, a disk, that has an inner radius of 89 mm and an outer radius of 150 mm, consists of 8 sectors. Each sector holds 6 modules (3 on the front, 3 on the back), which are identical to those on the staves. Table 4.1 provides

	Radius (mm)	**Staves**	**Modules**
Layer-0	50.5	22	286
Layer-1	88.5	38	494
Layer-2	122.5	52	676
	z-Position (mm)	**Sectors**	**Modules**
Disk-1	±495	8/8	48/48
Disk-2	±580	8/8	48/48
Disk-3	±650	8/8	48/48
Total			1744

Table 4.1: Dimensions and coordinates of the ATLAS pixel detector subparts.

the dimensions and coordinates for the subparts of the pixel detector. In total, there are 1744 modules or 80 363 520 readout channels, which represent more than

90% of all readout channels of ATLAS. As a consequence, it is very vital that the pixel detector is carefully and successfully operated as well as calibrated.

Figure 4.2 shows the structure of a pixel module. Each pixel is connected to a single electronic cell via a tiny solder bump (bump bond). When a charged particle traverses a pixel, free electrons and holes in the n^+ on n type sensor are produced. Over the bump bond, a capacitor is charged, which delivers the readout current.

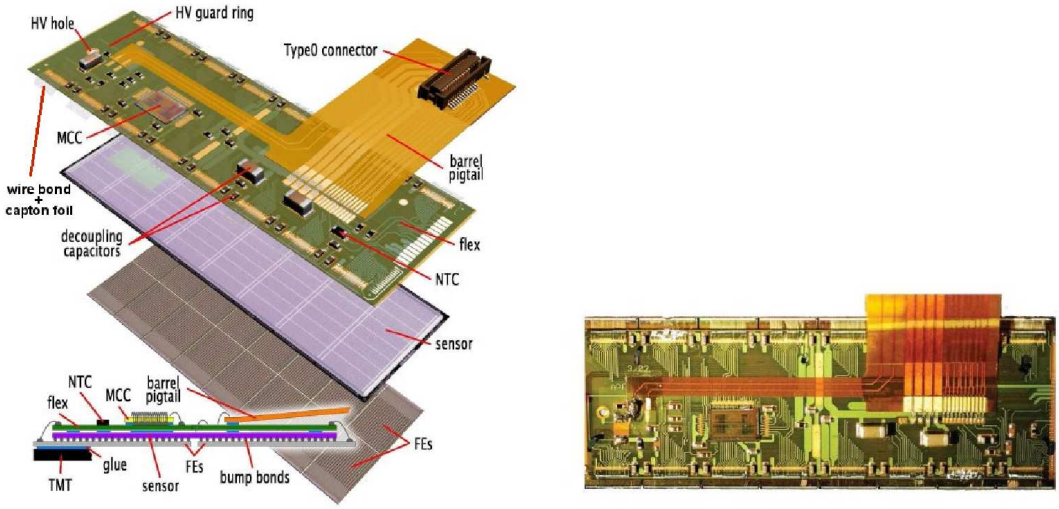

Figure 4.2: At the left: scheme of a pixel module. The high voltage (HV) for the sensor depletion is led through a hole in the flex to the pixels. A guard ring prevents sparkovers. For the thermal contact to the cooling pipe the pixel modules are glued onto the carbon-carbon thermal management tiles (TMT). More information can be gained from the text. At the right: top view photo of a pixel module. The top side points to the beam pipe. [34]

2880 electronic cells form one special ASIC, the Front-End (FE) chip [75]. On each module there are 16 FE chips, that amplify and digitize the signals from the sensors. The FE chips are connected via 25 μm thick wirebonds, held by capton foil, to the flexible copper-capton board, the "flex".
Several components are on the flex, including a negative temperature coefficient (NTC) temperature sensor, decoupling capacitors and the Module Control Chip (MCC) [76]. The MCC configures the 16 FE chips, sends to them received trigger signals and reads their hit data. These electrical trigger signals and hit data come to or leave the module via the "pigtail" and its Type0 connector. The electrical signals come from or go to the optoboard (OB). More about the readout of the pixel modules will be explained in the next section.

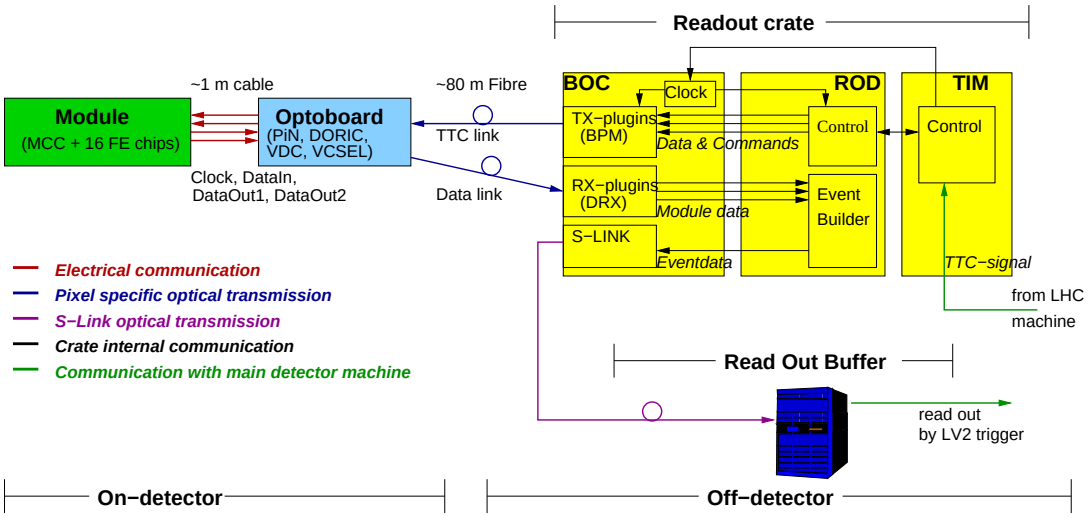

Figure 4.3: The ATLAS pixel dectector readout chain. [11]

4.2 The Pixel Detector Readout: The Optical Link

In figure 4.3, the readout chain of the ATLAS pixel detector is shown. It consists of two parts: the on-detector part, that is located inside the ATLAS detector in the cavern UX15, and the off-detector part, which is located outside the ATLAS detector in the cavern USA15. Due to the high data rate and for an electrical separation of both parts, the on- and off-detector sides are connected by fiber optical cables. Therefore, on each side an opto-electrical interface is needed.

As indicated in the previous Section 4.1, on the on-detector side the electrical signals from a module are converted by the optoboard (OB) [77] to optical signals and sent to the off-detector side. There, the Back Of Crate (BOC) card [78] receives the optical signals and converts them to electrical signals. Vice versa the OB converts optical signals coming from the BOC card to electrical signals and sends them to the MCCs. 1 BOC card is connected by long glass fiber ribbons to up to 4 OBs. On the BOC side, each ribbon consists of 8 fibers (on the OB side, two ribbons of 8 fibers each form one ribbon of 16 fibers). Each fiber represents the connection to 1 MCC. The data transmission is done unidirectional. This means that for each MCC 2 fibers (3 for a Layer-0 MCC) are necessary, one for sending commands to the MCC and one (two for a Layer-0 MCC) for receiving data from the MCC. All 8 fibers of 1 ribbon are either for sending data to or for receiving data from MCCs.

Each BOC card is connected back to back to a Read Out Driver (ROD) [79] in a 9U VersaModule Eurocard (VME) crate. As the name indicates, the Back of Crate card is located on the back side of the ATLAS pixel readout crates.

Each crate can hold up to 16 RODs/BOC cards and contains 1 Timing, Trigger and Control (TTC) Interface Module (TIM) [80] and 1 Readout Crate Controller (RCC). This is a Single Board Computer (SBC) and steers the cards in the crate. The TIM receives the optical TTC signal. It sends e.g. a Level-1 Trigger "accept" signal (75 KHz) electronically to the RODs. The TIM sends also a clock signal to the BOC cards. The signal has a frequency of ≈ 40 MHz, which corresponds to the bunch crossing rate, and it is tuned on the BOC card to compensate signal propagation delays.

The BOC card receives the readout commands from the ROD and encodes them with the clock into one Bi-Phase Mark (BPM) [77] signal. This encoding reduces the number of needed fibers and opto-transmitters/receivers. The electric BPM signal is converted to an optical signal and sent to the OB. The OB receives the optical signal and converts it back to an electrical signal. Then, the BPM signal is decoded, and clock and command data are restored. Both are sent separately from the OB to the modules.

Likewise, the OB receives the data from the MCCs and converts it to optical signals. The BOC card receives these signals, converts them to electrical data and sends it to the ROD. Here, the data format is checked, and a reformatting of the data is performed. Thereafter, the data is processed by the Event Fragment Builder. Afterwards, the event fragment data is sent to the BOC card, that converts the electronic signals to optical signals and sends them to the Read Out Buffer (ROB). This buffers event data with a frequency of 75 KHz until a Level-2 Trigger "accept" signal arrives (2 KHz). The further steps in the trigger and data acquisition chain were explained above in Section 3.2.

More details about the optical link can be found in [81]. The BOC card and the optical cables, their production, installation and commissioning, are described in the following section.

4.3 Production, Installation and Commissioning of Optoelectronics

4.3.1 Back Of Crate Cards

In the previous section, the BOC card, its position in the readout chain of the ATLAS pixel detector and its main tasks were introduced. This section summarizes the production, installation and commissioning of the BOC cards in detail. The main responsibility for these tasks was taken by the Wuppertal group.

In total, 132 BOC cards are needed for the ATLAS pixel detector. The cards are placed in 9 readout crates in the 5 racks Y25-05A2 to Y29-05A2 in the ATLAS counting room USA15.

The BOC card is a 9U VME 64X transition card, which was developed by the group of the Cavendish Laboratory in Cambridge [82]. The card was designed

for the SCT and the pixel detector, and the board design is the same for both. However, a section on the BOC card is unused and not assembled, if it is employed for the pixel detector. This is because of the different detector design. Figure 4.4 shows a photo of a pixel BOC card.

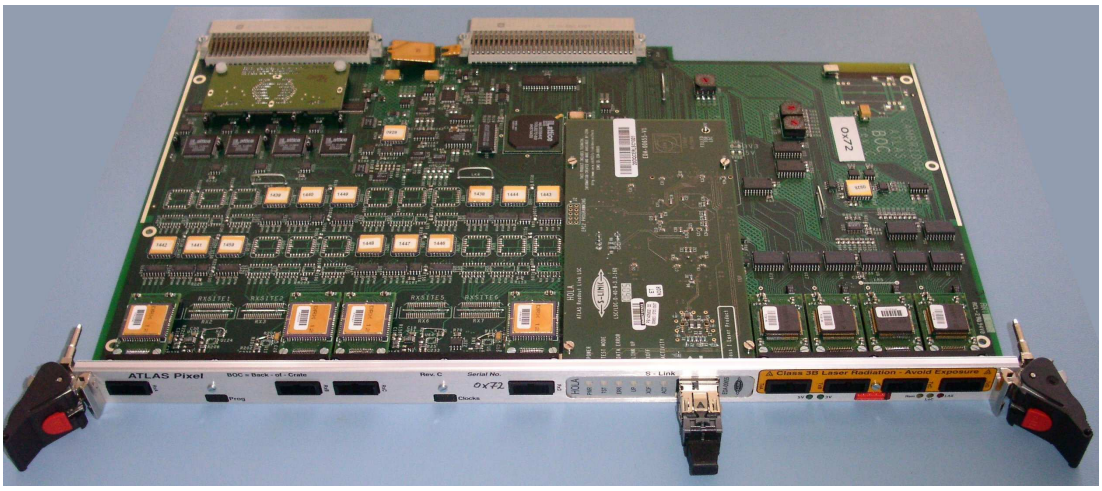

Figure 4.4: A Back Of Crate Card card in pixel assembly (Layer-2). [11]

Each pixel BOC card can serve up to 32 pixel modules (channels), but due to modularities not all channels of a BOC card will be used. Furthermore, the number of used channels depends on which part of the pixel detector the BOC card has to support. For Layer-0, one BOC card is connected to one OB by 3 optical 8-way ribbons (these are described in Section 4.3.2), 1 for the transmission of clock and commands to the OB, 2 for the transmission of data to the BOC card. For Layer-1 and Disks, one BOC card is connected to 2 OBs by 2 optical ribbons each, 1 for clock/commands, 1 for data. For Layer-2, one BOC card is connected to 4 OBs by 2 optical ribbons each, 1 for clock/commands, 1 for data. The BOC card can be operated in 40 or 80 MBit/s mode, determining the data transfer rate from the OB to the BOC card. For Layer-2 this rate will be 40 MBit/s, for Layer-1 and Disk-1,2,3 it will be 80 MBit/s. The 160 MBit/s rate for the Layer-0 is achieved by two connections each operated at 80 MBit/s. The numbers of required BOC cards for each configuration are shown in table 4.2.

The BOC card has to provide the following functionalities:

- Receive electrical commands from the ROD
- Receive electrical clock from the TIM
- Convert electrical commands and clock to Bi-Phase Mark encoded optical signals and send them to the OB(s); provide masking, timing and laser current adjustment functions for these signals

Detector part	OBs	BOC cards	bit data rate
Layer-0	44	44	160 MBit/s
Layer-1/Disk-1,2,3	76 + (16+16+16)	38 + (8+8+8)	80 MBit/s
Layer-2	104	26	40 Mbit/s
total	272	132	

Table 4.2: Number of BOC cards to read out the complete ATLAS pixel detector, split up for the different detector parts.

- Receive raw optical data from the OB(s) and convert them to raw electrical data; provide a timing and threshold adjustment and clock synchronization for these signals
- Send raw electrical data to the ROD
- Receive electrical event fragment data from the ROD
- Convert electrical event fragment data to optical data and send them via the S-Link path to the ROB
- React to the laser safety interlock

The layout of the BOC card and those tasks, which represent a signal flow on the BOC card, are sketched in figure 4.5. Several functionalities are separated into several sections on the board: the 2 VME bus connectors (1 for input signals, 1 for output signals) to the backplane; the main Complex Programmable Logic Device (CPLD), that controls the board; the clock section, where the timing can be tuned; the receive section with the mounted RX-plugins, which convert the incoming optical signals to electrical signals via PIN diodes (P-type, Intrinsic, N-type); the command and transmission section with the mounted TX-plugins, that encode the clock and the commands from the ROD with the BPM-chip [83] into one signal, which is converted by the Vertical Cavity Surface Emitting Laser (VCSEL) [83] arrays into the outgoing optical signals; the S-Link [84, 85] section with the mounted S-Link card, or High speed Optical Link for ATLAS (HOLA); the 3 voltage lines (not shown in the figure), one with 3.3 V, one with 5.0 V and one with 12.0 V; the 2 hexadecimal rotari switches (not shown in the figure), that define the serial number of the BOC card. More information about the design and the functionality of the BOC card can be found in [11].

The main assembly of the BOC circuit boards was done by the Turck company [86] in Germany. It delivered 156 BOC circuit boards, leaving 24 as spare boards. The postproduction at the University of Wuppertal consisted of 6 steps for every BOC circuit board:

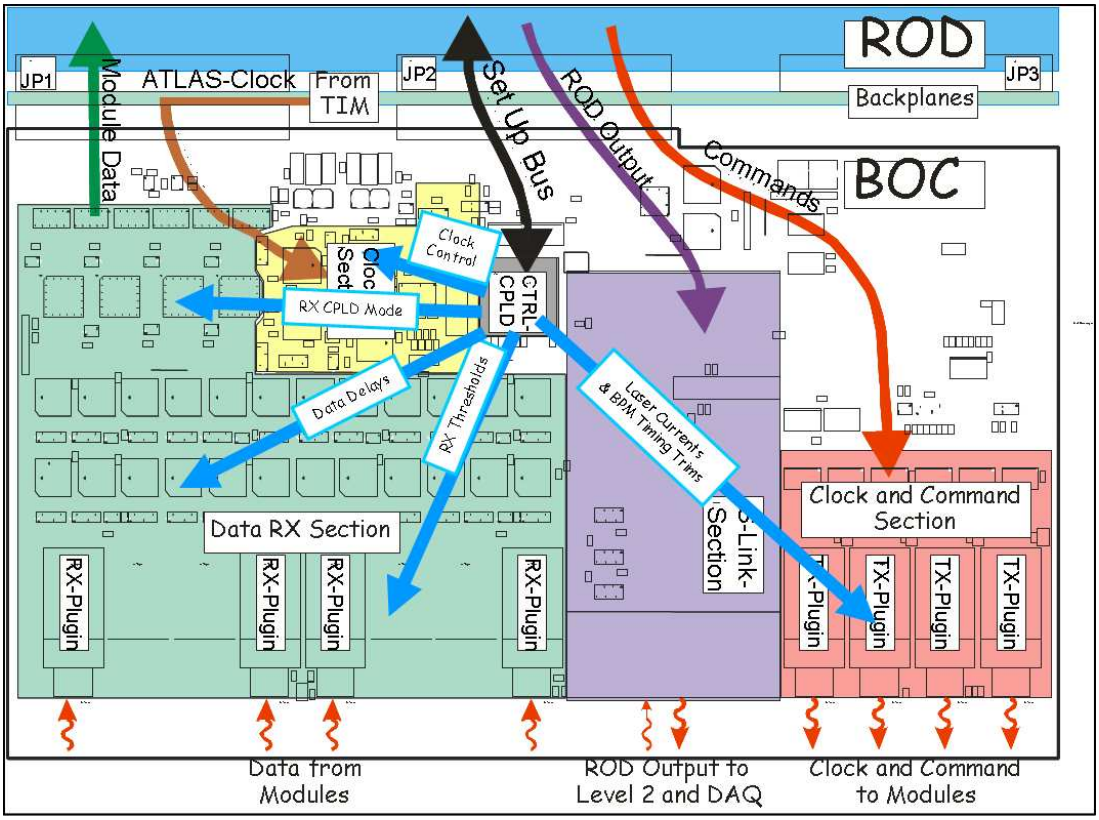

Figure 4.5: Sketch of the signal flow on the Back Of Crate card (Layer-2 assembly arbitrarily chosen). [87]

1. Optical Inspection, Assembly
2. Electrical Power Test/Programming
3. Main Electrical Test (E-Test I)
4. Mechanical Assembly
5. Optical Test (O-Test), Assembly
6. "S-Link" Test (E-Test II)

The numbers of required components for the postproduction in Wuppertal can be found in table 4.3. In step 1, the delivered board was inspected optically. If a problem was found, the board was sent back to the Turck company for repair. Common problems were: The packing was improper so that a board was slightly damaged (e.g. scratches) during transport. Components were shifted or rotated. There were missing or excess components. Some components were defect. Often there were shorts at adjacent soldering joints. Sometimes, cold soldering joints

component	total number of pieces	number of pieces needed for ATLAS
BOC Circuit Boards	156	132
Front-Panels	156	132
Routing Boards	156	132
S-Link-Cards	156	132
TX-Housings/RX-Housings	156/312	132/264
TX-Plugins/RX-Plugins	315/366	272/316
Washers/Screws	$\approx$2000/6000	$\approx$1700/5100

Table 4.3: Numbers of BOC card components.

were found. If the board was accepted, it was serially numbered, and a small routing board and handles were assembled.

In step 2, an electrical power test was done. Here, the voltages and currents of the 3 power lines of the board are measured. The corresponding currents and their allowed ranges depend on the assembly and are shown in tabular 4.4. If the voltages or currents were not in the requested ranges, the board was investigated further (mostly there were shorts at soldering joints).

current range	3.3 V line	5.0 V line	12.0 V line
main assembly	1.8–2.2 A	0.45–0.6 A	5–10 mA
complete assembly	$-$2.8 A	$-$0.9 A	$-$15 mA

Table 4.4: The current ranges for the three voltage lines of the BOC card. The ranges depend on the assembly stage of the BOC card.

Next, the 7 Complex Programmable Logic Devices (CPLDs) of the BOC card were programmed (firmware). If this failed, the board was investigated further.

In step 3, the main electrical test (E-Test I) was done. The 5 LEDs of the BOC card were tested: two green LEDs for the power lines U=5.0 V and U=3.3 V, two yellow LEDs for the external (Patch Panel 1, PP1) and internal (crate door) interlock, one red LED for lasers on (no interlock). With test plugins the TX-sites were connected electrically with the RX-sites ("loopback"). In a later step, the plugins for transmitting/receiving light to/from the OBs are mounted here. With the test plugins the following was tested: the data paths (100% success required), the control of the plugins (100% success required), the timing (clocks), some other functionalities (for example the PIN diode current limitation). Details of the electrical test procedure can be found in [11]. If the electrical test was not 100% successfully, further investigation of the board was carried out.

In step 4, the S-Link-Card, the RX/TX-Housings and the RX/TX-Plugins, which were produced in Taiwan and are illustrated in figure 4.6, were mounted on the BOC board.

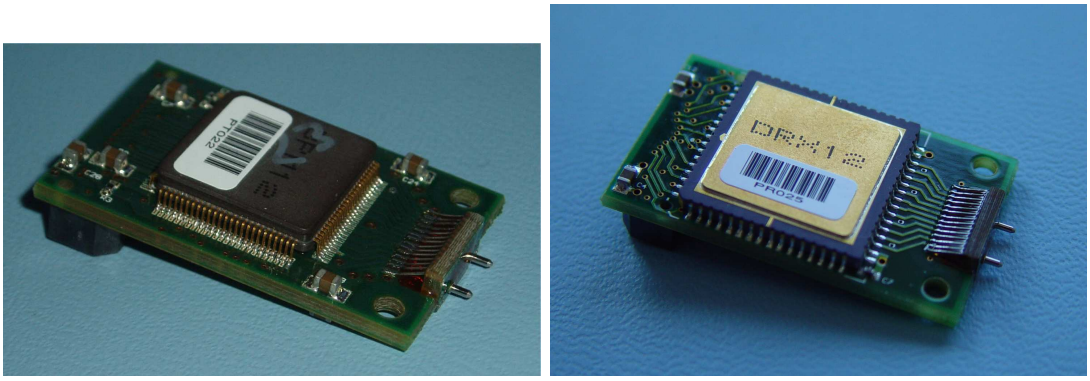

Figure 4.6: Left: a TX-Plugin. Right: an RX-Plugin. [11]

In step 5, the optical test was done. The TX-plugins were connected to the RX-plugins by optical ribbons ("loopback"). So, the mechanical alignment of the plugins/housings was verified, if light from the TX-plugins arrived for all 8 channels each at the RX-plugins. Furthermore, the TX/RX-chips could be tested. Details of the optical test procedure can be found in [11]. If the optical test was not passed, the plugins were realigned or changed. After passing the optical test, the front panel was assembled.

In step 6, the "S-Link" test (E-Test II) was done. During this, the remaining untested components of the BOC card were tested electrically by sending data from the ROD via the S-Link card of the BOC card to the ROB.

In table 4.5, the current status of the BOC card production is shown. An always up to date status can be found in [88, 89]. At the moment, there are 143 BOC cards ready for operation, but only 132 are needed for the ATLAS pixel

Table 4.5: Current status of BOC card production.

Step	Needed for ATLAS	Done
Main Assembly	132	156
Incoming Inspection	132	153
Electrical Power Test/Programming	132	153
Main Electrical Test (E-Test I)	132	153
Mechanical Assembly	132	143
Optical Test	132	143
"S-Link" Test (E-Test II)	132	143
completely finished	132	143

detector. So, there are 11 spares (8%), that can be used for testing purposes. If there would be more TX-plugins, the number of spares could be raised to 21. This issue of missing TX-plugins is discussed in more detail below.

The installation of the BOC cards in the ATLAS counting room USA15 started in April 2007. 9 crates in 5 racks were equipped with 132 BOC cards. The distribution of the BOC cards on the different crates is determined by the ATLAS pixel connectivity table [90]. One half of the BOC cards serves the A side of the pixel detector, the other half serves the C side. This is true for all BOC card types (Layer-0, Layer-1, Layer-2, Disk-1,2,3). BOC cards for the A and the C side are arranged alternatively in the crates instead of populating 1 crate with A side BOC cards and the next with C side BOC cards. Table 4.6 presents an overview of the installed BOC cards. Detailed up to date tables of the BOC card installation can be found in [91].

Table 4.6: Current status of BOC card installation in USA15.

Crate	needed BOCs/RODs	installed BOCs/RODs
Y25-05A2 upper	16	16
Y25-05A2 lower	16	16
Y26-05A2 upper	16	16
Y26-05A2 lower	12	12
Y27-05A2 upper	16	16
Y27-05A2 lower	16	16
Y28-05A2 upper	8	8
Y28-05A2 lower	16	16
Y29-05A2 upper	16	16
total	132	132

After installation of 1 TIM and 1 RCC (SBC) per crate, the installed BOC cards were powered up and showed a normal behavior of currents, voltages and temperatures. The communication between the RODs and BOC cards worked as well. Due to clock problems, the BOC card 0x53 was replaced by the BOC card 0x33, and the BOC card 0x49 was replaced by the BOC card 0x93. Due to a broken S-Link connection, the BOC card 0x72 was replaced by the BOC card 0x3A. More details about the commissioning of the BOC cards can be read in [92].

Since the installation in summer 2007 the BOC cards are operated. Occasionally, a laser channel on any TX-plugin dies (in figure 4.6 a photo of a TX-plugin is shown). This means that the emitted light power gets far below the necessary threshold or is even not measurable anymore. This problem cannot be recovered, so the TX-plugin has to be changed. During roughly two years, more than 52 TX-plugins had to be replaced. All information about failing TX-plugins such as

Table 4.7: Status of numbers of TX/RX-plugins for BOCs after their installation. The sum of "mounted and working", "defect" and "spares" gives "existing".

	needed for ATLAS	existing	mounted and working	defect	spares
TX-Plugins	272	345	284	61	0
RX-Plugins	316	366	348	4	14

detection date or serial number of the concerned TX-plugin were collected and are available in [93].
On the BOC cards of the SCT detector some TX-plugins (that have 12 channels instead of 8) failed with the same problem. Whereas, the failure rate of the RX-plugins is nearly zero for the pixel and the SCT detector. This means that there is a TX-plugin specific problem; perhaps the naked VCSEL array was damaged during production by electrostatic discharge (ESD). The problem is investigated amongst others in Taiwan, where the plugins have been produced, and in Wuppertal, where the plugins have been assembled. The status of the Taiwanese TX/RX-plugins after the installation of the BOC cards is shown in table 4.7. If the TX-plugin failure rate continues, not all channels of the pixel detector could be read out anymore, because spares are available in limited quantities only. Therefore, it was decided to produce carefully new optimized TX-plugins. All existing plugins of the pixel and SCT detector have been replaced in summer 2009, so that the optolinks are ready for the first beam collisions of the LHC.

4.3.2 Optical Cables

As mentioned in the overview of Section 4.2, optical fibers connect the ATLAS pixel on-detector electronics (OBs) in the cavern UX15 to the off-detector electronics (BOC cards) in the cavern USA15 [94]. More precisely, 80 m long glass fibers go from the BOC cards to the detector endplates at PP1. From there, further glass fibers with a length of 2.5 m go to the OBs at PP0. That is illustrated in figure 4.7. The Wuppertal group has taken a considerable responsibility for the installation and commissioning of the optical cables.
Due to the high levels of radiation inside the detector, when the LHC is in operation, the fibers must be radiation hard. This can be achieved by using Stepped Index Multi Mode (SIMM) fibers, which have a pure silica core. For the part outside the detector, Gradient Index (GRIN) fibers are used, which are doped and radiation tolerant. Corresponding to the connection layout in figure 4.7, the 2.5 m long fibers between PP0 and PP1 are SIMM fibers. The fibers between PP1 and the BOC cards consist of two spliced fibers, one 8 m long SIMM fiber and one 72 m long GRIN fiber. Since one single fiber represents an unidirectional

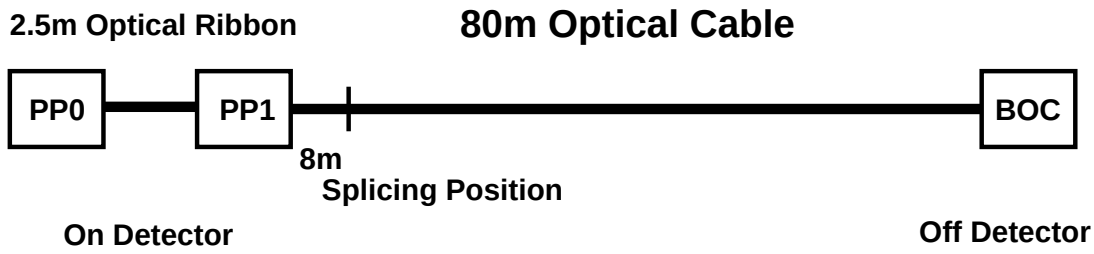

Figure 4.7: The ATLAS pixel optical connection between OBs and BOC cards.

connection of one pixel module (MCC) to a ROD, the light signals can go either from the GRIN part into the SIMM part ("timing" fiber from the BOC card to the OB) or from the SIMM part into the GRIN part ("data" fiber from the OB to the BOC card). For the "timing" case, the diameter of the GRIN part of the fiber is 62.5 μm. For the "data" case it is 50 μm. The diameter of the SIMM part of the fiber is always 50 μm. For more details look at [95].

Except for Layer-0, each of the 1744 modules requires 2 fibers, 1 to receive the clock/command signals and 1 to send the data signals. Layer-0 (for which there are 286 modules) has an extra send fiber due to the higher occupancy. Thus, 3774 fibers are needed in total. For handling reasons, the fibers are organized as ribbons. On the BOC card side, 1 ribbon consists of 8 fibers, which are glued to each other in one row (Mechanical Transfer ferrule MT8 as terminating connector). On the on-detector side, 1 ribbon consists of 16 fibers in two rows (MT16 as terminating connector). This means that 2 ribbons on the BOC card side form 1 ribbon on the PP1 side.

On the BOC card side, the male Sub Miniature C (SMC) connector, which contains the MT8 ferrule, of the ribbon is plugged into the SMC housing (TX- or RX-plugin) mounted on the BOC card. On the other side (at PP1) of the two 80 m long MT8 ribbons, one male MT16 connector is connected to the female MT16 connector of the 2.5 m long MT16 ribbon. The OB has also a female MT16 connector.

Because of these MT8/MT16 and other modularities, 6, 7 or 8 fibers of one ribbon are used. Due to further modularities it even occurs 52 times that one MT8 ribbon of two, both belonging to the same MT16 ribbon on the on-detector side, is not used. Such a ribbon is called a "dark" ribbon. Furthermore, 32 spare MT8 ribbons were produced. In total, there are 336 MT16 or 672 MT8 ribbons between the on- and off-detector side (this corresponds to 5376 fibers). Table 4.8 presents an overview of the numbers of optical MT8 ribbons.

It would be too elaborate to lay 672 thin ribbons in the cavern. Furthermore, each time there would be a risk of damage. Therefore, 8 MT8 ribbons are arranged together to form a cable. At the end of a cable, the last 3 m (1 m) of the 8 MT8 (4 MT16) ribbons have a protection to be laid safely to their destinations. Inside the cable no additional protection for each ribbon is necessary. So, there are 84

Table 4.8: Numbers of optical MT8 ribbons per destination and type (data/timing). These correspond to the numbers of needed BOC card plugins (RX/TX). Since 2 MT8 ribbons belong to 1 MT16 ribbon, the corresponding numbers of optical MT16 ribbons are half as much. Of course, a ribbon can be laid only to one detector side (A or C), i.e. the corresponding numbers of MT8 ribbons for one detector side are half as much.

Destination	Data (RX)	Timing (TX)	total
Layer-0	88	44	132
Layer-1	76	76	152
Disk-1,2,3	48	48	96
Layer-2	104	104	208
subtotal	316	272	588
dark	36	16	52
spare	16	16	32
total	368	304	672

cables in total (including spares). There are three types of cables: data, timing and hybrid, which contain 4 data and 4 timing MT8 ribbons. The numbers of cables per type and detector side are patent from table 4.9.

The cables were laid according to the ATLAS pixel connectivity table [90]. Thereafter, they were tested to check, if any of their fibers had been damaged. This was done by an Optical Time Domain Reflectometer (OTDR) which is shown in figure 4.8. Such a device can localize places in a glass fiber, where light is more strongly reflected than on average. This can happen due to damage or a connection of 2 fibers. The OTDR sends light into a fiber and measures the time t until a reflection comes back. So, the distance l to the reflecting place is given by formula 4.1, where n is the refraction index of the glass fiber and c the speed of light.

$$l = \frac{c \cdot t}{2 \cdot n} \tag{4.1}$$

The single mode (SM) OTDR, that was used, has a sensitivity of $-110\,\mathrm{dBm}$ and

Table 4.9: Numbers of optical cables per type and detector side.

	Data Cables (yellow)	Timing Cables (blue)	Hybrid Cables (gray)	total
Side A	20	16	4 (+2 spares)	40 (+2)
Side C	20	16	4 (+2 spares)	40 (+2)
total	40	32	8 (+4)	80 (+4)

Figure 4.8: Optical Time Domain Reflectometer (OTDR). [96]

operates at a wavelength of 1310 nm. The 1 point resolution is 5 mm, the 2 point resolution is 10 cm. A computer (PC) steers the OTDR via an USB interface. The OTDR sends and receives light over a Stab and Twist (ST) connector.

The ribbons were tested by the OTDR from the off-detector side, where they are more easily accessible. Since there are in one MT8 ribbon 8 fibers, which should be tested simultaneously in order to save time, a passive splitter was used. It splits the 1 channel output of the OTDR into 8 channels. To distinguish the signals of the different channels, extension fibers are used so that each channel has a different length. The passive splitter and the extension fibers have ST connectors, but the ribbon between on- and off-detector side has a SMC connector at the BOC card side. So, a fanout (1 MT8 to 8 ST) was used. The MT connector (ferrule) of the fanout was connected via 2 steel guide pins to the SMC connector of each ribbon. A sketch of the developed test setup is presented in figure 4.9.
A measurement of a ribbon is shown in figure 4.10. On the x-axis the distance in meters is plotted and on the y-axis the relative backreflection signal in dBm. The OTDR is at 0 m. Every connector gives 1 peak. The marker A indicates the start of the first channel (fiber) of the ribbon, marker B indicates its end. In between, no peaks can be recognized. So, there should be no damages of the ribbon. In figure 4.11 it is zoomed to the end of a ribbon (i.e. 80 m plus extension cables/setup from the OTDR). All 8 peaks for the 8 channels can be clearly seen. This means that light is coming back from the end of the ribbon. Due to the 2 point resolution of 10 cm there could be undetected damage inside the 10 cm from the PP1 end. This can be excluded only by connecting a further ribbon to the far end or by testing also from the PP1 side.

After they have been laid, all 84 optical cables were tested in early 2007 from the off-detector side (racks Y25-05A2 to Y29-05A2 in USA15). Only 4

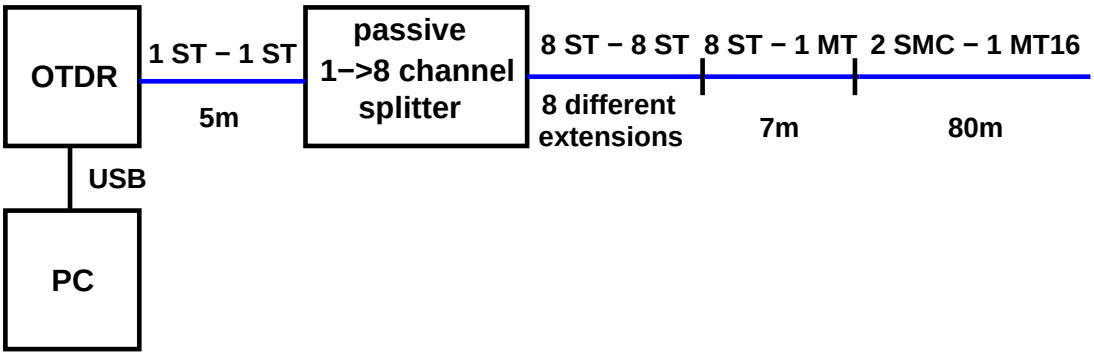

Figure 4.9: Developed test setup for the 80 m long optical ribbons of the ATLAS pixel detector between on- and off-detector side. The connector types and lengths of the ribbons/fibers are noted. 2 MT8 ribbons (SMC connector), going to the same MT16 connector, are tested one after the other.

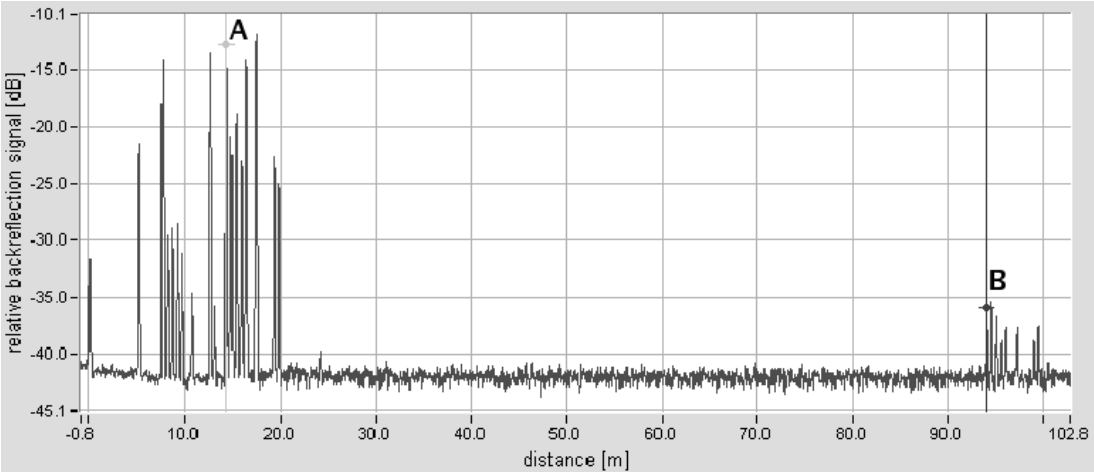

Figure 4.10: OTDR measurement of a ribbon.

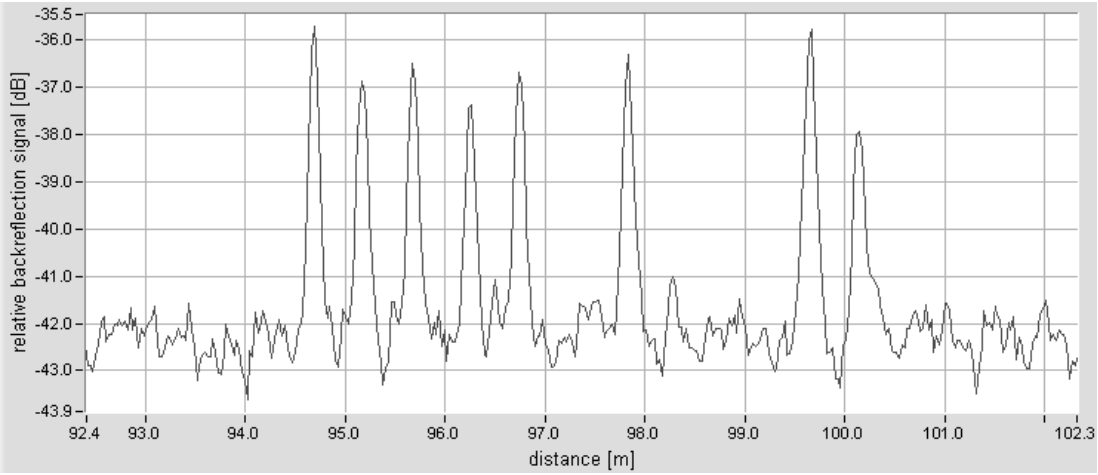

Figure 4.11: OTDR measurement at ribbon end (PP1).

MT8 ribbons indicated problems (CR1-BAT12, CR4-YCD06, CR1-YAD16, CR2-YAD16). All of them do not show the expected reflection after 80 m (end of ribbon) for all channels. Furthermore, CR1-BAT12 shows an unexpected reflection in all channels after 15 m from the BOC card side, CR4-YCD06 after 11 m, CR1-YAD16 and CR2-YAD16 after 9 m.
Three of these 4 could be tested also from the on-detector side at PP1. Since the ribbons have MT16 connectors at this end, a fanout ribbon with 2 MT8 connectors and 1 MT16 connector had to be inserted between the test setup and the ribbon to be tested. The 3 investigated ribbons showed neither the expected reflection after 80 m nor the unexpected reflections for all channels. So, these 4 MT8 ribbons had to be replaced by spares. As 2 of the faulty ribbons have the same MT16 end connector at PP1, only 3 of the 16 MT16 spare ribbons were needed (CR1-/CR2-BAT12 were replaced by CR1-/CR2-GCH04, CR1-/CR2-YAD16 by CR5-/CR6-GCH04, CR3-/CR4-YCD06 by CR5-/CR6-GCH05). Due to the described failures the ends of the optical cables were rearranged at the off-detector side. 2 spare cables were laid to rack Y26-05A2 to replace the problematic ribbons.

In summer 2007, the 3 m long ends of the ribbons at the off-detector side were laid to their destinations (according to the results of the pixel connectivity test in the SR1-building [97] and the resulting connectivity table [98]) in the racks Y25-05A2 to Y29-05A2 in USA15. Figure 4.12 shows a rack backside (BOC card

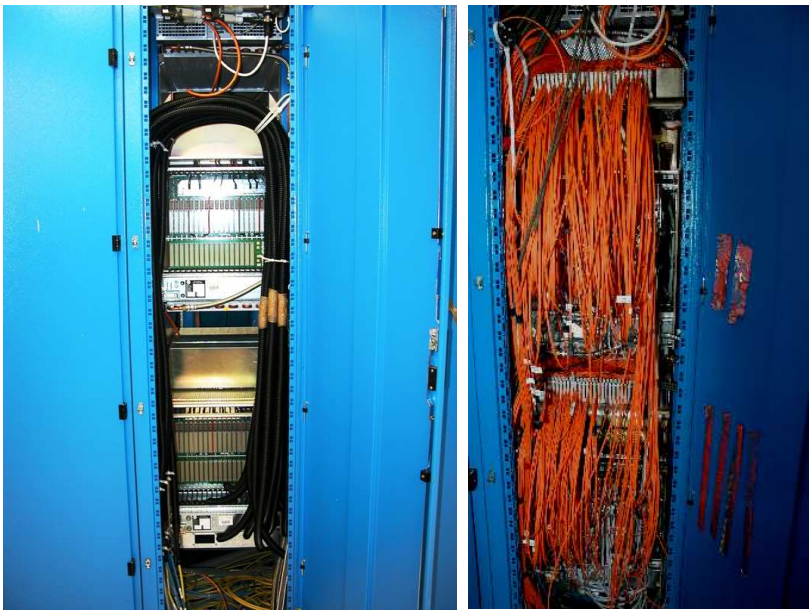

Figure 4.12: Back side of a rack with 2 crates in ATLAS counting room USA15. Left: after installation of the optical cables but before installation of the optical ribbons. Right: after installation of the optical ribbons.

side) before and after the installation of the ribbons. Every cable ends in one of the five racks, but each of the 8 ribbons of a cable can stay in the corresponding rack or go to the left or right neighboring rack. If a ribbon goes to a neighboring rack, it can go only to the upper (U28) crate, because the ribbon end is only 3 m long. If a ribbon stays in that rack, where the corresponding cable ends, it can go to the upper (U28) or lower (U07) crate.
Before the ribbons were connected to the BOC cards, all were tested again with a single mode OTDR as stated.
Three ribbons showed problems in the OTDR tests:

1. Only fiber number 8 of ribbon CR1-BAT09 showed a reflection at 80 m. All other 7 showed no reflection. So, this ribbon and ribbon CR2-BAT09 were replaced by the spare ribbons CR3-/CR4-GCH04.

2. Fiber number 8 of ribbon CR8-YCD17 showed no reflection. No action was necessary as fibers number 1 and 8 are not used in this ribbon.

3. The two fibers number 7 and 8 of ribbon CR1-BCT03 showed no reflection. So, this ribbon and ribbon CR2-BCT03 were replaced by the ribbons CR1-/CR2-GAH05.

The reason for the malfunctioning of these 3 ribbons is unknown.
A multi mode (MM) OTDR was used on some ribbons in order to confirm the single mode OTDR measurements. The results were as follows:

- "Low" reflections at 80 m in the single mode show "clear" reflections in multi mode.

- Ribbons indicate no measurable attenuation (< 1 dB) over 80 m when all 8 fibers are measured at the same time.

- When one single fiber of one ribbon is measured alone, a step in the attenuation (< 2 dB) can be observed at about 72 m (transition by splice of SIMM and GRIN fibers). When 8 fibers are measured at the same time, the corresponding 8 steps are not visible, because of the very low attenuation at the transition and of the overlap of the 8 light signals. A further reason is that the position of the splicing is not excactly the same for all fibers.

- All other measurements done in single mode are confirmed by multi mode measurements (especially missing reflections in fibers remain missing).

So, the connectivity table for the optical ribbons was updated and used for the connection of the MT16 connectors at PP1 in early 2008. More details of the OTDR tests can be found at [99], all data of the described OTDR tests at [100].

During the connectivity tests [101] in March 2008, it turned out that the connectivity table had a wrong mapping (swap) of the MT16 connectors at PP1 to

their corresponding 2 MT8 connectors at the BOC cards, so the 2 MT8 connectors were swapped for all pairs. To get the correct mapping, some database entries were swapped. However, many MT8 ribbons had to be relaid in the racks of USA15. Fortunately, it seems (some OTDR checks were done again) that no ribbon was damaged during this delicate work. More details about this recabling and the results of the connectivity tests are set forth in [102].

After the successful commissioning of the optoelectronics the pixel detector was commissioned and set into operation successfully. On 14 September 2008, the first cosmic muon was observed in the ATLAS pixel detector. Since then, the pixel detector is being calibrated — e.g. by the analysis of observed cosmic muon events — and prepared for the data acquisition of LHC beam collision events.

Chapter 5

Event Reconstruction

In order to select and analyze collision events, the recorded raw data stored in RDO files, as introduced in Section 3.2, is used to obtain physics objects such as tracks. This data processing step after the detection and recording of the collision events is called event reconstruction. The common objects needed for physics measurements are tracks, vertices, jets, electrons, photons, muons and missing transverse energy. Each of these objects and its reconstruction are summarized in the following sections, more details can be found in [26, 34].

After the reconstruction of the physics objects, they are saved together with some raw information in ESD files, which are mainly used for detector calibration. Physics analyses such as top quark precision measurements are more suitably performed with AOD or DPD files, which are derived from ESDs and small in size, because they contain almost no raw information. More information about the general reconstruction step in ATLAS can be found in [103]. In Section 7.1, the corresponding processing steps are explained for Monte Carlo simulation.

5.1 Tracking

The muon spectrometer allows a track reconstruction over the range $|\eta| < 2.7$ for muons with momenta between 3 GeV and 3 TeV. The inner detector contributes also to the measurement of muons. Moreover, vertices and tracks are reconstructed there, which are important for the jet tagging as we will see in Section 6.1. Typically, charged particle tracks with transverse momentum $p_T > 0.5\,\text{GeV}$ and $|\eta| < 2.5$ are reconstructed in the inner detector.

The inner detector track reconstruction software [104] is responsible to recognize tracks in a large number of measured space points. Tracks represent charged particles traversing the detector and depositing energy ("hits"). The software uses generally for track fitting global χ^2 and Kalman filter techniques. Furthermore, dynamic noise adjustment [105], Gaussian sum filters [106] and deterministic annealing filters [107] may be applied, especially for electron tracks. Other tools

such as those for track extrapolation are provided by the inner detector track reconstruction software as well.

Before tracks can be found or fitted, the raw data from the pixel and SCT detectors is converted into clusters, and the TRT raw timing information are turned into calibrated drift circles. The default strategy to find tracks works outwards: Space points in the three pixel layers and the first SCT layer are combined to make track seeds, which are extended throughout the SCT to form track candidates. These candidates are fitted, outlier clusters are removed, ambiguities in the cluster-to-track association are resolved, and fake tracks are rejected. Then, the selected tracks are extended into the TRT to associate drift circle information. Furthermore, left-right ambiguities are resolved. Finally, the extended tracks are refitted with the full information of the pixel, SCT and TRT detectors. The quality of the refitted tracks is compared to the silicon-only track candidates. Hits on track extensions resulting in bad fits are labelled as outliers.

Muon tracks can be reconstructed by the muon spectrometer data only (standalone fit), by the combination of the two tracks from the muon spectrometer and the inner detector (combined fit) or as a track from the inner detector to an inner muon station (segment fit). Similar to the strategy for inner detector tracks, the raw data of the muon spectrometer is first preprocessed to form drift circles or clusters. Then, patterns are found, segments are made and combined, and finally tracks are fitted.

However, the tracking may not reconstruct the tracks of all charged particles that traverse the detector. Furthermore, some tracks may be reconstruced though no corresponding particle travelled there. Such tracks, which pass some quality cuts but are not matched to any particle, are called fake tracks. The ratio of such tracks compared to all reconstructed tracks is the fake rate. This rate can be estimated by MC simulation: Generated MC particles traverse the simulated detector, and the corresponding detector response is simulated. The reconstruction software runs over this data in the same way as it is done for real data. Then, the reconstructed tracks are matched to MC particles, if at least 80% of their detector hits were created by a MC particle. The remaining tracks that have no match are fake tracks.
The standard quality cuts of the tracking require at least seven precision hits in the pixel or SCT detector. The fraction of MC particles that are matched to reconstructed tracks passing these quality cuts is called the tracking efficiency. Stricter selection cuts are used for the tracks in the b-tagging and require at least two hits in the pixels and one in the vertexing layer. For example, the efficiency is decreased and the fake rate is increased by cavern background and pile-up.

The track resolution strongly depends on the alignment of the detectors and the determination of the magnetic field. The first affects more the inner detector, whereas the second is more important for the muon spectrometer.
The magnetic field in the muon spectrometer is continuously monitored by 1730 Hall cards [38, 108], which make a 3D measurement. The absolute Hall-card

accuracy on $|\vec{B}|$ is 0.2 mT up to $|\vec{B}| = 1.4$ T and 1 mT up to 2.5 T; the angular accuracy achieved on the measured field direction is 2 mrad. The field of the solenoid was mapped [109] before the inner detector was installed but with all calorimeters in their final positions. In order to monitor the ID field strength throughout the lifetime of ATLAS, the inner detector is equipped with four NMR probes. These are fixed to the wall of the inner warm vessel near $z \approx 0$ and equally spaced in azimuth. They measure the field strength with an accuracy of about 0.01 mT.

The alignment procedure of the inner detector determines the positions in space of the silicon modules (1744 pixel and 4088 SCT modules) as well as of the straws in the 136 TRT modules every 24 hours. Therefore, a system with almost $3.6 \cdot 10^4$ degrees of freedom ($1744 + 4088 + 136 = 5968$ modules times $3 + 3 = 6$ shifts/rotations) has to be solved. With one million good tracks in this period the precision on the silicon module positions may reach 10 μm. Track based alignment techniques use the minimization of hit residuals from high momentum tracks. Since multiple scattering decreases with increasing momentum, high momentum tracks have lower distortions and are well suited for the purpose of alignment. In order to determine all global distortions, tracks with different topologies are used, for example: tracks from the interaction point, cosmic ray tracks, tracks from beam halo, and tracks passing through the overlap regions of adjacent modules.

The absolute positions of the muon chambers are also determined by track based techniques. Whereas, the relative positions of the precision chambers are continuously monitored by optical in plane alignment sensors. Thereby, an accuracy of $\leq 30\,\mu$m can be achieved for the chamber locations [26].

Since the magnetic field is not perfectly homogenous and the track particles lose energy by traversing the detector material, a track is not an exact helix. Normally, the track parametrization is chosen in a frame, where the z-axis is parallel to the magnetic field. Figure 5.1 illustrates the five track parameters used in ATLAS:

- the charge over momentum magnitude q/p
- the transverse impact parameter d_0, which is the distance of the closest approach of the track projected into the $x-y$ plane relative to the nominal interaction point $(0, 0, 0)$. During track reconstruction primary vertices cannot be chosen as reference points, because they are found and fitted afterwards.
- the longitudinal impact parameter z_0, which is the z value of the point of closest approach determined as above
- the azimuthal angle ϕ_0 of the momentum at the point of closest approach determined as above. ϕ_0 is measured in the range $[-\pi, \pi)$.
- the polar angle θ in the range $[0, \pi]$

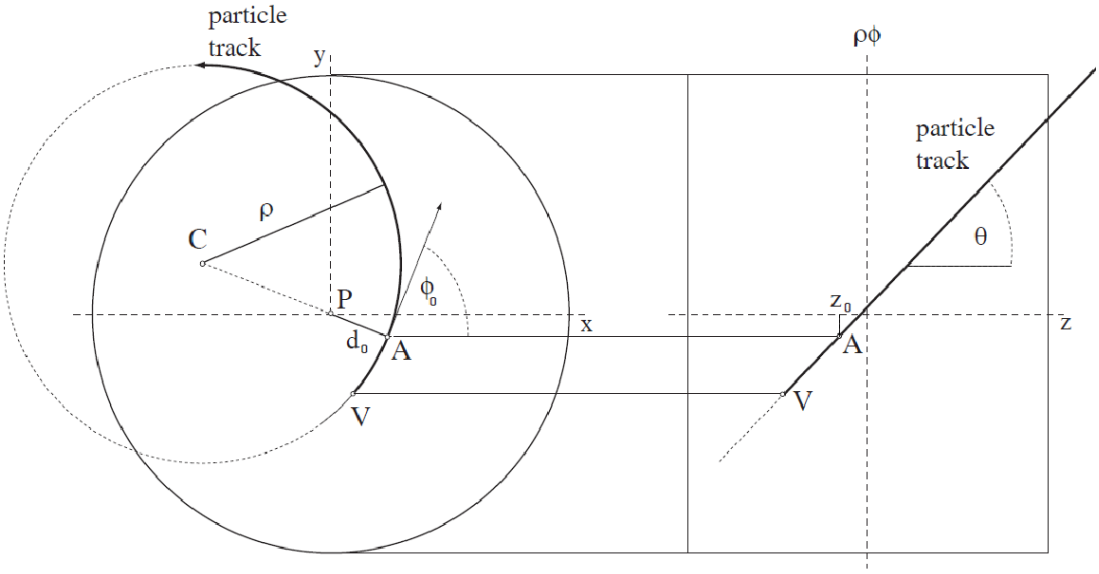

Figure 5.1: The five track parameters, split into three transverse ones ($x-y$ plane) and two longitudinal ones ($r-z$ view) for a helix track in a cylindrical detector geometry. [110]

For the consideration of their resolutions, three of the track parameters are modified as q/p_T, $\cot(\theta) = \frac{p_z}{p_T}$, $z_0 \cdot \sin(\theta)$ (z_0 projected onto the plane transverse to the track direction), because then they are more useful for many applications such as electron identification or particularly jet tagging. However, the multiple scattering of a charged particle makes the resolution of the track parameters worse than the intrinsic detector resolution. Since the multiple scattering decreases with larger momentum, the resolution of the track parameter X can be estimated as:

$$\sigma_X(p_T) = \sigma_X(\infty) \, (1 \oplus p_X/p_T) \tag{5.1}$$

The resolution $\sigma_X(\infty)$ is the asymptotic resolution expected at infinite momentum. The parameter p_X is a constant for the parameter X under consideration and represents the value of p_T, for which the intrinsic and multiple scattering terms are equal. Table 5.1 provides the expected resolutions of the five track parameters according to the mentioned detector layout and accuracies of the alignment as well as of the magnetic field. We will see in Chapter 6 that an accurate track reconstruction and high resolutions of their impact parameters are very important for the selection, identification and analysis of top quark events.

track parameter	$0.25 < \|\eta\| < 0.50$		$1.50 < \|\eta\| < 1.75$	
	$\sigma_X(\infty)$	p_X [GeV]	$\sigma_X(\infty)$	p_X [GeV]
inverse transverse momentum q/p_T	$0.34\,\text{TeV}^{-1}$	44	$0.41\,\text{TeV}^{-1}$	80
azimuthal angle ϕ	$70\,\mu$rad	39	$92\,\mu$rad	49
polar angle $\cot(\theta)$	$0.7 \cdot 10^{-3}$	5.0	$1.2 \cdot 10^{-3}$	10
transverse impact parameter d_0	$10\,\mu$m	14	$12\,\mu$m	20
longitudinal impact parameter $z_0 \cdot \sin(\theta)$	$91\,\mu$m	2.3	$71\,\mu$m	3.7

Table 5.1: Expected track parameter resolutions (RMS) at infinite transverse momentum, $\sigma_X(\infty)$, and transverse momentum, p_X, at which the multiple scattering contribution equals that from the detector resolution. The values are shown for two η regions, one in the barrel of the inner detector, where the amount of material is close to its minimum, and one in the end cap, where it is close to its maximum. [34]

5.2 Vertexing

When the reconstruction of tracks is finished, vertices can be found and fitted, which represent the positions, where particles decay. Especially the reconstruction of primary vertices, where interactions between beam particles occured, and conversions of photons into positron-electron pairs in the detector are important for many physics analyses.

There are several algorithms to find and fit primary vertices [34,111,112]. All primary vertex finders use tracks, which originate from the beam crossing area. The expected resolution of the primary vertex in top-antitop events is $18\,\mu$m in the $x-y$ plane and $41\,\mu$m in the z-direction. Due to a beam constraint, the resolution in the transverse plane can be improved to $11\,\mu$m [34]. The situation when collision events have more than one primary vertex is called pile-up, as already introduced in Section 2.3. Especially in such a case a good tracking performance is important in order to associate the tracks to the primary vertex they originate from. A correct association reduces backgrounds of physics analyses and is vital for the b-tagging of jets, as we will see in Chapter 6.

Secondary vertices represent particles decaying at a point displaced from any primary vertex. Especially when a particle containing a b-quark decays, the reconstruction of the corresponding vertex may be important to identify a jet as a b-jet. The resolution of the radial position of secondary vertices strongly depends on where the particle decays: if the decay is in front or behind of a detector or detector layer may cause a factor of up to 10.

5.3 Electrons and Photons

Electrons and positrons are both referred as electrons due to their undistinguishable behavior in the calorimeter. The inner detector may determine the sign of the electric charge of an electron measured in the calorimeter, if an electron track can be associated to it.

For the reconstruction of $t\bar{t}$ events with one top quark decaying leptonically in the electron channel, it is crucial to find the electron in the large amount of jets: at the LHC the electron-to-jet ratio is expected to be about 10^{-5} at $p_T = 40\,\text{GeV}$ [34]. Moreover, the accurate reconstruction of electrons and their discrimination from photons are important for precision measurements.
The reconstruction and identification of electrons/photons starts with a seed cluster from the electromagnetic calorimeter. Electron and photon candidates are separated reasonably cleanly by requiring the electrons to have an associated track but no associated conversion into an electron/positron pair. Tracks and conversions are both reconstructed in the inner detector. Similarly, photons have no matched track but may have a matched reconstructed conversion. Shower shape variables, for example lateral and longitudinal shower profiles, help to separate electrons from jets. Finally, calorimeter energy isolation or observables from the inner detector such as TRT high threshold hits are used to identify electrons and photons.

These observables from the inner detector and from the calorimeters also help to reject jets. The corresponding rate is between 500/1300 for low energetic electrons with loose/medium cuts (shower-shape or track cuts) and 10^5 for high energetic electrons with tight cuts (high ratio between high-threshold and low-threshold hits in the TRT detector or calorimeter energy isolation). However, the efficiency strongly depends on the physics process, for example the efficiency of "tight isolated" high-energetic electrons may be 64% for $Z \to ee$ but only 16% for electrons, which are near heavy flavor jets [34].

For photons the jet rejection rate is not as high, because jets can fake photons, especially when they contain single or multiple neutral hadrons. Nevertheless, a quark jet rejection rate of about 1500 without and 3000 with track isolation is achievable for high energetic photons. For gluon jets these numbers are by a factor of 10 larger, because they have a broader lateral extent due to the softer fragmentation [34].

Thus, the inner detector contributes a lot to the reconstruction and identification of electrons/photons.

5.4 Missing Transverse Energy

In order to measure the top quark mass from $t\bar{t}$ events, in which one top quark decays leptonically, a good measurement of the missing transverse energy $\not{E}_T$ in

terms of linearity and accuracy is important. $\not\!\!E_T$ is determined by the muon spectrometer and the calorimeters using the same global calibration of calorimeter cells as for jets. The uncalibrated $\not\!\!E_T$ is the energy that sums up with the transverse energy of all calorimeter cells and the p_T of the muons, measured by the muon chambers, to zero transverse energy. To avoid fake muons, only muons with a track match in the inner detector are selected. Since a lot of energy is lost in the cryostat between the barrel LAr electromagnetic and tile calorimeter, a correction has to be applied, which represents about 5% per jet with $p_T > 500\,\text{GeV}$ [34]. A refined calibration of $\not\!\!E_T$ is obtained, if each reconstructed high p_T object is associated to its globally calibrated cells.

Over a wide range of the total transverse energy deposited in the calorimeters, the resolution σ of the missing transverse energy follows an approximate stochastic behavior: A simple fit to the function $\sigma = a \cdot \sqrt{\sum E_T}$ yields values of about 0.55 ± 0.02 for the parameter a, when $\sum E_T$ is in the range between 20 and 2000 GeV. Above 2 TeV, the constant term in the jet energy resolution dominates, as we will see in Section 5.5. Monte Carlo simulations provide an estimation of the azimuthal accuracy, which is defined as the angular difference between the directions of reconstructed and true missing transverse energy. For a missing transverse energy above 100 GeV a resolution of about 100 mrad is expected.
However, several effects may change the direction and amount of the missing transverse energy: beam gas scattering and other machine backgrounds, displaced interaction vertices, hot/dead/noisy cells (or regions) in the calorimeters, and mismeasurements in the detector itself, for example due to high p_T muons escaping outside the acceptance of the detector and due to large losses of deposited energy in cracks or inactive material. Normally, the transverse missing energy is increased by such effects. Therefore, the fake $\not\!\!E_T$ is defined as the difference between reconstructed and true missing transverse energy. The ratio between fake and true $\not\!\!E_T$ can be reduced by applying cuts. For QCD dijet events a reduction of a factor 10 over the whole energy range can be gained, if the distance in azimuth between the reconstructed $\not\!\!E_T$ vector and the direction of any high p_T jet is required to be larger than 17°. [34]

5.5 Jets

At the LHC, jets will be present in almost every event, whether they originate from QCD background processes or represent signal, for example in top quark pair events. As introduced in Section 3.2.2, jets are detected by the calorimeters. However, there is a long chain of processing steps from calorimeter signals to jets, which are finally calibrated to parton level, as illustrated in figure 5.2.

Jets are formed from signals of all calorimeters. For a fast jet finding, cell signals are combined into clusters. There are two possibilities for such a combi-

nation: tower clusters and topological clusters. Towers are formed by collecting cells into bins of a regular $\Delta\eta \cdot \Delta\phi = 0.1 \cdot 0.1$ grid and summing up their signals. If a cell overlaps partly with the tower bin, a fraction of its signal is added, which corresponds to the overlap area fraction.
Topological cell clusters represent an attempt to reconstruct three-dimensional energy depositions in the calorimeters. Starting from seed cells, which have a high absolute signal, close cells are collected, if they have absolute signals above certain thresholds.

There are a lot of jet-clustering algorithms, which reconstruct jets from the towers or topological clusters. Two of them are very common in ATLAS: a seeded fixed-cone algorithm and a successive recombination algorithm (k_T). Both can produce narrow ($\Delta R = 0.4$) or wide ($\Delta R_{\text{cone}} = 0.7$, $\Delta R_{k_T} = 0.6$) jets. The cone algorithm is the default for the high-level trigger, because it is very fast. As parameters it uses the transverse energy threshold for a seed and the cone size ΔR. A split-and-merge step follows the actual cone building, with an overlap fraction threshold of 50%. The successive recombination algorithm is preferred for offline physics analyses. However, the choice of the jet algorithm depends also on the physics channel under consideration.

After the jets are found, they have to be calibrated. First, the cell signals are weighted to obtain the hadronic calibration. The standard approach is similar to the one developed for the H1 calorimeter [113]. The weighting functions depend on the calorimeter signal types (towers or topological clusters), on the jet algorithm (for example cone or k_T) and its specific configuration, and on the physics calibration samples. Therefore, jet energy scale corrections have to be applied to get physics jets, which are calibrated to particle level.
Finally, physics jets are refined to interaction/parton level by in-situ measurements, which include for example pile-up. Good choices for this task are final states such as γ+jet(s), Z+jet(s) or hadronically decaying W bosons ($W \to qq$), because they have a well measured electromagnetic object, which balances one or more jets in transverse momentum or because the mass of the W boson constrains the energy scale of the two quark jets.

In order to analyze the performance of the jet reconstruction, efficiency and fake rate are determined in a similar way as described in Section 5.1. Generated MC particles are clustered, usually by cone algorithms, into "truth" jets. Therefore, the jet reconstruction efficiency is defined as the ratio between the number of matches of "truth" jets with reconstructed jets and the total number of "truth" jets. Whereas, the purity of the jet reconstruction can be expressed as the ratio between the number of matches of "truth" jets with reconstructed jets and the total number of reconstructed jets. The fake rate is then one minus purity. The corresponding values depend on the chosen matching radius, which is typically $\Delta R_m = 0.2$. Moreover, the two calorimeter signal definitions (towers or topological clusters) as well as the various jet algorithms may lead to different efficiencies or purities. Finally, pile-up and electronic noise may decrease the effi-

ciency. However, efficiency and purity are almost 100% for jets with a transverse energy above 50 GeV.

In order to analyze efficiencies and purities of jet reconstruction and jet tagging in more detail, a jet label is defined, which uses MC information. This label indicates from which type of parton the jet originates. However, this labelling procedure is ambiguous. Usually, a jet is labelled as a b-jet, if a b-quark with transverse momentum $p_T > 5\,\text{GeV}$ is found in a cone of size $\Delta R = 0.3$ around the jet direction. If there is no b-quark found but a c-quark or a τ-lepton, which fulfills the same conditions, the jet is labelled accordingly. Otherwise, the jet is labelled as a light jet. Especially in Chapter 6, these labels are needed for performance studies.

After global calibration the fractional jet energy resolution, whose performance is quite important for the reconstruction of $t\bar{t}$ events, may be fitted by following function:

$$\frac{\sigma}{E} = \sqrt{\frac{a^2}{E} + \frac{b^2}{E^2} + c^2} \tag{5.2}$$

For cone jets made from towers with $\Delta R = 0.7$ and $0.2 < |\eta| < 0.4$, the stochastic term $[a/\sqrt{E}]$ is $\approx 60\%\sqrt{\text{GeV}}$, while the high-energy limit of the resolution, expressed by the constant term c, is $\approx 3\%$. The parameter b represents the noise, which increases from 0.5 GeV to 1.5 GeV when going from the barrel to the end cap.

Another important observable of the jet reconstruction is the jet axis. A MC study shows that the choice of the calorimeter signal (towers or topological clusters) does not significantly affect the reconstructed jet direction. Though, the reconstruction of the jet axis is degraded at transverse energies below 100 GeV. However, the precise measurement of the jet direction is very important for b-tagging, as we will see in the next chapter.

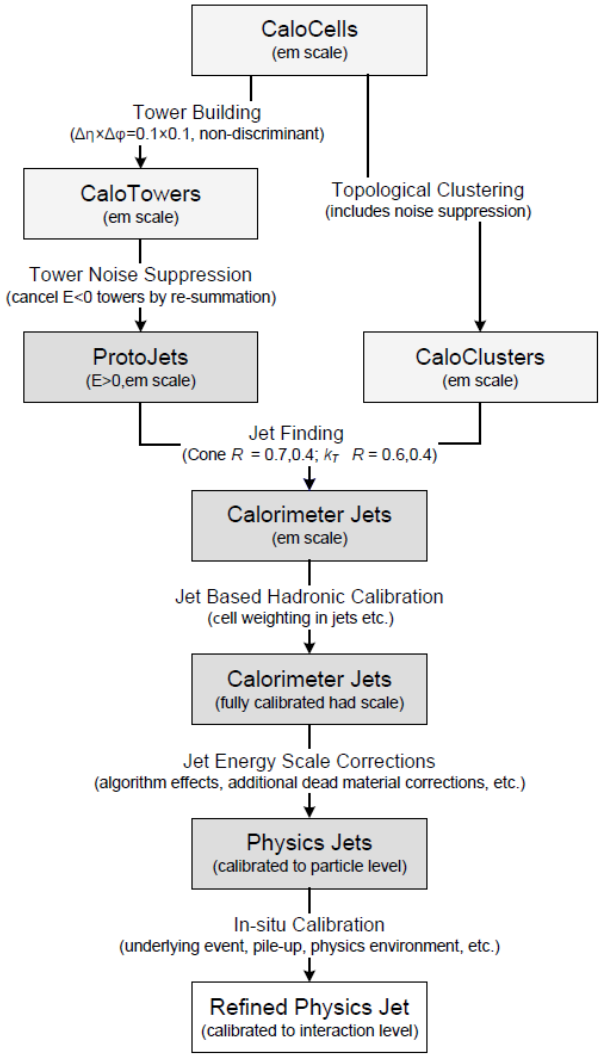

Figure 5.2: Jet reconstruction flow for calorimeter jets from towers or clusters. [34]

Chapter 6

Tuning of Track Impact Parameters for b-Tagging

Track impact parameters are important for b-tagging as we will see in Section 6.1. In order to understand the detector and to have reliable measurements of impact parameters, it is necessary to perform simulations. If these are in agreement with data, maximal systematic uncertainties can be estimated. However, track impact parameter distributions will quite likely not agree between first data and Monte Carlo simulation because of a misalignment of the detector, detector inefficiencies, etc. Therefore, the MC distributions are tuned to account for such effects. After the tuning, the deviations between simulation and data distributions provide an estimate of the systematic uncertainties of track impact parameters. This chapter overviews the b-tagging, presents the development and tests of an impact parameter tuning approach, and points at eventual alternatives.

6.1 b-Tagging

For many physics analyses like Higgs boson searches or top quark measurements it is important to identify jets stemming from the fragmentation and hadronization of b-quarks. Hence, there are jet tagging algorithms, which calculate for each jet a probability that the jet originates from a certain quark flavor.

b-tagging algorithms use the fact that b-jets contain a B-hadron, that has a relatively long lifetime of the order of 1.5 ps ($c \cdot \tau \approx 450\,\mu$m). A B-hadron in a jet with $p_T = 50\,\text{GeV}$ will therefore have a significant flight path length $\langle l \rangle = \beta \cdot \gamma \cdot c \cdot \tau$ [114], typically 3 mm in the plane transverse to the beam before decaying. The tracks from the B-hadron decay products can be used to reconstruct the decay vertex, a so called secondary vertex (SV). If such a vertex is near or inside a jet, this jet may be tagged as a b-jet. A jet can be also identified by using the measured impact parameters (IP) of its tracks. Whereas in the tracking, the track parameters are usually defined relative to the nominal interaction point

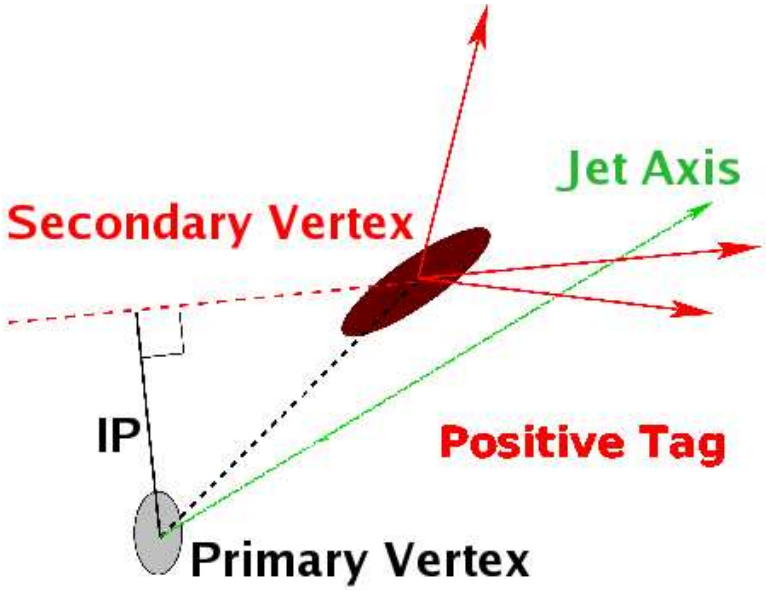

Figure 6.1: Example of a positive jet tag: The jet may be tagged due to the reconstructed secondary vertex, which lies near the jet axis, or the large impact parameters of its tracks. [115]

$(0, 0, 0)$, the jet tagging uses definitions relative to the associated reconstructed primary vertex: The transverse impact parameter, d_0, is the distance of closest approach of the track to the primary vertex (PV) point in the $r - \phi$ projection. The longitudinal impact parameter, z_0, is the z coordinate of the track at the point of closest approach in $r-\phi$. [114] As illustrated in figure 6.1, tracks from a B-hadron decay tend to have large impact parameters which can be distinguished from small impact parameters of tracks originating from the primary vertex. Therfore a jet, to which large impact parameter tracks are associated, can be identified as a b-jet. As a consequence, b-tagging relies highly on tracking and vertexing.

For further discrimination of tracks from B-hadron decays from tracks originating from the primary vertex the impact parameter is signed. As illustrated in figure 6.2, the sign is positive, if relative to the primary vertex the point of closest approach lies on the side of the jet; otherwise the sign is negative. More precisely, the sign is defined using the jet axis $\vec{P}_j$ as measured by the calorimeters, the direction $\vec{P}_t$ and the position $\vec{X}_t$ of the track at the point of closest approach to the primary vertex and the position $\vec{X}_{pv}$ of the primary vertex [114]:

$$\text{sign}(d_0) = \text{sign}\left[(\vec{P}_j \times \vec{P}_t) \cdot \left(\vec{P}_t \times (\vec{X}_{pv} - \vec{X}_t)\right)\right] \tag{6.1}$$

Similarly, the sign of the longitudinal impact parameter z_0 is given by the sign of $(\eta_j - \eta_t) \cdot z_{0t}$, where η_j is the pseudorapidity of the jet, η_t is the pseudorapidity of the track and z_{0t} is the impact parameter relative to the primary vertex.

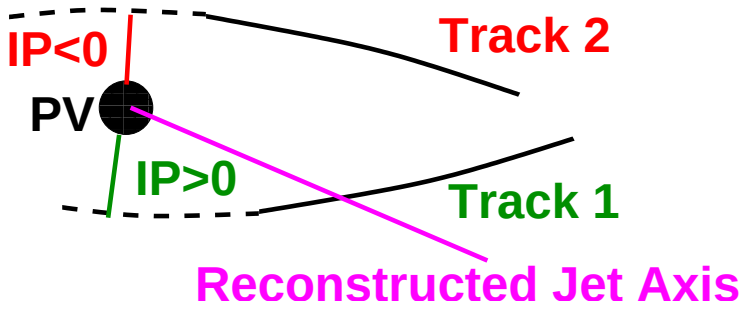

Figure 6.2: Visualization of the definition of the signed track impact parameter.

Due to the limited experimental resolution the sign of the impact parameters of tracks originating from the primary vertex, whatever jet they are associated to, is random in an ideal experiment. However, particles can be deflected by material interactions. If this deflection is not corrected for, the reconstructed track may have a positive, larger impact parameter than the particle would have without any material interaction. Similarly, impact parameters may be shifted toward higher values, if some decay vertices are not resolved. Hence, the impact parameter distribution of tracks associated to the primary vertex has larger tails on the positive side.

The signed impact parameter distribution of tracks is even more asymmetric because of long lived particles (e.g., K^0, Λ^0, D^0, B^0, etc.), that decay displaced from the primary vertex and lead to much more entries on the positive side of the distribution. Although B-hadrons have a shorter lifetime than lighter hadrons, the positive side of the signed impact parameter distribution is more raised for b-jets than for light jets, as illustrated in figure 6.3. This is due to the proportion of tracks originating from long lived particles. Though the negative side of the impact parameter distribution is mainly determined by the intrinsic detector resolution, negative tails arise, if the reconstructed jet axis deviates from the primary hadron flight axis. This is illustrated in figure 6.4 and happens especially in b-jets, when the distance between B- and C-hadron decay vertices is significant compared to the vertex resolution in flight direction. The larger the deviation is, the more sizeable are the negative tails. If a track from the B-hadron decay vertex lies between both axes (straight track assumed), it has a negative impact parameter.

Due to multiple scattering, the track impact parameter distribution is a function of the transverse momentum p_T. There is also a dependence on the pseudorapidity η of the track, because of the geometry of the detector (different resolutions and distances between measurement points) and the distribution of the material. These dependencies are illustrated in figures 6.5 and 6.6. Similarly, the error of the impact parameter can vary significantly and depends on p_T, η, cluster topology and hit pattern. In order to give more weight to precisely measured

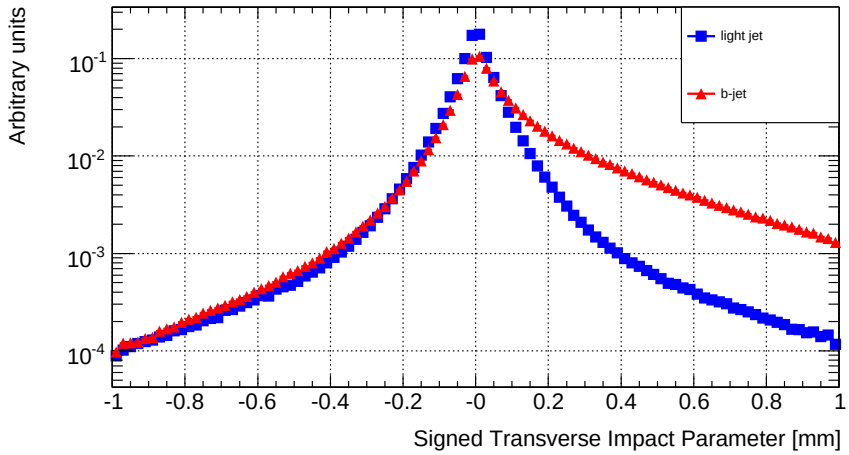

Figure 6.3: Signed transverse track impact parameter distribution for b- (red) and light (blue) jets in $t\bar{t}$ events.

tracks, the impact parameter IP is divided by its error σ_{IP}. Consequently, a better discrimination is reached, whether a track stems from the primary vertex or not. This discriminating variable d_0/σ_{d_0} or z_0/σ_{z_0} is called the signed impact parameter significance.

Track impact parameter significances are used by several b-tagging algorithms, for example IP1D (z_0/σ_{z_0}), IP2D (d_0/σ_{d_0}) and IP3D (combination of IP1D and IP2D) [114]. The first two algorithms use a likelihood ratio method: The significance S_i of a track i, e.g. d_0/σ_{d_0}, is compared to predefined smoothed and normalized distributions (so called reference histograms) for both the b- and light jet hypotheses, $b(S_i)$ and $u(S_i)$. The ratio of the probabilities $b(S_i)/u(S_i)$ defines the track weight, which can be combined into a jet weight W_{jet} as the sum of

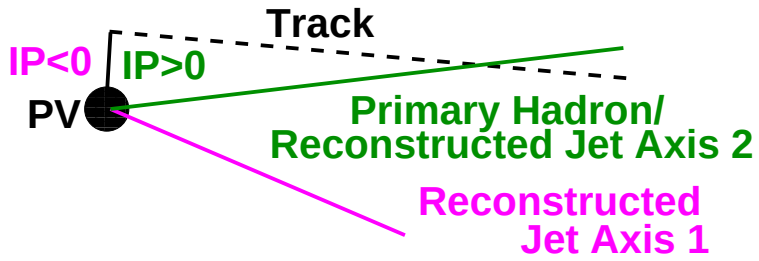

Figure 6.4: The sign of the track impact parameter depends on the reconstruction of the jet axis. Jet axis 2 (green) may be the flight direction of the B-hadron, and the sign would be positive, but jet axis 1 (purple) is reconstructed and this leads to a negative sign.

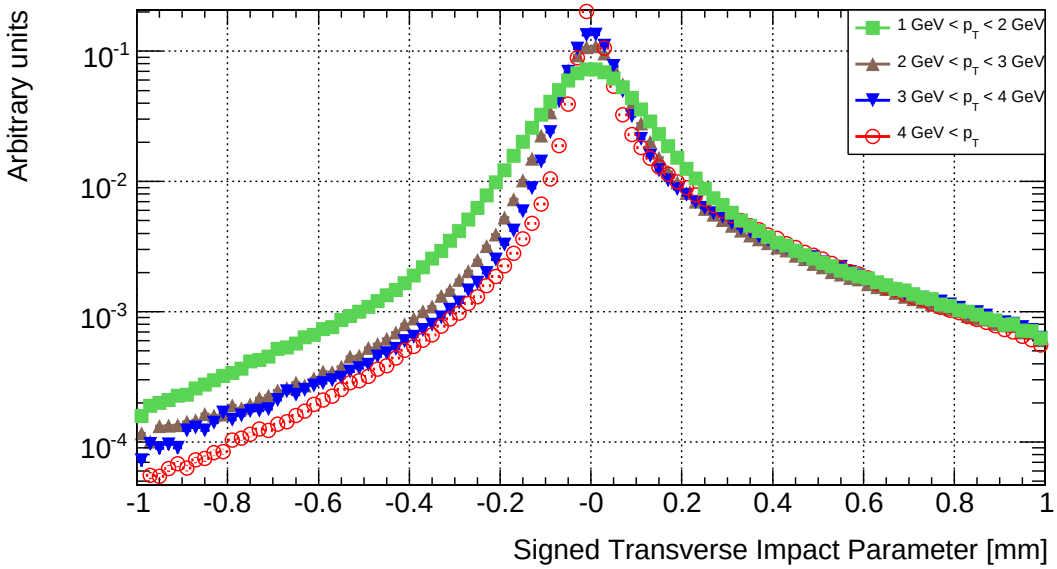

Figure 6.5: Dependence of the track impact parameter on the transverse momentum of the track. The tracks are associated to any jet of $t\bar{t}$ events.

the logarithms of the N_t individual track weights W_i:

$$W_{Jet} = \sum_{i=1}^{N_t} \ln(W_i) = \sum_{i=1}^{N_t} \ln\left(\frac{b(S_i)}{u(S_i)}\right) \tag{6.2}$$

To select b-jets, a cut value on W_{Jet} must be chosen and corresponds to a given b-tagging efficiency, which is defined as follows: Out of all jets labelled as b-jets (according to the jet labelling explained in Section 5.5) how many of them are tagged as b-jets? The performance of a b-tagging algorithm is then given by its efficiency and its rejection rate, that is the inverse of the mistagging rate. Similar to the efficiency, the mistagging rate answers: out of all jets labelled not as b-jets, how many of them are tagged as b-jets? However, the efficiencies and the corresponding rejection rates of many b-tagging algorithms depend for example on the jet transverse momentum and pseudorapidity.

Some algorithms such as IP2D require an a priori knowledge of the properties of both b- and light jets, namely the reference histograms of the track impact parameter significances. The b-tagging algorithms were implemented and calibrated by using MC simulation data samples. But these samples may have different distributions of track parameters (d_0, z_0, ϕ, θ, q/p and their errors or the corresponding error matrix V) than real data samples. Some possible reasons are: misalignment of the detector, detector inefficiencies and detector material description; different heavy flavor content for different generators of the simulation. A different impact parameter distribution would lead to a change of the

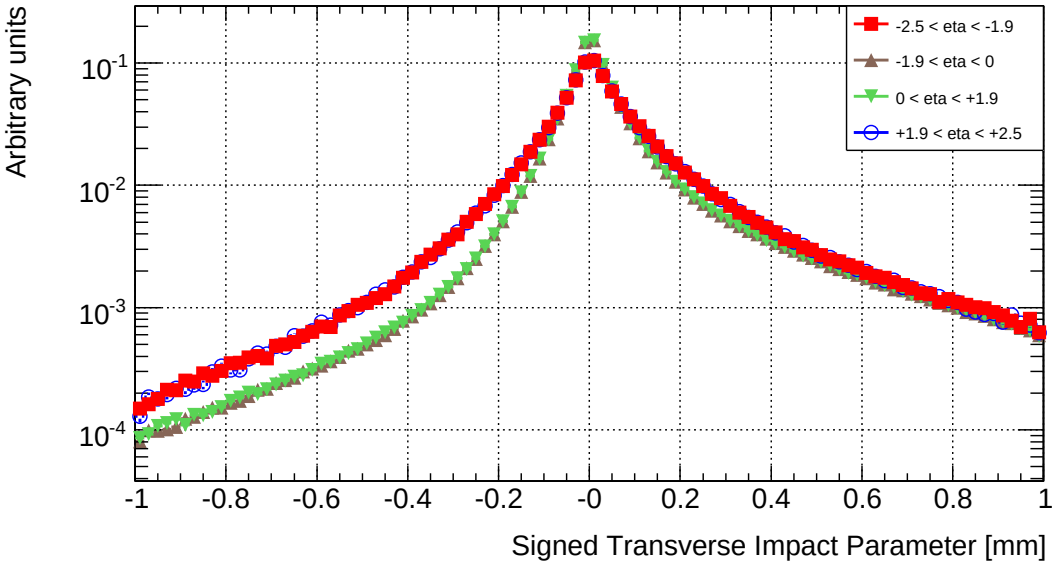

Figure 6.6: Dependence of the track impact parameter on the pseudorapidity of the track. The tracks are associated to any jet of $t\bar{t}$ events.

tagging efficiency and therefore of the signal amplitude of the physics process in concern. To get closer agreement with data, the MC generator or the detector simulation may be tuned. Such a procedure can be quite complicated and elaborate. Instead, the track impact parameters of the MC samples may be adjusted, a procedure which can be done more rapidly. A possible approach of such a method is presented in the next section.

6.2 Adjustment Strategy

To get closer agreement with data, the MC track impact parameter distribution may be forced to the shape of the real data (RD) distribution, as illustrated in figure 6.7. The reconstructed MC impact parameter d_0^{MC} (likewise z_0^{MC}) is the sum of the true impact parameter $d_0^{\mathrm{MC,true}}$ and its correction Δd_0^{MC}:

$$d_0^{\mathrm{MC}} = d_0^{\mathrm{MC,true}} + \Delta d_0^{\mathrm{MC}} \tag{6.3}$$

The correction Δd_0^{MC} depends mainly on the detector resolution, and the true impact parameter is zero ($d_0^{\mathrm{MC,true}} = 0\,\mathrm{mm}$) for tracks (simulated particles) from the primary vertex. When the distributions of such tracks do not coincide for MC and RD, the MC correction (Δd_0^{MC}) has to be adjusted for all MC tracks. After the MC correction is set to the corresponding RD resolution ($\Delta d_0^{\mathrm{MC}} := \Delta d_0^{\mathrm{RD}}$), the distribution of the adjusted MC impact parameter ($d_0^{\mathrm{MC,AD}}$) should be mapped

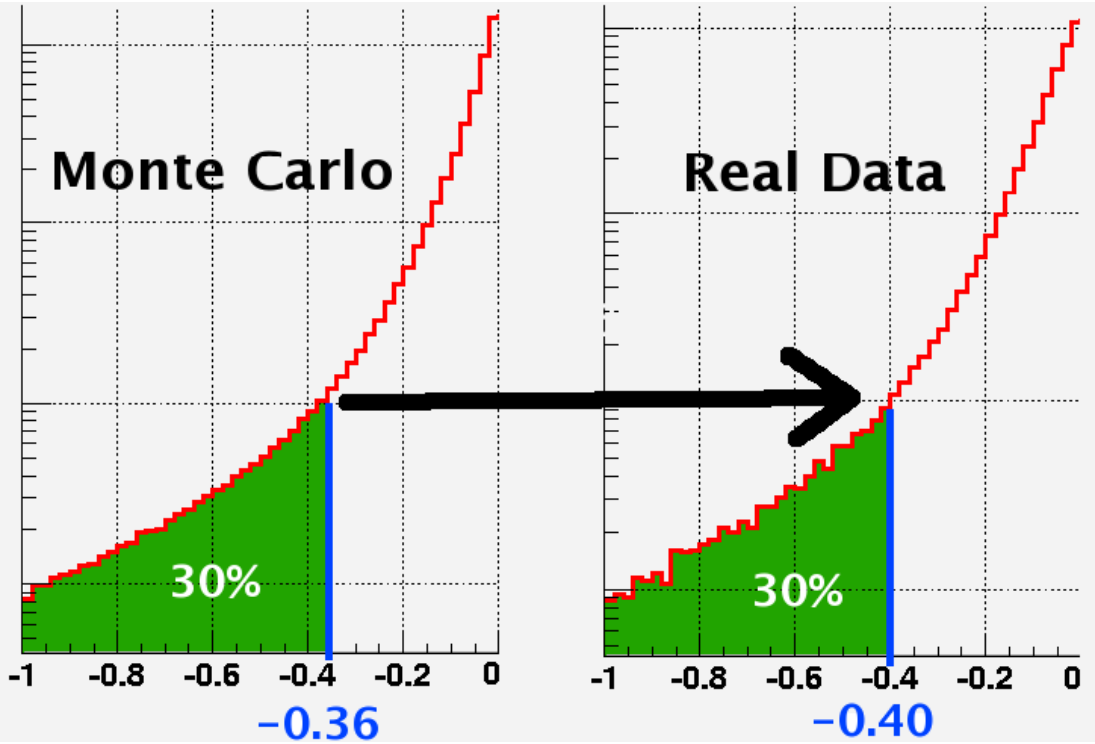

Figure 6.7: Sketch of the mapping approach. A MC track has an impact parameter of −0.36 mm. This corresponds to a position in the distribution, where the integral is 30% (whole distribution normalized to 1). The position in the real data distribution (also normalized to 1), where the integral is also 30%, is −0.40 mm. So, the MC track impact parameter is forced to the value −0.40 mm.

on the RD distribution (hence the name mapping approach) and they should agree well.

Using the RD impact parameter distribution, the resolution Δd_0^{RD} is obtained as follows: The negative side of the RD impact parameter distribution (d_0^{RD}) provides a good estimation of the resolution (Δd_0^{RD}), because it is dominated by the intrinsic detector resolution, as indicated in the previous section. When a symmetric resolution around zero is assumed, the full resolution distribution (Δd_0^{RD}) is obtained by just mirroring the negative side to the positive. For a certain MC track the corresponding RD resolution (Δd_0^{RD}) is defined, when the integral over the normalized MC distribution (obtained by the same mirror technique as for RD) up to the MC correction value (Δd_0^{MC}) is equal to the integral over the RD distribution (for both distributions the lower/upper integration limit is ±1 mm, which is the default impact parameter cut for $b-$tagging in ATLAS [114]):

$$\int_{-1}^{\Delta d_0^{\mathrm{MC}}} \frac{1}{N}\frac{d\,n}{d\,d_0^{\mathrm{MC}}} d\,d_0^{\mathrm{MC}} = \int_{-1}^{\Delta d_0^{\mathrm{RD}}} \frac{1}{N}\frac{d\,n}{d\,d_0^{\mathrm{RD}}} d\,d_0^{\mathrm{RD}} \tag{6.4}$$

Then, the adjusted MC impact parameter for tracks from the primary vertex is:

$$d_0^{\mathrm{MC,AD}} = d_0^{\mathrm{MC,true}} + \Delta d_0^{\mathrm{MC,AD}} = 0 + \Delta d_0^{\mathrm{RD}} \tag{6.5}$$

Since the track impact parameter distribution depends on observables like the transverse momentum p_T (due to multiple scattering) or the pseudorapidity η (due to detector geometry) of the track (as shown in figures 6.6 and 6.5), the adjustment has to be made in bins of these observables.

For tracks, which do not originate from the primary vertex, the adjustment strategy is the same. Since it is assumed, that the resolution does not strongly depend on the track type (from primary vertex or not), Δd_0^{RD} can be retrieved in the same way as above. The true impact parameter $d_0^{\mathrm{MC,true}}$ is not zero, but the equations 6.3, 6.4, 6.5 still hold.
However, the following two approximations for both types of tracks are reasonable as the results below indicate: The true and the reconstructed primary vertex are exactly coinciding. The true and the reconstructed impact parameter, which is actually a vectorial quantity, are parallel.

Most b-tagging algorithms do not use the track impact parameter itself but its significance. When the reconstructed MC impact parameter is adjusted as explained above, the significance will change accordingly. For b-tagging algorithms relying only on track impact parameters the efficiencies and rejection rates may agree between MC and data, after the adjustment was done.
However, it is still possible, that the impact parameter significance distribution of MC does not coincide with that of data. A reason may be an unaccurate estimation of the individual track parameter errors. Then, it may be applicative to scale the error of the impact parameter as well. A first approach is presented in [116].

The implementation of the track impact parameter mapping approach within the Athena framework of ATLAS is described in Appendix A.1.

6.3 Performance and Tests

The performance tests of the presented approach use the default tracking [104], jet definition and reconstruction [114] of ATLAS. During these studies LHC and ATLAS had not collected any collision data. But the procedure of tuning track impact parameters needs to be tested before collision data is available. Instead of using MC and real data, one can use two different MC data samples. A sample with a perfectly aligned detector geometry may act as MC data, whereas a MC data sample with a misaligned detector geometry can act as real data 'RD' (pseudo data). In figure 6.8 signed track impact parameter distributions of such samples and their difference are shown. As expected, the distribution of the MC dataset, which has a misaligned detector geometry, is broader (corresponding to

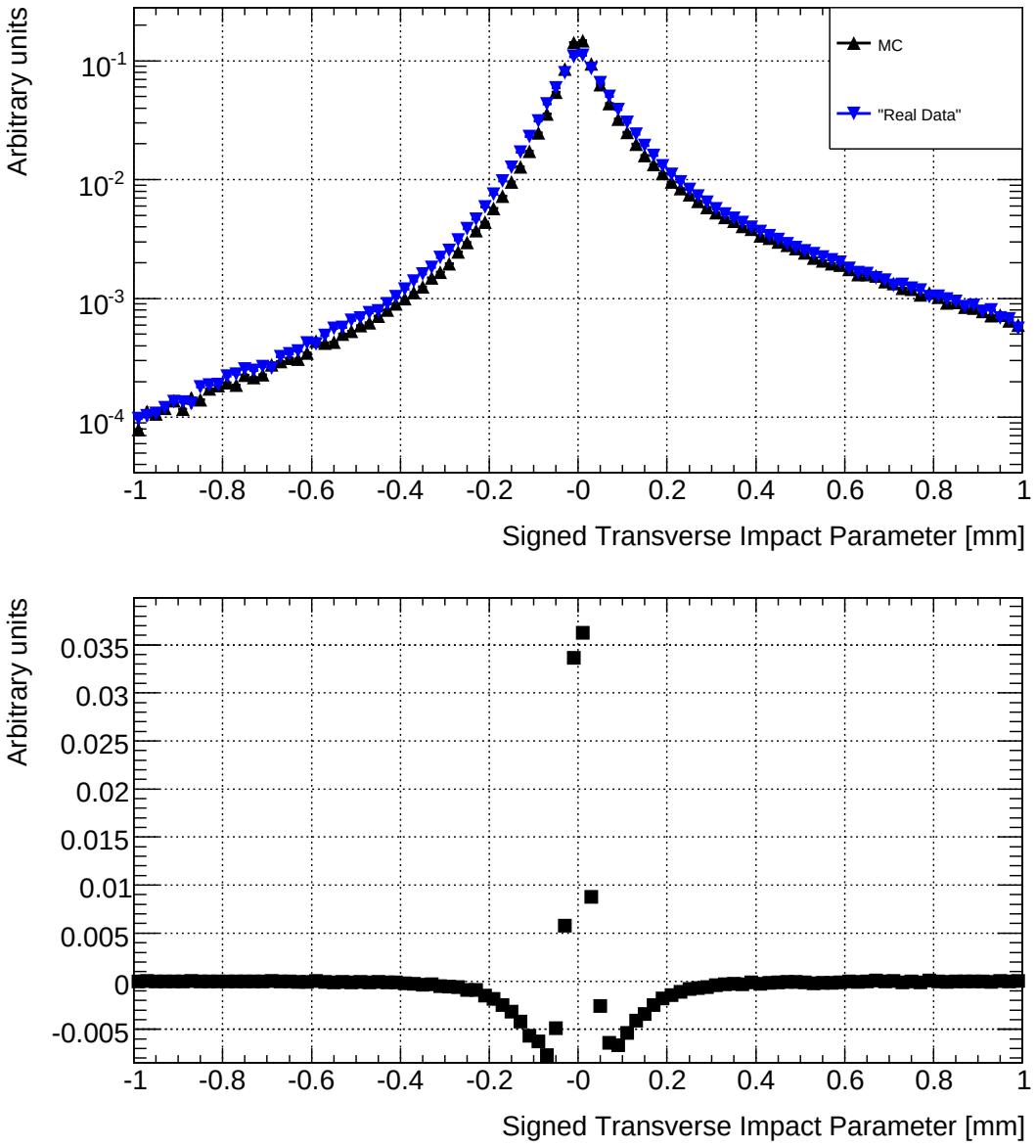

Figure 6.8: At the top: signed transverse track impact parameters for MC and pseudo real data ("Real Data"); At the bottom: difference of MC and pseudo data.

a worse resolution) than the one having a perfect geometry: at small impact parameters, the difference between MC and pseudo data is positive.

A first test of the presented method is illustrated in figure 6.9: a Gaussian distribution (black) is mapped (red) to another Gaussian distribution (blue) having a different width. As expected, the blue and the red distribution are coinciding very well.

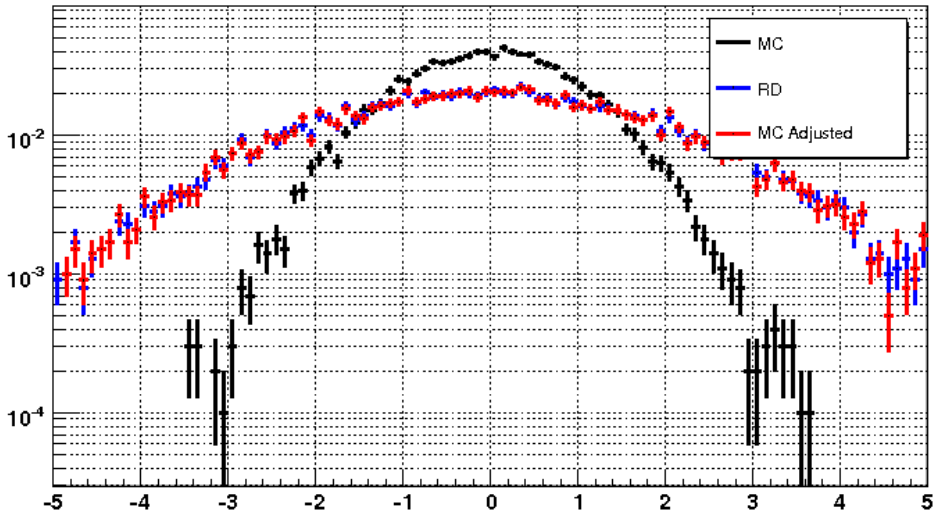

Figure 6.9: First test of the method: a gaussian distribution (black) is mapped (red) to another gaussian distribution (blue).

Since b-tagging is very important for the selection and analysis of top quark events, especially precision measurements of top quark properties, the mapping approach was applied using two top quark pair datasets. Their generated events are the same, but they differ in the alignment used for the reconstruction. The one serving as MC [117] has $6 \cdot 10^5$ events, the other serving as pseudo RD [118] has $5 \cdot 10^4$ events, which is a factor 12 less. In order to provide similar conditions, the MC and 'RD' distributions ("mapping histograms"), which are used to determine Δd_0^{RD} for each MC track, have the same statistics of $5 \cdot 10^4$ events. The bin size of the histograms is chosen as $20\,\mu$m, so that it is a bit larger than the expected resolution of the transverse impact parameter, see table 5.1. Since the signed impact parameters depend on the transverse momentum p_T and pseudorapidity η of the track as explained in Section 6.1, several mapping histograms were created for different p_T and η ranges: 4 bins are used for the p_T (in units of GeV): $1-2$, $2-3$, $3-4$ and >4. In η the following 5 bins were chosen: $(-2.5;-1.9)$, $(-1.9;-0.1)$, $(-0.1;+0.1)$, $(+0.1;+1.9)$, $(+1.9;+2.5)$. As a result, there are 20 mapping histograms, both for the MC and the pseudo RD dataset. Moreover,

these distributions were created for the signed transverse track impact parameter d_0 as well as the longitudinal z_0. Apart from transverse momentum and pseudorapidity the impact parameter distribution depends also on other track properties. For example, tracks that share some of their hits with other tracks represent a set yielding a worse resolution. Such effects are not considered in this first approach.

Figure 6.10 shows that a statistics of about $5 \cdot 10^4$ tracks is sufficient for a successful mapping in the range of $|d_0| < 0.4\,\text{mm}$. When the two difference distributions before and after adjustment (pseudo RD − original MC versus adjusted MC − pseudo RD) are compared, it can be seen that the agreement in this range is up to a factor of 4 ($|-0.008/4| = 0.002$) better after the mapping procedure. However, the adjustment overshoots in this region, because the sign of the difference inverts. Up to now, the reason for this behavior is unknown. For large impact parameters there is nearly no improvement because of the small number of tracks, which are available for the mapping histograms. The statistics in the range $|d_0| > 0.4\,\text{mm}$ is two orders of magnitude lower than for small impact parameters. Furthermore, the statistics is limited by the pseudo RD sample, which has 12 times fewer events than the MC one.

The results are confirmed by further plots provided in Section A.2. In particular, the argument that a failing mapping is caused by low statistics is supported: At high pseudorapidities ($1.9 < |\eta| < 2.5$) the mapping leads to smaller improvements than at low pseudorapidities ($0.1 < |\eta| < 1.9$), where the number of available tracks is about 6 times larger.

Section A.2 provides all 40 histograms, which present the results of the mapping approach done for the chosen datasets: for the transverse and longitudinal signed impact parameter, for the 5 bins in the pseudorapidity and the 4 bins in the transverse momentum of the track.

Since the mapping approach leads to a better agreement between MC and pseudo RD by up to a factor of 4, it is a good candidate to correct MC datasets from first data. Therefore, development and tests of the implemented method are ongoing.

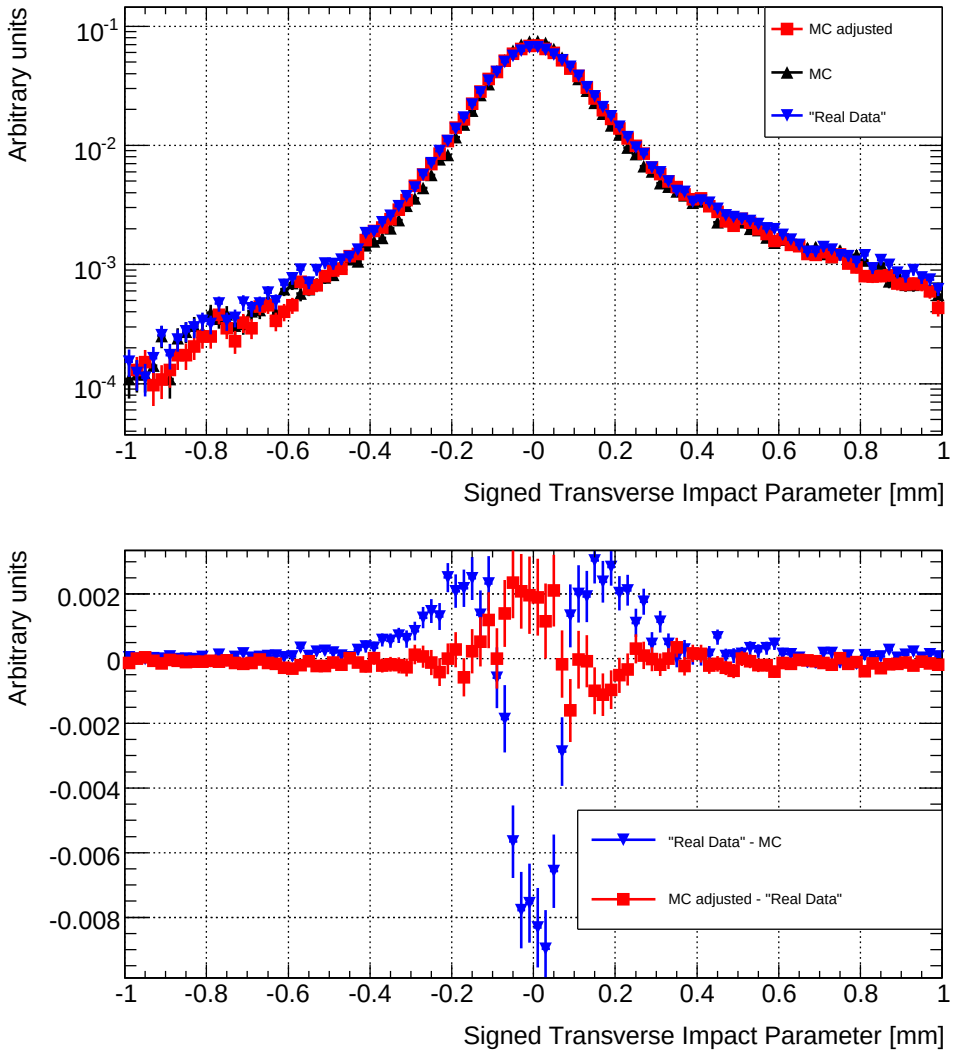

Figure 6.10: Proof of method using about $5 \cdot 10^4$ tracks: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). The mapping is done for tracks with $0.1 < \eta < 1.9$ and $1\,\text{GeV} < p_T < 2\,\text{GeV}$ for all jet types (b- and light). At the bottom: the differences of the distributions on the left. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping (adjusted MC − pseudo real data).

6.4 Alternatives

A different approach is explained in [119] and outlined in figure 6.11. This approach was adopted for DØ [120] and is also in progress for ATLAS [116]. There,

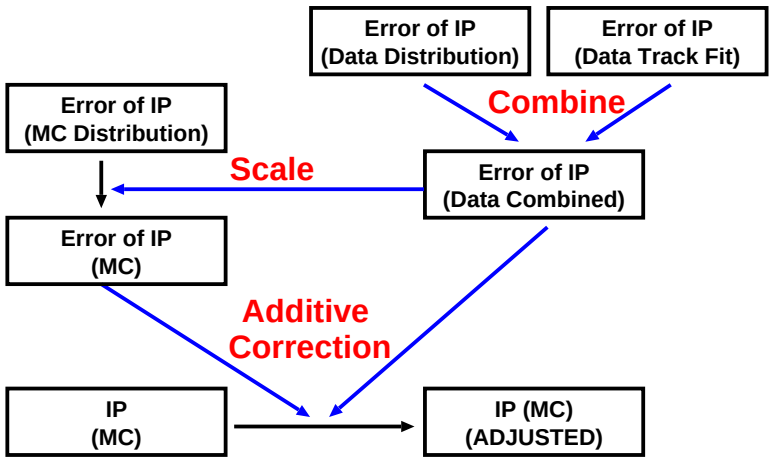

Figure 6.11: Track impact parameter correction approach of Borisov and Mariotti at DELPHI.

the errors of the track impact parameters from the real data distribution and the track fit are combined first to reflect the real precision/understanding of the tracking system. According to this combined error, the error of the MC distribution is corrected. After having corrected the error of the impact parameter, the IP itself can be corrected additively (assuming a gaussian distribution) by using those errors of MC and real data. Finally, the impact parameters of a subset of tracks are adjusted to correct for the non-gaussian tails of the impact parameter distribution. This approach may be also used to correct data, if the error calculated by the reconstruction needs some correction.

It is also possible to make a correction on the clusters associated to tracks. This method would provide adjusted track fits with corresponding errors. It requires rerunning the reconstruction processing step: the corrections could be applied to ESDs containing the cluster information or by reprocessing the raw data. From this adjusted data, AODs can be produced and analyzed afterwards. Therefore, such a procedure may easily take a lot of resources, and more time will be consumed to do the correction iterations, until a sufficient tuning is reached. Hence, the presented mapping approach may be the preferred way of correcting MC track impact parameters, if the correction performance is similar for both alternatives.

Chapter 7

Systematic Uncertainties of Acceptance Corrections

Top quark pair production has a considerable cross section at the LHC ($\sigma(t\bar{t}) \approx 800\,\text{pb}$ in proton-proton collisions at design energy). Due to the quite easy selection and the sizeable signal to background ratio for the semileptonic channel, the ATLAS detector can be in-situ calibrated by large high purity samples aggregated in a few months of data acquisition. Moreover, the top quark properties measured at the Tevatron or predicted by the Standard Model can be validated with high statistics never reached before. Eventually, new physics may be pointed at by analyzing top events.

The precise measurement of top quark properties suffers from limited resolution, detection efficiency less than 100% and the presence of background processes. According to [121], a true distribution $\vec{g}$ (generated) which can be never exactly measured, of a top quark observable may be estimated from the data distribution $\vec{d}$, by subtracting the expected background $\vec{b}$ and applying a multiplicative correction factor $\vec{C}$. This has to be done in bins i of the distribution:

$$g_i = C_i \cdot (d_i - b_i) \tag{7.1}$$

The factors C_i can be determined by a simulation of the experiment. The Monte Carlo program is run once with and once without the detector simulation, yielding model predictions for the observed and true values of each bin, a_i^{MC} (accepted, i.e. generated with reconstruction cuts) and g_i^{MC}. Here $\vec{g}^{\text{MC}}$ refers to the signal process only, i.e. background is not included. The correction factor is then simply the ratio

$$C_i = \frac{g_i^{\text{MC}}}{a_i^{\text{MC}}} \tag{7.2}$$

Statistical errors in the correction factors can be made negligibly small by producing large MC statistics. If, moreover, the effects of the resolution are insignificant, an accepted event stays in the same bin as the corresponding generated

one. Thus, the expected number of entries (signal) in bin i without background is the measured difference between accepted and background entries:

$$a_i^{\text{sig}} = a_i - b_i = \frac{1}{C_i} g_i \tag{7.3}$$

In the ideal case, $C_i = 1$ and therefore $g_i = a_i^{\text{sig}}$. However, every real experiment has a bias, that is the deviation from the true distribution, is

$$s_i = \left(\frac{g_i^{\text{MC}}}{a_i^{\text{MC}}} - \frac{g_i}{a_i^{\text{sig}}} \right) \cdot a_i^{\text{sig}} \tag{7.4}$$

Since g_i is unknown, only a rough estimate of the systematic uncertainty due to this bias can be obtained by computing the correction factors with different Monte Carlo models. Equation 7.4 implies that the systematic uncertainty of the correction procedure vanishes ($s_i = 0$), if the ratio of g_i/a_i agrees for data and MC ($g_i^{\text{MC}}/a_i^{\text{MC}} = g_i/a_i^{\text{sig}}$). However, this can only be checked by comparing the data to the accepted MC. If perfect agreement is reached here and if the detector simulation represents the experiment, it is justified to assume that the correction method is free of any bias.

Though, forthcoming data may differ from Monte Carlo simulations by more than the statistical and systematic uncertainties. When new physics or further systematic errors are excluded, the deviations of the Monte Carlo simulations may arise from an imperfect detector simulation or from the physics simulation (MC generator). Assuming the correctness of the detector simulation, the average variation of the MC generators corresponds to an uncertainty of the physics modelling. The deviations have to be reduced by either tuning the generators or, initially, by reweighting their events such that the model events agree with the data.

The latter approach is chosen and presented in this thesis. The observed number of top/antitop quark events, which corresponds to their total selection efficiency when the luminosity is known from measurements, may be different for MC and data. For example, the systematic uncertainty due to the physics model of MC may be 20%, whereas for data it may be 5% corresponding to detector and theoretical uncertainties of $\approx 2.5\%$ each. The statistical uncertainty can be neglected because of the high production rate at the LHC and if enough MC events are available. But if all MC events are reweighted in an appropriate way, the systematic uncertainty of MC is reduced to the level of data. Hence, the agreement between MC and data is improved for the selection efficiency.

Since the integral over a differential distribution $\frac{dN}{dx}$ provides the total number N,

$$N = \int \frac{dN}{dx} dx \tag{7.5}$$

the total numbers agree, if the differential distributions do. However, the full range of an observable x is normally not accessible in an experiment and extrapolations have to be done, which may be inaccurate. If these are accurate, the approach of reweighting depends on whether the differential distributions of the observables, which are used for selection cuts, are correlated or not. When they are uncorrelated, the MC events can be reweighted by a total weight, which is the product of all weights obtained for the individual distributions. If the observables are correlated, a reweighting from one distribution may suffice.

The ratio of MC and data provides the inverse of the weight, which may be applied bin by bin to the MC sample in order to achieve a better agreement with data. However, such a bin by bin reweighting procedure is unphysical (not smooth) and leads generally to instabilities. Thus, the distribution can be instable, when the method is applied iteratively. Therefore, the ratio of the MC and data distribution is fitted by a smooth function, which may be even extrapolated into regions that are experimentally not accessible. The fit function can be arbitrarily chosen, if it is smooth, fits the ratio of the distributions and leads to an improvement of the agreement. The systematic uncertainty due to the choice of the fit function can be estimated by using different functions for the reweighting of the MC events.

The next sections explain the simulation of $t\bar{t}$-events, the selection of semileptonic events and the corresponding efficiencies, the distributions of the relevant observables, the reweighting approach and its results.

7.1 Monte Carlo Simulation

As explained in Section 2.3, Monte Carlo computer programs also called MC generators, that approximate the analytical theory by empirical models, are used to simulate the measurement of single collision events. In ATLAS, several MC generators and many related computer packages, which contain for example PDF sets, are used [122]. Typically, the MC generator output are the four-vectors of all final state particles of a collision event. However, the Monte Carlo models do not take statistical and systematic uncertainties, which result from the limited resolution of the measuring device, into account. So, the measured distributions will be different from the initially predicted distributions.

To reconcile the experiment with the theory, the detector must be simulated. In figure 7.1, all steps of the simulation data processing, also called the ATLAS MC full chain, are sketched. The common steps for real data processing are noted as well.
The detector simulation gets the predicted observables (i.e. the four-vectors of all final state particles) from the MC generator and returns a simulated detector response. In ATLAS, the detector simulation GEANT4 [123] is used. The exact detector material description including geometry and masses is provided to this

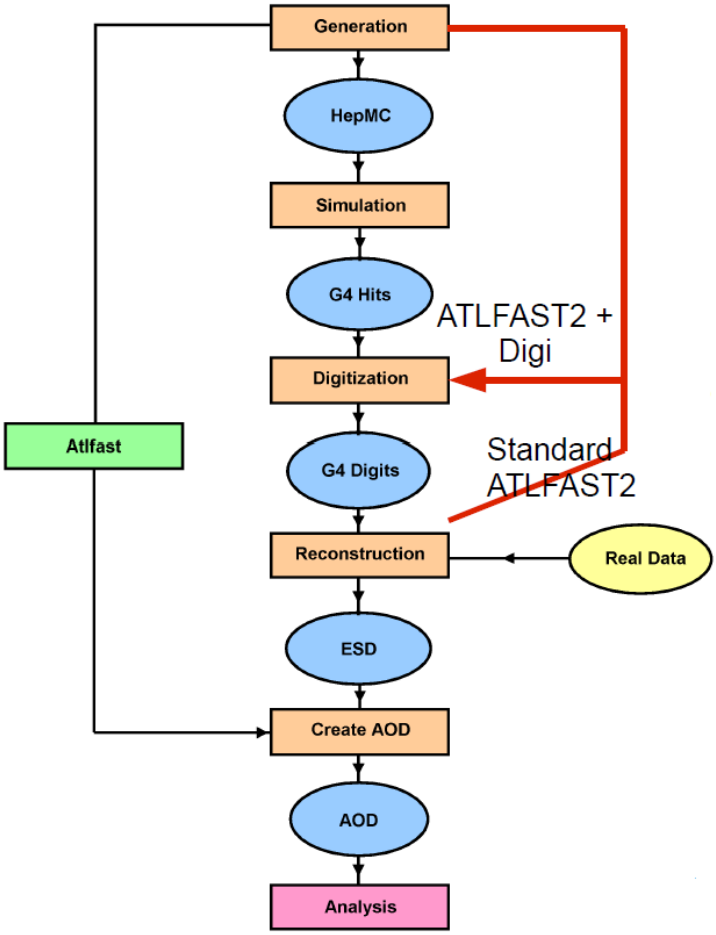

Figure 7.1: Program and data flow from MC event generation to analysis (the ATLAS full chain Monte Carlo production). The flow of fast simulation (ATLFAST) and real data is noted as well. [124]

program. GEANT approximates the exact geometry by a composition of many simple geometrical objects as cuboids. For each particle provided by the MC generator GEANT produces "hits", which are energy deposits in sensitive detector volumes according to the Bethe stopping power formula or other models.
Another computer program, called digitization, converts these energy deposits into detector responses, "digits", typically voltages or times on preamplifier outputs. More details about this digitization step can be found in [125]; the overall documentation of the generation and the digitization step, also called simulation, can be read in [126]. The simulated and digitized detector response should provide the same distributions of the experimental observables as the measurement.

Since such a simulation of the ATLAS detector is very time consuming, it may be appropriate to simulate the detection of particles only in a simplified way, which is much faster. Therefrom, the names of the corresponding AT-

LAS packages were chosen: ATLFAST [127] and FATRAS [128], which simulates tracks. To find a balance between speed and consumption of resources on the one hand and accuracy on the other, full and fast simulation can be combined by ATLFAST2 [124], which may also do the digitization but skip the intensive simulation step.

Instead of comparing raw detector responses, the MC event is reconstructed just like a real data event as explained in Chapter 5: Raw data stored in RDO files is used to reconstruct common physics objects like tracks, vertices, jets, electron, photons, muons, missing transverse energy. These objects can be saved together with some raw information in ESD files, wich are mainly used for detector calibration. Physics analyses prefer AOD or DPD files, which are small in size because they contain almost no raw information.

Since the reconstruction algorithms are identical for simulated and real collision events, the distributions of the experimental observables should be the same for simulated and real data. However, MC simulation can neither predict unknown physics nor reproduce precisely all detector effects. Moreover, various MC generators or detector simulations use different models. Therefore, deviations from real data represent the systematic uncertainty caused by the modelling, if systematic uncertainties in the data are minimized.

In order to study such uncertainties, 5 MC generators were chosen to simulate $t\bar{t}$-events in proton-proton collisions with a center of mass energy of 14 TeV. Each of the generators mentioned in Section 2.3, Pythia, AcerMC, Sherpa, Herwig, MC@NLO, simulated about 10^5 events each. Since this chapter presents a method, how systematic uncertainties due to the physics modelling can be reduced by reweighting MC events from (pseudo) data, numbers of events were not normalized to an integrated luminosity, and any effect of pile-up was neither simulated nor considered. Though for AcerMC the hadronization step was done by Pythia and Herwig was used for this purpose for MC@NLO, we stick in this thesis to the short notations "AcerMC" and "MC@NLO".

The configuration values of the generators were chosen as default values either of the generator for LHC or of ATLAS. After the generators had been configured in a similar way, for example they used the same top quark mass of 175 GeV or similar PDF sets (CTEQ6L1 or CTEQ6M [129, 130]), they were run within the same version of the Athena framework to guarantee a maximal consistency. The detector geometry and simulation as well as the event reconstruction are exactly the same for all 5 generators and defined by the corresponding version of the Athena environment. All configuration files and scripts, which were used to generate the 5 datasets [131], can be obtained at [132]. Table 7.1 overviews all steps of the simulation. In the following section we will explain the event observables, selection and analysis.

Step	Athena version	generator	generator version
generation	14.4.0.1	Pythia	6.419
		Herwig	6.510
		AcerMC	3.5
		MC@NLO	3.31
		Sherpa	1.1.2.2
detector simulation	14.4.0		
digitization	14.5.1		
reconstruction	14.5.1		
analysis	14.5.0		

Table 7.1: Configuration of 5 Monte Carlo $t\bar{t}$ simulations. The steps detector simulation to analysis use the same Athena version and same detector geometry (ATLAS-GEO-02-01-00) for all 5 generators.

7.2 Total and Differential Efficiencies

As mentioned in section 2.2, about 44% of top quark pairs decay semileptonically. Usually, the final state consists of two b-jets and two light jets, one lepton and one neutrino, apparent as missing transverse energy. High level reconstructed objects such as electrons, muons or jets, whose reconstruction was explained in Section 6.1, can be defined in several ways. To achieve comparability, the standard definitions of ATLAS for top quark physics were chosen [114, 133], for example a cone jet algorithm with an opening angle of 0.4 rad was used. Since reconstruction algorithms of different objects run independently over the raw data, the same hit pattern may be reconstructed as more than one object. For example, the same calorimeter hits can be identified as an electron and as a jet, or generated events are crowded by soft Bremsstrahlung photons (and therefrom also by soft electron/positron pairs due to pair production) from electrons, which would be seen altogether in the detector only as one narrow electromagnetic shower. Therefore, a procedure may be necessary to decide, which of the overlapping objects are selected and which are removed, hence its name "overlap removal". In this analysis, it was implemented also according to [114, 133] and referred as selection cut 1.
Like in [114, 133], the following observables and corresponding cuts due to kinematics (and for a reduction of background events) are used to select semileptonic $t\bar{t}$-events:

- one isolated lepton (only e or μ because a τ can decay hadronically) with high transverse momentum $p_T > 20\,\mathrm{GeV}$ (selection cut 2)

- large missing transverse energy $\not{E}_\mathrm{T} > 20\,\mathrm{GeV}$ (selection cut 3)

- at least 4 jets with high transverse momentum $p_T > 20\,\text{GeV}$ (selection cut 4)
- of which at least 3 have an even larger transverse momentum $p_T > 40\,\text{GeV}$ (selection cut 5)

Further requirements can be that 2 jets have an invariant mass inside a window around the W boson mass or there are one or two b-tagged jets (only useful for later data when all detector effects are fully understood). Although jets can be measured in the range of the pseudorapidity $|\eta| < 3.2$ by the ECAL/HCAL or even up to $|\eta| < 4.9$ by the FCAL, b-tagging is possible only in the range $|\eta| < 2.5$, which corresponds to the acceptance of the inner detector. Therefore, only those events are selected, whose aforesaid objects (electrons, muons, jets) are all within $|\eta| < 2.5$. Since this study is a first approach, no trigger information are used. The resulting selection efficiencies of the 5 cuts above for events without (generator level) and with detector simulation (reconstruction level) out of a semileptonic sample are illustrated in table 7.2. For a certain MC generator and selection cut, the relative deviations between generator and reconstruction level are up to 13%. For example, by the fourth cut (≥ 4 jets with $p_T > 20\,\text{GeV}$) 78% (19.1/24.5) of the events on generator level are kept, whereas 68% (19.4/28.5) are kept on reconstruction level, corresponding to a relative deviation of 13% for Herwig. Such deviations may be explained by the acceptance and resolution effects of the detector simulation and the fact, that there is for example no precise counterpart of a jet in both levels. Likewise, the large acceptance differences between generator and reconstruction level in the first cut are caused by soft particles, which are removed on generator level.
Similarly, the relative deviations between two MC generators for the same selection cut represent up to 11%, which can be explained by the different modelling of the physics process by the generators. The maximal deviation between the total efficiencies is 20% (12.7/15.8) for Pythia versus Sherpa after the fifth cut on reconstruction level. This difference is mainly due to Pythia's new parton shower algorithm, which is much more radiative. The statistical and systematic uncertainty for data will probably be much smaller ($\lesssim 5\%$). Therefore, a reweighting of MC events is already necessary from the point of view of total efficiencies.

The presumably most precise reconstruction of a semileptonic $t\bar{t}$-event is a kinematic fit. As indicated in Section 2.2, two light jets must come from a W decay and have a corresponding invariant mass. Similarly, the missing transverse energy and a charged lepton have to come from another W decay. The charge of the lepton determines, if it comes from the top or antitop quark. Furthermore, two jets have to be b-tagged. Each of them must have together with one of the reconstructed W bosons an invariant mass of a top quark. Such a fit may be challenging, especially for early data, when the detector is not well understood. Hence, this reconstruction strategy is not chosen in this analysis. Since the reconstruction of the leptonically decaying top quark depends on three or four

different object types (missing transverse energy, jet, electron or muon), it is neglected here. Especially the missing transverse energy will be not well measured

Gen. Cut [%]	Pythia	Herwig	AcerMC	MCNLO	Sherpa
none no. of events	100 (≡ 38348)	100 (≡ 40658)	100 (≡ 41777)	100 (≡ 28198)	100 (≡ 96711)
stat. error	±0.2	±0.2	±0.2	±0.2	±0.1
overlap removal	91.8	89.8	91.1	90.6	90.0
isol. lepton $p_T > 20\,\text{GeV}$	25.8	25.7	25.9	25.8	26.2
$\not\!\!E_T > 20\,\text{GeV}$	24.5	24.5	24.6	24.6	24.7
≥ 4 jets $p_T > 20\,\text{GeV}$	20.2	19.1	20.1	19.4	18.9
≥ 3 jets $p_T > 40\,\text{GeV}$	15.2	13.4	15.0	13.6	12.9

Reco. Cut [%]	Pythia	Herwig	AcerMC	MCNLO	Sherpa
none no. of events	100 (≡ 38345)	100 (≡ 40648)	100 (≡ 41758)	100 (≡ 28188)	100 (≡ 96677)
stat. error	±0.2	±0.2	±0.2	±0.2	±0.1
overlap removal	99.9	99.9	99.9	99.9	99.9
isol. lepton $p_T > 20\,\text{GeV}$	29.7	31.0	30.0	30.5	30.5
$\not\!\!E_T > 20\,\text{GeV}$	27.3	28.5	27.6	27.9	27.8
≥ 4 jets $p_T > 20\,\text{GeV}$	20.1	19.4	20.0	19.6	18.1
≥ 3 jets $p_T > 40\,\text{GeV}$	15.8	14.4	15.7	14.6	12.8

Table 7.2: Event selection efficiencies out of semileptonic samples of five MC generators. As well, the numbers of events and the corresponding statistical error are noted, which applies approximately to the efficiencies of a column. At the top: without detector simulation (generator level). At the bottom: with detector simulation (reconstruction level).

in early data. Whereas, the hadronically decaying top quark relies only on the object type jet (1 b-jet and 2 light jets from the W decay). According to [114], we define this top quark as that combination of three jets (out of N jets, hence there are $\frac{N!}{3!\cdot(N-3)!}$ possible combinations), which has the highest vector sum of their transverse momenta:

$$\sum_{j=1}^{3} \vec{p}_{T_j} = max \tag{7.6}$$

However, this loose definition selects quite probably a wrong jet combination near the threshold of $t\bar{t}$ pair production and neglects b-tagging and other constraints. In spite of the corresponding combinatorial background, this simple definition may be sufficient for first data and represent a stable analysis.

For the differential distributions of the top quark the expected resolutions were determined from two dimensional plots: The difference of the reconstructed and generated observable is plotted against the generated observable. An example of a resolution plot for the transverse momentum of the top quark is illustrated in figure 7.2. The binning of the plots of the differential cross sections was chosen a

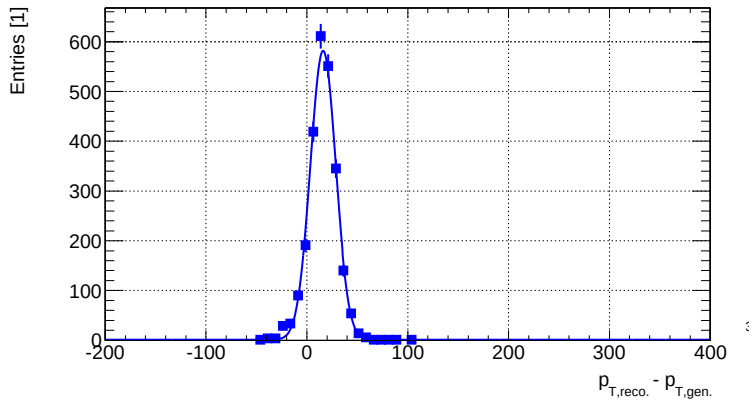

Figure 7.2: Example of a resolution plot: The expected resolution (reconstructed – generated) of the transverse momentum of the hadronically decaying top quark for Pythia in the range of 200 to 220 GeV (generated) is about 13 GeV.

bit coarser and more regular than the obtained resolutions, see figure 7.3. This binning ensures that bin by bin migration between generated and accepted events in the MC simulation are small.
Since LHC and ATLAS had not collected any collision data during these studies, Sherpa was arbitrarily chosen to represent pseudo data. As patent in figure 7.4, relative to pseudo real data there are deviations only up to 5% in bins of the pseudorapidity of the hadronically decaying top quark. Hence the MC generators agree quite well in this observable, even with detector simulation. However, in bins of the transverse momentum there are deviations up to a factor of 3, see figure 7.5. Due to this discrepancies, a reweighting of the MC events, for example from this observable, is mandatory.

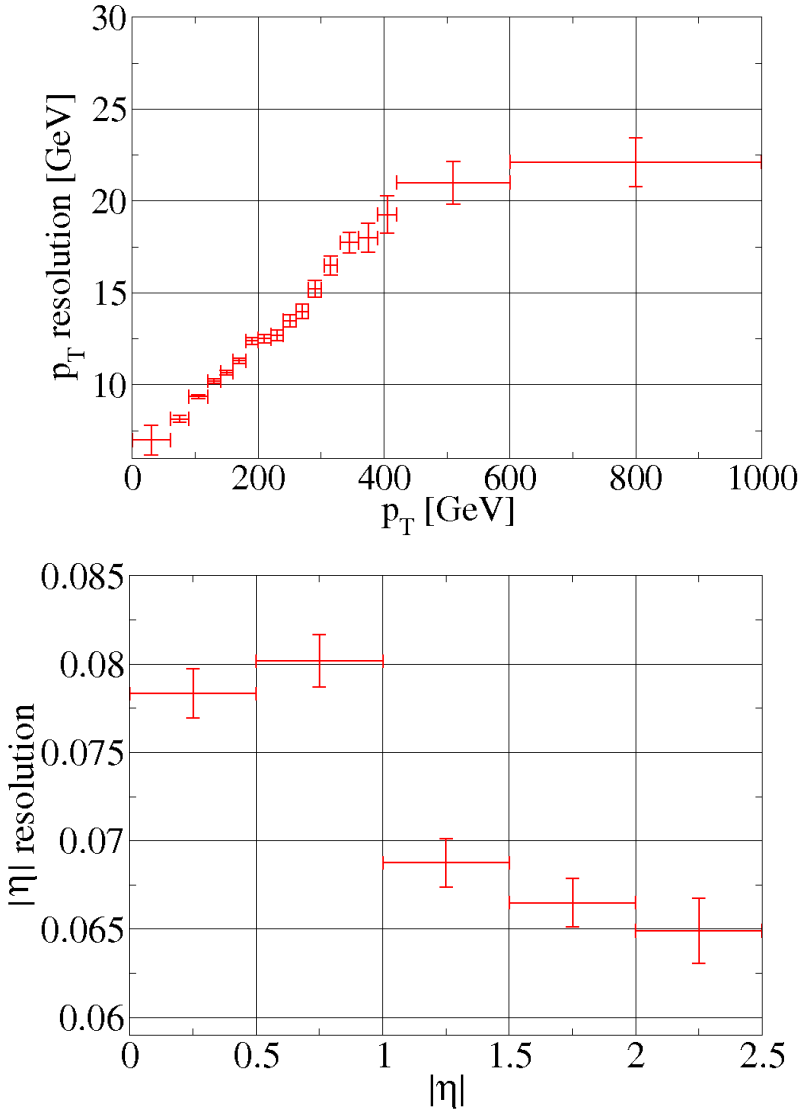

Figure 7.3: The expected resolutions (reconstructed – generated) of the transverse momentum and pseudorapidity of the hadronically decaying top quark for Pythia.

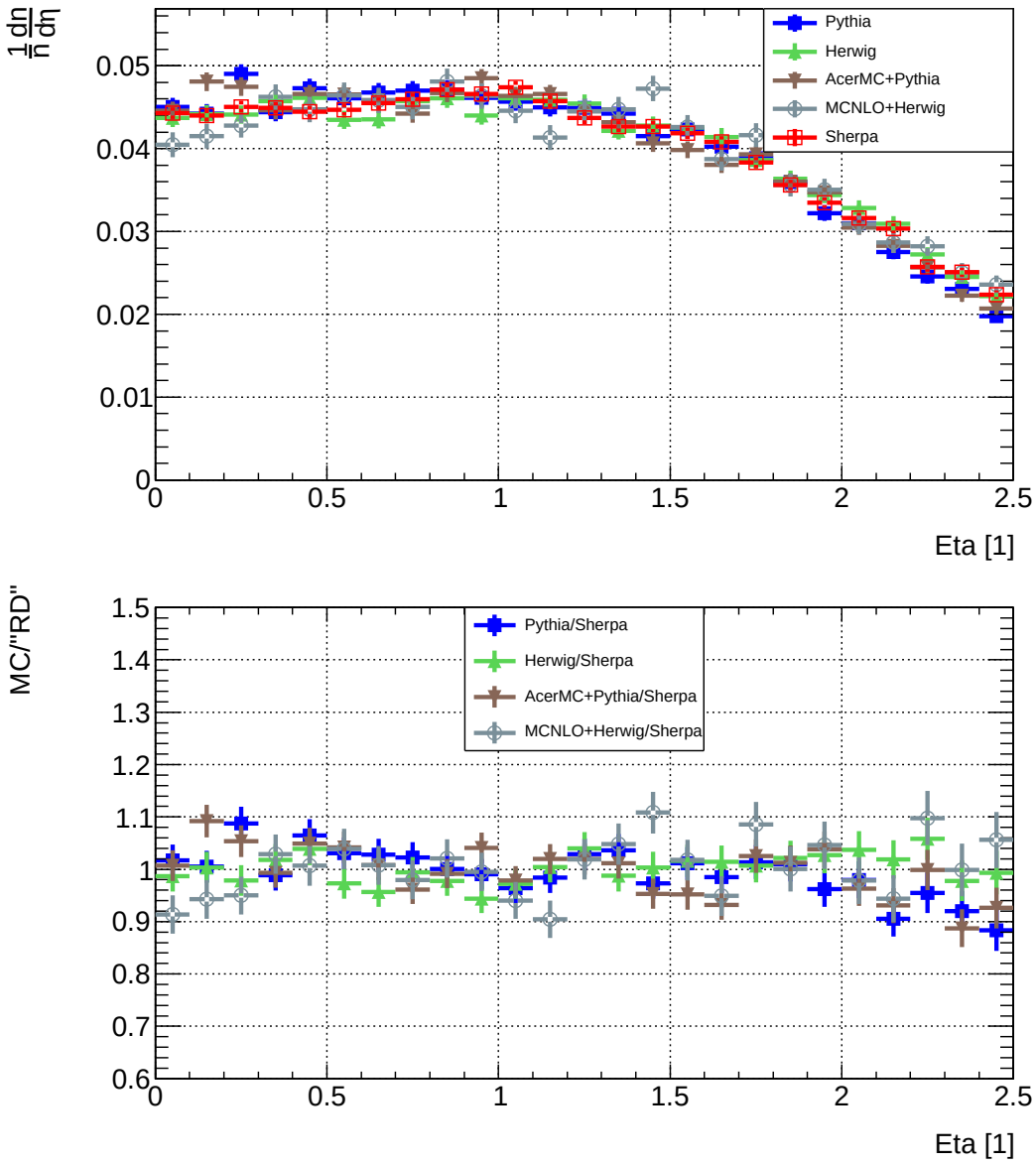

Figure 7.4: Distributions of the pseudorapidity of the hadronically decaying top quark on generator level without any cuts. On reconstruction level (with cuts) the plots are very similar, see B.3. At the top: for five Monte Carlo generators. At the bottom: four of them relative to the last (Sherpa arbitrarily chosen as pseudo data).

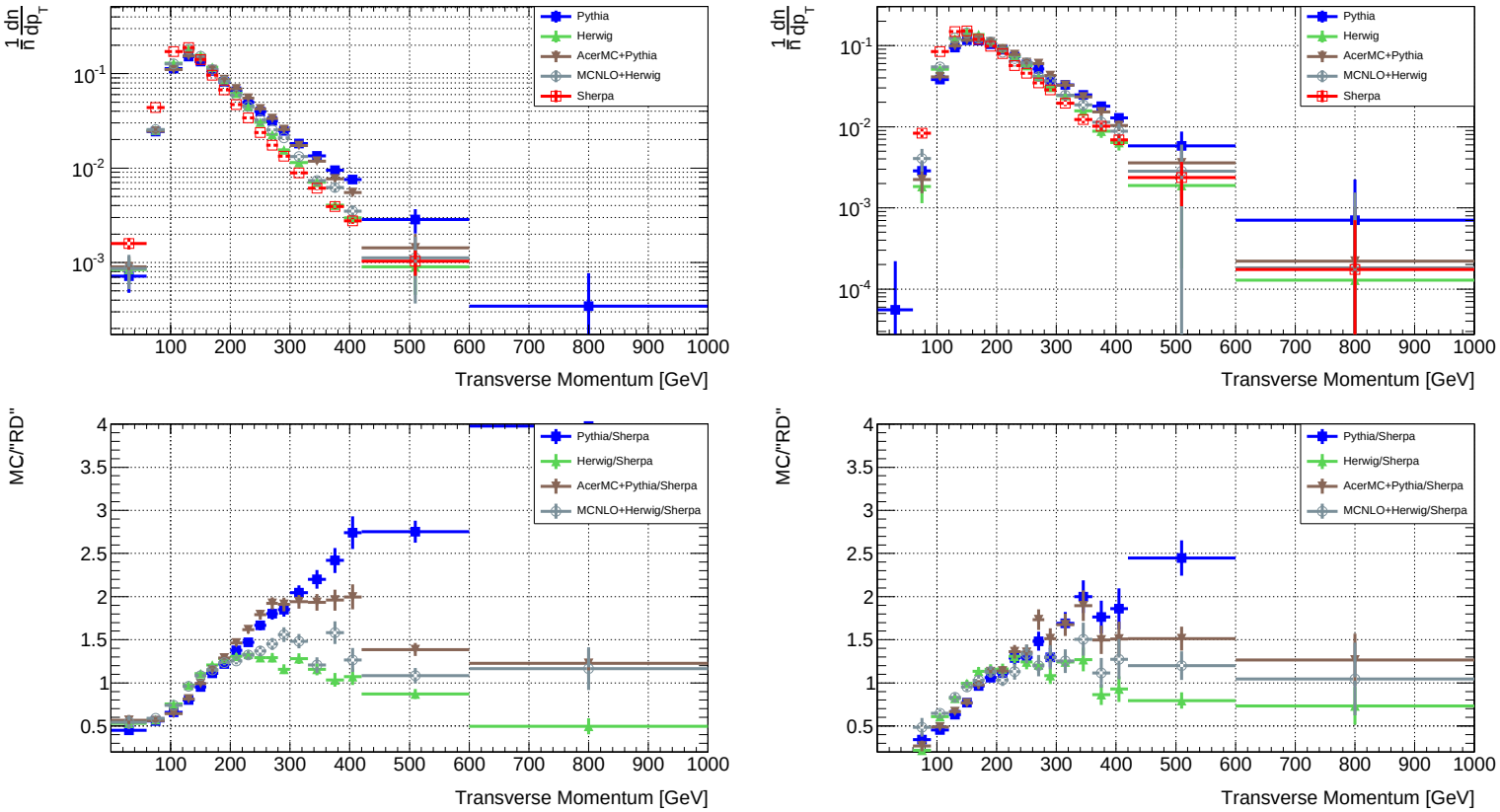

Figure 7.5: Distributions of the transverse momentum of the hadronically decaying top quark for five Monte Carlo generators without detector simulation (generator level without any cuts) at the left, with detector simulation (reconstruction level with cuts) at the right. at the bottom. The corresponding ratio plots are shown as well (Sherpa arbitrarily chosen as pseudo data).

7.3 Reweighting of MC Events from (Pseudo) Data

As announced in the previous sections, a reweighting of the MC events from (pseudo) data may be appropriate and sufficient to reduce the uncertainty of the physics modelling (several MC generators with different predictions for the same process). Though, it is quite unprobable to obtain good agreement between MC and (pseudo) data for the total and all relevant differential efficiencies. This may be even so after the MC events have been reweighted. The main reasons are limited experimental resolution, noise and the fact, that only parts of the full momentum and rapidity ranges can be measured with ATLAS. Particularly, distributions have to be extrapolated into the low energy/momentum range, where measurements are not possible. However, if the MC generators have a similar approach for the important stages of event generation (explained in Section 2.3), and differ only in one specific domain or variable, it is likely that the different MC models agree reasonably well after the reweighting according to a single, well chosen observable.

In fact, the above situation is realized. There are three major elements in the generation of a $t\bar{t}$ event: the underlying PDFs, the hard strong interaction $t\bar{t}$-production process, and the weak W decay. The PDFs are by construction similar in the performed simulation, and should mainly influence the longitudinal properties (η distribution) of the $t\bar{t}$ event. The decay of the W is theoretically very well understood and mis-simulation is unlikely. The largest differences are expected for the strong production process, already due to the differing values of the strong coupling and the ansatz (matrix element, parton shower, matching of both) of the models.

Variations due to the hadronization model should be of minor significance, as the mass scales of the involved processes — the top mass with respect to typical hadron masses — are substantially different. Thus, an observable for the reweighting is suggested, which shows large differences between the models and is sensitive to the strong production process. Therefore, the hadronically decaying top quark was chosen, which is obviously sensitive to the strong production process and is defined according to formula 7.6. Since the distributions of the pseudorapidity and the azimuthal angle for that observable agree quite well between (pseudo) data and MC, they are not considered in this first approach. However, the transverse momentum of the top quark differs up to a factor of 3 from pseudo data (Sherpa arbitrarily chosen) and is therefore the obvious candidate.

The reweighting procedure is neither stable nor physical, when the MC events are reweighted bin by bin from the ratio of the MC and pseudo data distribution (this implies detector simulation and all selection cuts mentioned in Section 7.1 and 7.2). Therefore, it is fitted for each MC generator. A combination of 2

straight lines ($y_{1,2} = a_{1,2} \cdot x_{1,2} + b_{1,2}$ for two $x = p_T$ ranges 1 and 2) was chosen, representing the most simple smooth function describing the ratio. Figure 7.6 illustrates such fits for Herwig. The corresponding figures for the other MC generators are in Appendix B.2. Some data points do not lie on the fit function,

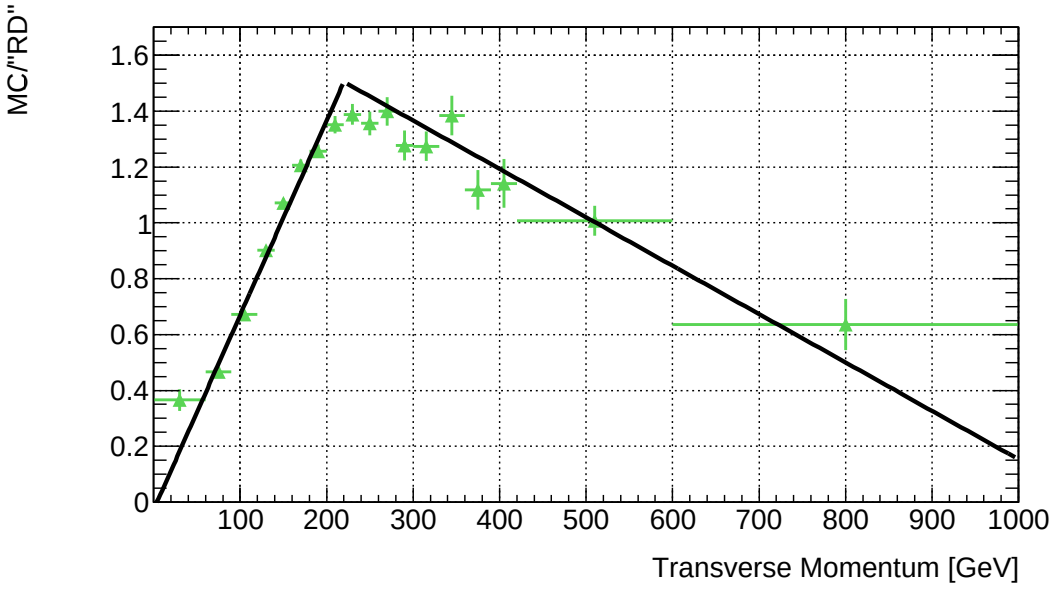

Figure 7.6: The fits to the distribution of the transverse momentum of the hadronically decaying top quark for Herwig relative to pseudo data (Sherpa). The fit parameters are chosen so that the overall function is smooth.

especially at low transverse momentum the fit function lies beneath the data. Though an extrapolation to values near zero may lead to bad results for this reason, a reweighting was tried.
The weight, which has to be given to the MC events, is provided by following formula, where a and b are coefficients from the straight line fit and $p_T^{\mathrm{MC,0c}}$ is the transverse momentum of the hadronically decaying top quark (as defined by equation 7.6) without detector simulation and without any cuts:

$$w = \left(a \cdot p_T^{\mathrm{MC,0c}} + b\right)^{-1} \tag{7.7}$$

If the MC event has already a weight (for example MC@NLO produces weighted events), the total new weight is just the product of the old and the calculated weight.

After all MC events had be reweighted, the ratio between MC and (pseudo) data is nearly flat and as expected about one for the transverse momentum of the hadronically decaying top quark. Figure 7.7 illustrates the improvement with and without detector simulation. This shows that the chosen simple fit strategy works well.

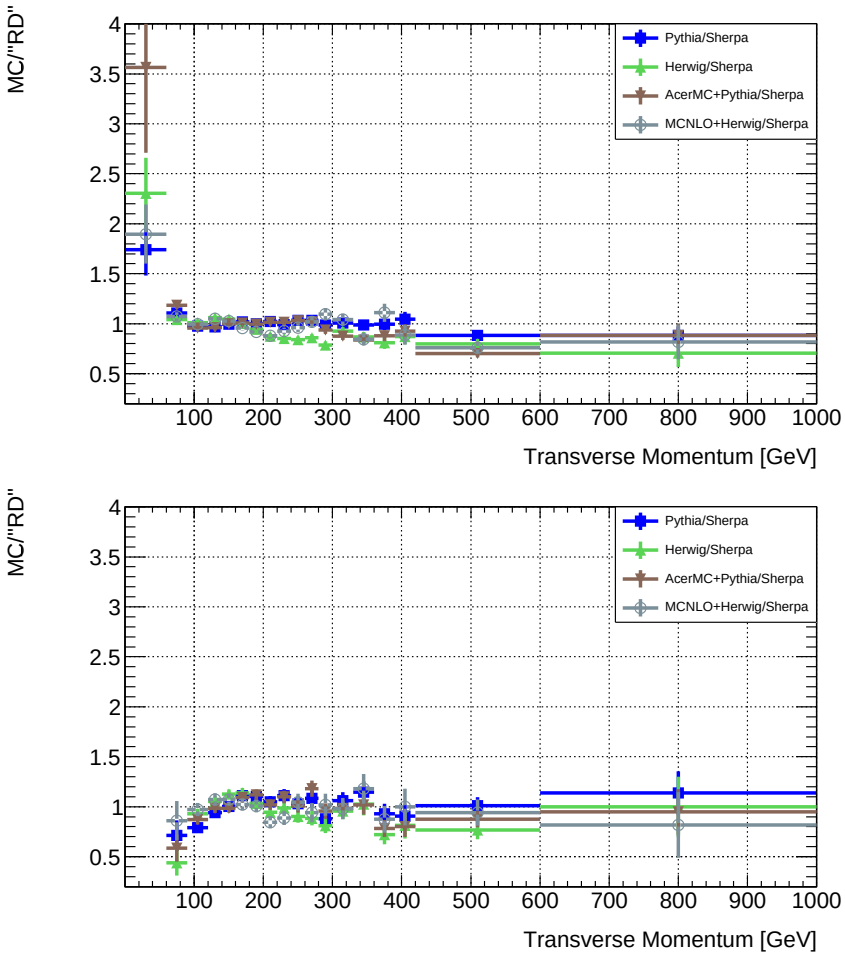

Figure 7.7: Reweighted distributions of the transverse momentum of the hadronically decaying top quark for four Monte Carlo generators relative to pseudo data, without detector simulation (generator level) and any selection cuts at the top, with detector simulation (reconstruction level) and selection cuts at the bottom.

Especially at the connection point of both straight lines, where the fit function is not differentiable, the reweighting provides not worse results than in other ranges. Moreover, the extrapolation in the low momentum region seems not to fail, because the distribution with detector simulation is still below pseudo data. Though the MC events were strongly weighted up in that range, what can be seen in the distributions without detector simulation, a migration of reconstructed events to higher bins is apparently unprobable.

Moreover, the distribution of the pseudorapidity of the hadronically decaying top quark does not change by the reweighting, as it is independent of the transverse momentum. Finally, all distributions of observables serving as selection cuts show for each bin deviations between MC and pseudo data of about the same

size as before reweighting or even up to a factor of 10 smaller. The corresponding figures are patent in the appendix B.2. The improvement is very pronounced for the transverse momentum of the jets (figure B.10), especially the leading four jets (figures B.11, B.12, B.13, B.14). This is expected due to the definition of the hadronically decaying top quark by the transverse momentum of three jets, see formula 7.6. An improvement can be also seen for the reconstructed missing transverse energy (figure B.8). Since the transverse energy is mainly calculated from the calorimeter signal after a subtraction of the jets, an improvement of that observable is reasonable.

However, the deviations for the p_T spectrum of electrons and muons (figures B.4 and B.6) stay in the same range. As high energetic electrons/muons originate from the leptonically decaying top quark, there may be only a reduced correlation to the transverse momentum of the hadronically decaying top quark, and hence no improvement for those distributions is visible. Similarly to the hadronically decaying top quark, the deviations between the distributions of the pseudorapidities of all object types remain as before reweighting.

Overall, the observed behavior is consistent with the conjecture on page 88 that the differences between the models are predominantly due to differences in the hard production process. The same is to be expected for comparisons of the models to real data as weak W decays are very well understood and longitudinal properties of top production seem less important in view of the reconstruction efficiencies.

To analyze the improvement of the total efficiency, we define it as the number of reconstructed data events accepted after N selection cuts (D^{Nc}) divided by the number of generated (no detector simulation) events without any cuts (MC^{0c}):

$$\epsilon = \frac{D^{Nc}}{MC^{0c}} \tag{7.8}$$

Table 7.3 illustrates the resulting improvements gained by the reweighting. For example, the deviation of the total efficiency of Herwig from pseudo data drops from +13% to 0%. The mean improvement is obtained by averaging over the four generators: the deviation of this average generator from pseudo data drops from $\approx +18\%$ to $\approx +3\%$. The mean deviation over the four generators is $\approx +3$ $(= (+4 + 0 + 6 + 3)/4)$. Therefore, the aimed uncertainty smaller than 5% is reached, except other models (generators) deviate even more than the chosen four. However, another generator that deviates even stronglier than Pythia or Herwig, which are very mature, should not be taken too serious.

Hence, there is a quite unexpectedly large improvement for the total and many differential efficiencies by such a simple reweighting approach using only the distribution of one observable. This may be explained as above. However, the improvement may depend on the extrapolation into the low momentum range.

In order to estimate the error of the approach, the reweighting is tried by using the transverse energy H_T of the $t\bar{t}$ pair, represented by the four leading

$\frac{D^{Nc}}{MC^{0c}}$ [%]	Pythia	Herwig	AcerMC	MCNLO	av. Gen	Sherpa
no. of events	35070	36412	37837	24412	33433	86989
stat. err.	±0.2	±0.2	±0.2	±0.2	±0.2	±0.1
unweigh.	15.8	14.4	15.7	14.6	15.1	12.8
rel. to S.	+23	+13	+23	+14	+18	0
reweigh.	13.3	12.8	13.6	13.2	13.2	12.8
rel. to S.	+4	0	+6	+3	+3	0

Table 7.3: Total selection efficiencies of semileptonic $t\bar{t}$ events before and after reweighting from pseudo data (Sherpa) for four Monte Carlo generators. Their relative deviations from pseudo data are given as well. The sixth column represents the average generator of these four and its deviation from pseudo data (not the average deviation of the four generators). The corresponding statistical errors are also noted, which apply approximately to the efficiencies of a column.

jets, the charged lepton and the missing transverse energy. Since the reweighting on the basis of the transverse momentum of the hadronically decaying top quark shows no improvement for the differential efficiencies of the transverse momentum of muons and electrons, a reweighting on the basis of H_T seems a priori to be doomed. The corresponding fit for Herwig, again with two straight lines, is illustrated in figure 7.8 and looks not better than the one in figure 7.6.

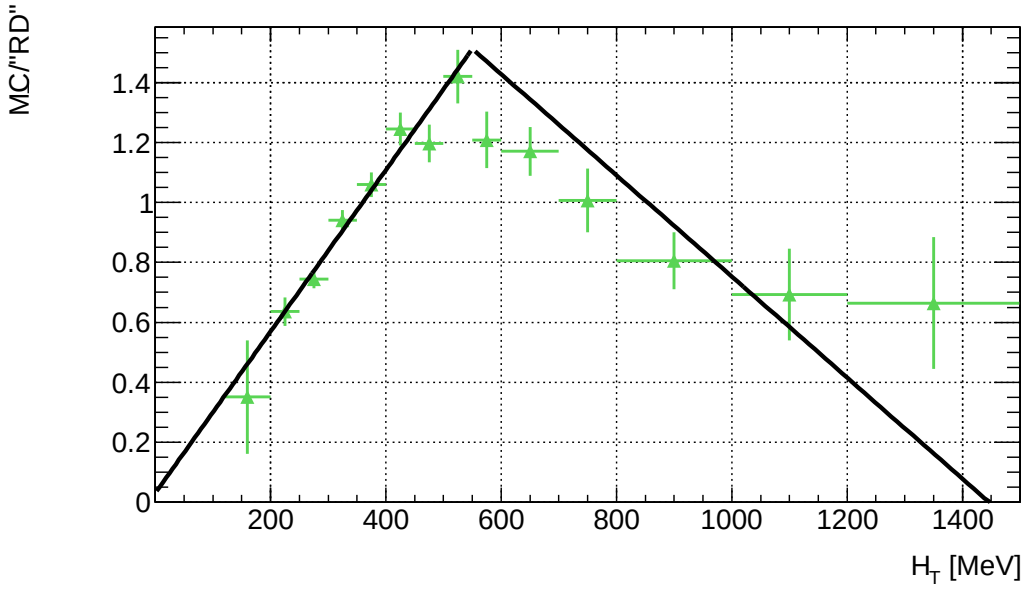

Figure 7.8: The fits to the distribution of the transverse energy H_T for Herwig relative to pseudo data (Sherpa). The fit parameters are chosen so that the overall function is smooth.

The obtained results are similar to those of the standard approach: as illustrated in figure 7.9, the reconstructed H_T (with detector simulation and cuts) is

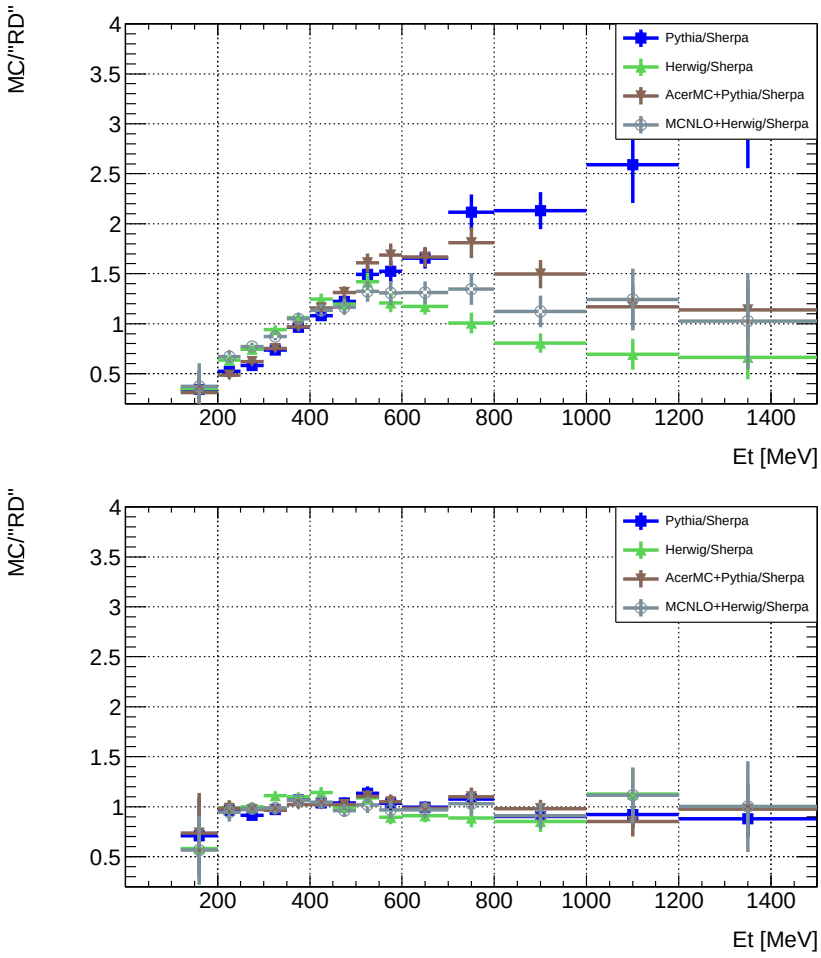

Figure 7.9: Original (at the top) and reweighted (at the bottom) distributions of the transverse energy H_T for four Monte Carlo generators relative to pseudo data, with detector simulation (reconstruction level) and selection cuts.

quite flat and near 1 after the reweighting procedure. For the total efficiencies even a slightly better improvement is obtained. As noted in table 7.4, the deviation of the average generator from pseudo data, which is obtained by averaging over the four generators, drops from $\approx +18\%$ to $\approx -1\%$. The mean deviation over the four generators is $\approx -1\%$ $(= (0 - 4 - 2 + 3)/4)$.

Therefore, the estimated obtainable systematic uncertainty of $\approx 3\%$ is reliable, and the first reweighting approach looks very promising to reduce potential deviations of MC from first data.

$\frac{D^{Nc}}{MC^{0c}}$ [%]	Pythia	Herwig	AcerMC	MCNLO	av. Gen	Sherpa
no. of events	35070	36412	37837	24412	33433	86989
stat. err.	±0.2	±0.2	±0.2	±0.2	±0.2	±0.1
unweigh.	15.8	14.5	15.8	14.6	15.2	12.9
rel. to S.	+22	+12	+22	+13	+18	0
reweigh.	12.9	12.4	12.7	13.3	12.8	12.9
rel. to S.	0	−4	−2	+3	−1	0

Table 7.4: Total selection efficiencies of semileptonic $t\bar{t}$ events before and after reweighting from H_T of pseudo data (Sherpa) for four Monte Carlo generators. Their relative deviations from pseudo data are given as well. The sixth column represents the average generator of these four and its deviation from pseudo data (not the average deviation of the four generators). The corresponding statistical errors are also noted, which apply approximately to the efficiencies of a column.

7.4 Alternatives

A different approach was suggested to estimate the systematic uncertainty of the total selection efficiencies of top/antitop quark events: a chi-square method. The chi-square of an observable is defined as the ratio between the squared deviation of MC from data RD and the sum of their squared errors. For a distribution with bins i, this has to be summed over all bins N:

$$\chi^2 = \sum_i^N \frac{(MC_i - RD_i)^2}{\sigma^2_{\mathrm{MC},i} + \sigma^2_{\mathrm{RD},i}} \tag{7.9}$$

If the chi-square is calculated for several MC simulations, an estimate of the systematic uncertainty can be obtained by fitting a plot of the chi-square of the observable in dependence of the observable: The distance between the position of the RD entry (data has $\chi^2 = 0$) and the position of the minimum of the parabola, which is fitted to the MC values and forced to be also at $\chi^2 = 0$.

Such a chi-square method is usually applied, for example, for one MC generator: a chosen variable of the MC generator, for example the top quark mass, is varied and after the simulation different distributions of an observable are obtained.

Nevertheless, the total event selection efficiencies (reconstruction level) taken from table 7.2 was chosen as an observable ($N = 1$), the chi-square was calculated for the four MC generators ("overall" chi-square) and Sherpa served as pseudo real data like above. As a second try, the distributions of the 10 selection observables of Section 7.2 (η of electron, muon, jet; $\not\!\!E_{\mathrm{T}}$; p_T of electron, muon, 1st, 2nd, 3rd, 4th leading jet) were chosen (N_j = number of bins for each distribution j). The sum over the computed chi-squares was calculated, which, however,

neglects correlations between the 10 observables. The chi-squares of both tries are illustrated for all MC generators in figure 7.10. Parabolas ($a \cdot x^2 + b \cdot x + c$)

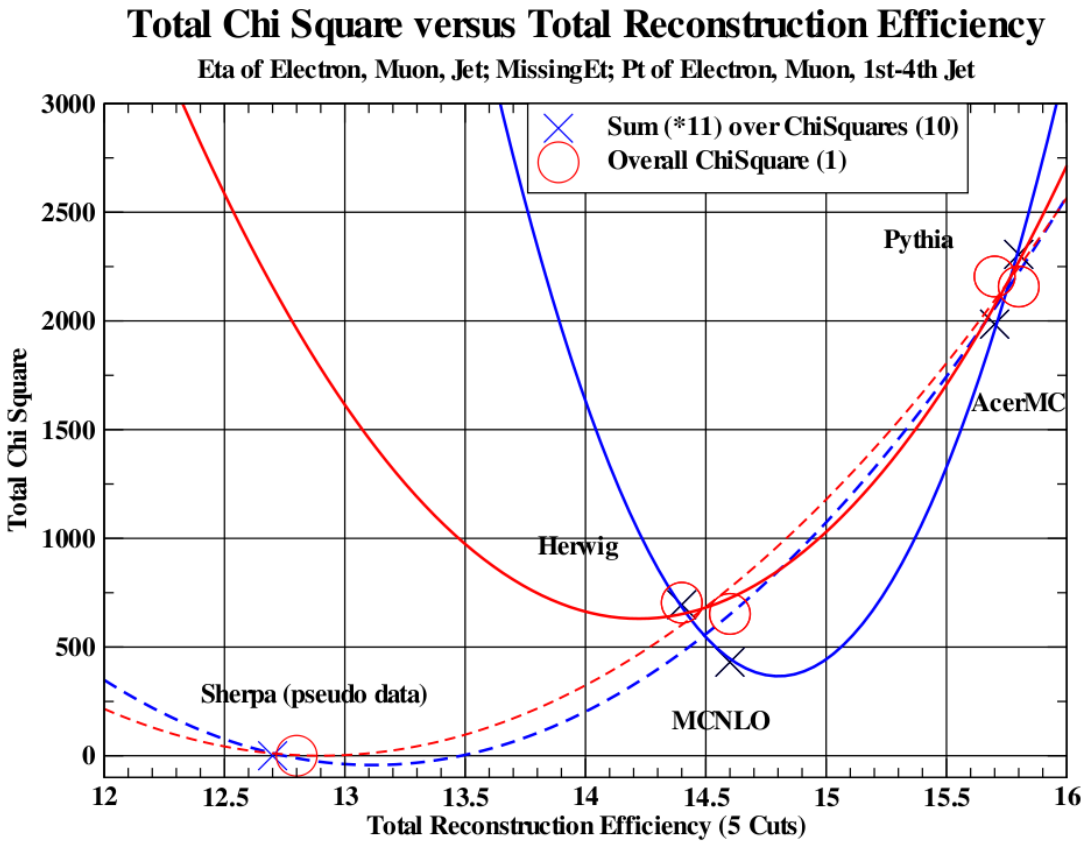

Figure 7.10: Chi-square in dependence of the total selection efficiency for five MC generators. The red symbols correspond to the overall chi-square, the blue symbols correspond to the sum over 11 chi-squares. In order to have a comparability between both, the overall chi-square is multiplied by 11. Solid lines indicate a parabolic fit to 4 MC generators, dashed lines represent such a fit to 5 MC generators. For a visibility of both Sherpa entries the red one was displaced by +0.1.

are fitted to both chi-squares. However, these fits have few significance, because the entries of MCNLO and Herwig and of Pythia and AcerMC are almost on top of each other. In order to obtain significant results, more MC generators are needed.

Furthermore, the pseudo real data entry lies outside the entries of the MC generators. Therefore, the estimate of the systematic uncertainty with corresponds at least to the distance to the next entry, which results in $\approx$ 13% ($= 100 \cdot (14.4 - 12.7)/12.7$). This value is almost by a factor of 4 larger than the value obtained by the reweighting approach in the previous section.

Eventually, first data is between the entries of the MC generators, i.e. in the minimum of the dashed parabolas (about 13.0). Only then, the systematic uncertainty would be at the 1% level, and the chi-square method would be preferred to the reweighting approach.

Chapter 8

Summary and Outlook

The LHC and its experiments will open new possibilities to study precisely the properties of known particles such as the top quark and to find particles never observed before.

Careful operation, maintenance and tuning of the detectors and their readout systems are mandatory, in order to obtain high statistics data samples and perform precision measurements. The control and the readout of the ATLAS pixel detector is performed through an optical transmission line. The installation and commissioning of the optical cables as well as the steps from the production to the operation of the Back of Crate cards, which are the optical interface on the off-detector side, are presented in this thesis. Tests of the optical cables have been developed and realized. These tests ensure a complete functionality of all optical cables and prove its operation within given limits. Similarly, the full operability of the Back of Crate cards is assured by a precise production, extensive tests and a careful installation. Since spring 2008, the readout of the pixel detector is ready for operation. Data was taken successfully during cosmics runs, and the tuning of detector and readout is ongoing.

Track impact parameters, which are mainly determined by the pixel detector, are important observables for b-tagging and therefore for many physics analyses such as top quark studies. Systematic uncertainties are reduced, if distributions of MC simulation and data agree. A method to map MC distributions of track impact parameters on data was developed, implemented and tested by using misaligned MC samples as data. The performance of the approach is promising for the correction of MC samples from first data. Especially for impact parameters $|IP| < 0.4\,\text{mm}$ the agreement between MC and pseudo data was improved by up to a factor 4. The work to prepare the corresponding software package for first data is ongoing as well as tests with current MC data samples.

Precision measurements of observables of top quark pair events will be done by ATLAS using the LHC as a top factory. From the experience gained by the Tevatron experiments CDF and DØ Monte Carlo generators were tuned to predict top quark observables at the LHC. A comparison of five common MC generators was

done for the total selection efficiencies and for many distributions, for example the transverse momentum and pseudorapidity of the hadronically decaying top quark. The deviations between the total efficiencies of the MC generators are on average 18%, the maximum is 23%. The deviations between the MC distributions of the observables serving for the event selection depend on the observable. For the pseudorapidity the deviations are typically small ($\approx 2\%$). However, there are large deviations for the transverse momenta, especially of jets and the hadronically decaying top quark (up to a factor of 4 in certain bins). These deviations represent the systematic uncertainty caused by the physics modelling.
Likewise, detector effects on these distributions were studied by using a full detector simulation and event reconstruction. As expected, the distortions of the distributions by detector effects are marginal or linear.
A method was developed, implemented and tested to reweight the MC distributions from one observable, for example the transverse momentum of the hadronically decaying top quark, such that they agree better with pseudo data. Thereby, an improvement of the average systematic uncertainty of the total efficiencies by a factor of > 5 to $\lesssim 3\%$ was reached. This was confirmed by using a different observable for reweighting. Such a performance may be also obtained by applying this method on first data. Therefore, the work to prepare the corresponding software package is ongoing.

Appendix A

Implementation and Plots of the Track Impact Parameter Tuning

A.1 Implementation of the Track Impact Parameter Mapping Approach

A new ATLAS offline software package ImpactParameterTuning was created to adjust the track impact parameters of Monte Carlo simulation Analysis Object Data (AOD) files. The package is implemented as a C++ library mostly according to the coding rules [134–137]. It can be downloaded at [138]. The necessary algorithms can be run inside the Athena framework [66] by using a python job option file.

The package is split into parts:

- "TrackSelector" is an Athena algorithm, which runs over AODs (Monte Carlo or Real Data — so, a comparison of simulated and real data can be easily done) and stores event/jet/track information according to chosen selection criteria (for example optimized for $t\bar{t}$ events) in a ROOT [139] file.

- "TrackIpPlotFitter" is an Athena algorithm, which runs over the .root file generated by "TrackSelector" and creates distribution histograms of signed impact parameters and significances. These histograms may be fitted by a provided function.

- "TrackAdjustor" is an Athena algorithm, which runs over MC AODs, copies their content into new AODs and adds a new `Rec::TrackParticleContainer`. This contains the copied tracks whose track impact parameters and error matrix were adjusted.

- "TrackAdjustorTool" is an Athena tool, which is called by "TrackAdjustor". It creates a new `Rec::TrackParticle` (with the adjusted track impact pa-

rameters) by using the histograms of "TrackIpPlotFitter", which are provided by two .root files (one for the MC distributions, one for the real data distributions) in the current directory. The "TrackAdjustorTool" uses two further subtools: "TrackParametersAdjust", which adjusts the track impact parameters, and "TrackErrorMatrixAdjust", which adjusts the track error matrix.

Figure A.1 illustrates the program and data flow. More detailed documentation of these parts can be found in their doc directories.

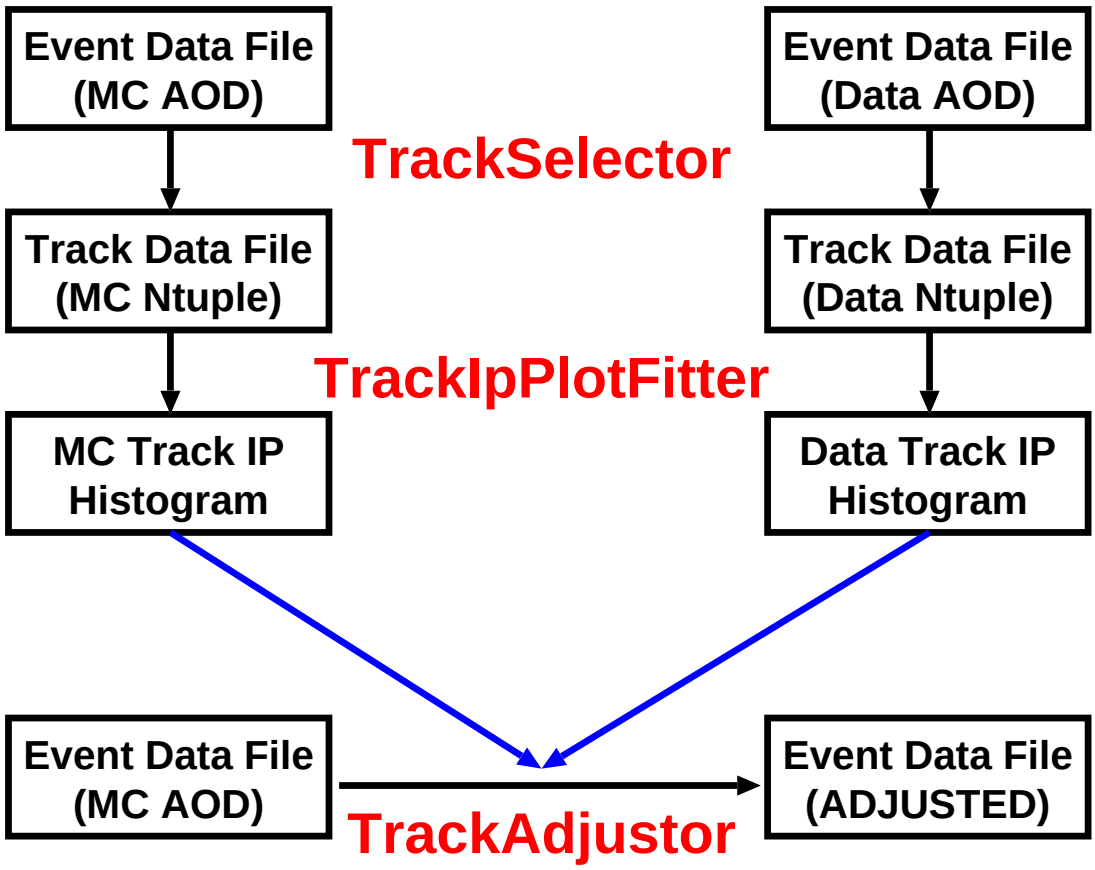

Figure A.1: Program and data flow of the track impact parameter tuning package.

A.2 Plots of the Track Impact Parameter Mapping Approach

The distributions of the signed transverse and longitudinal impact parameter are shown for MC (perfect detector geometry), pseudo data (misaligned MC) and for MC after tuning. There are 4 ranges in the transverse momentum of the track: $1-2\,\mathrm{GeV}$, $2-3\,\mathrm{GeV}$, $3-4\,\mathrm{GeV}$, $>4\,\mathrm{GeV}$. For the pseudorapidity of the track 5 ranges are chosen: $(-2.5;-1.9)$, $(-1.9;-0.1)$, $(-0.1;+0.1)$, $(+0.1;+1.9)$, $(+1.9;+2.5)$. Hence, there are 20 plots for each signed impact parameter (d_0 and z_0).

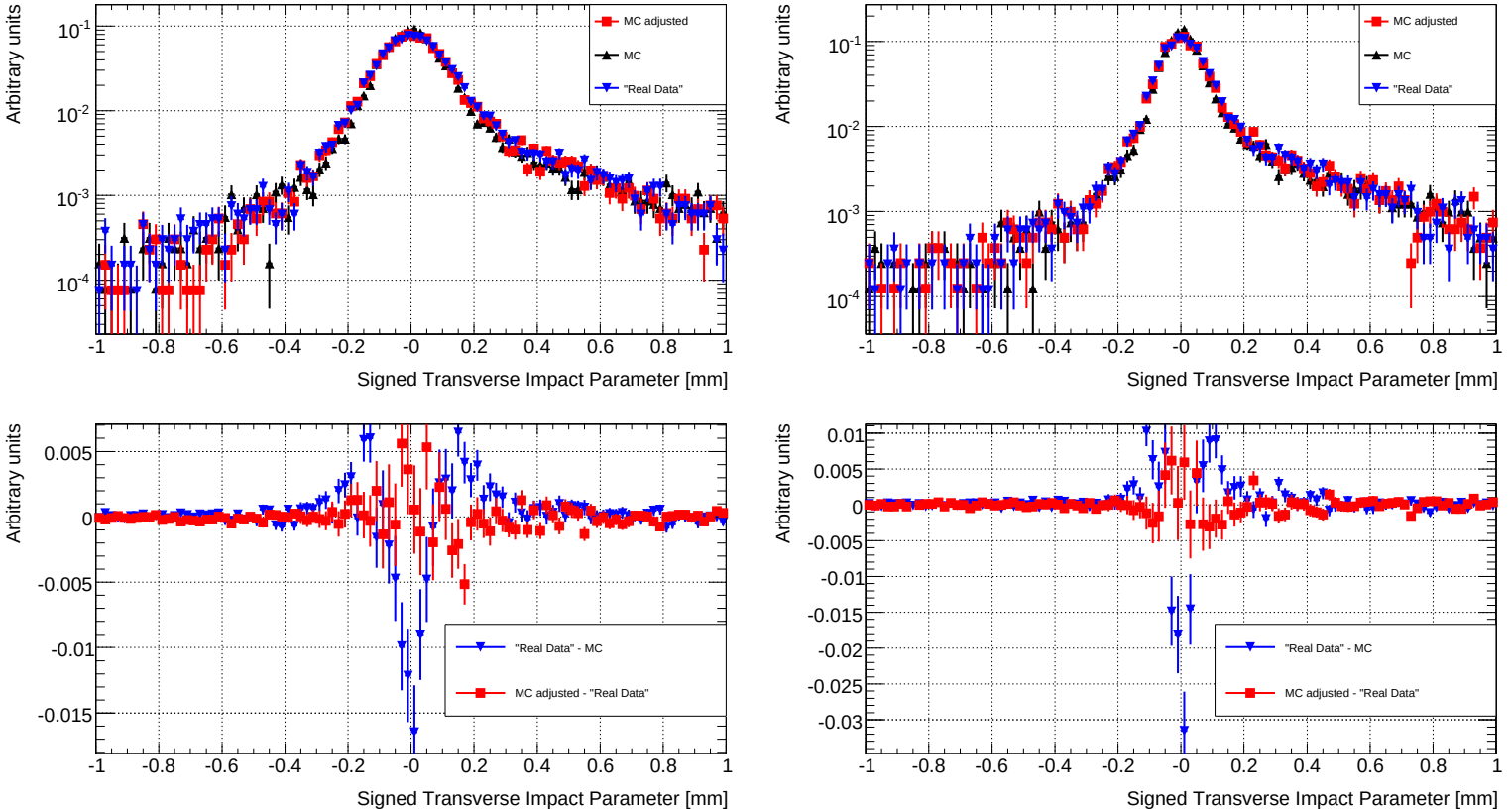

Figure A.2: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $-0.1 < \eta < 0.1$ and $1\,\text{GeV} < p_T < 2\,\text{GeV}$ (at the left) and $2\,\text{GeV} < p_T < 3\,\text{GeV}$ (at the right).

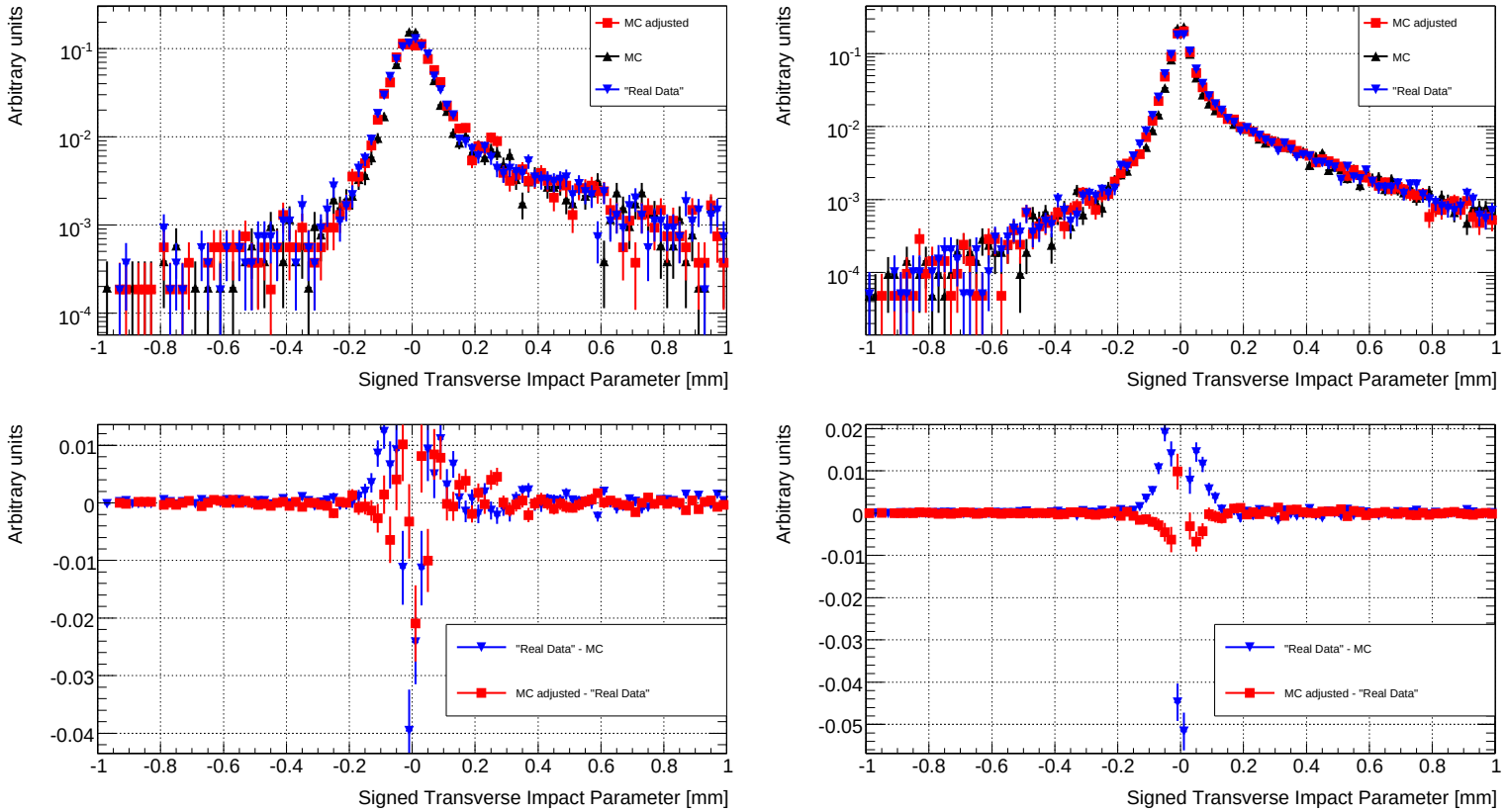

Figure A.3: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $-0.1 < \eta < 0.1$ and $3\,\mathrm{GeV} < p_T < 4\,\mathrm{GeV}$ (at the left) and $4\,\mathrm{GeV} < p_T$ (at the right).

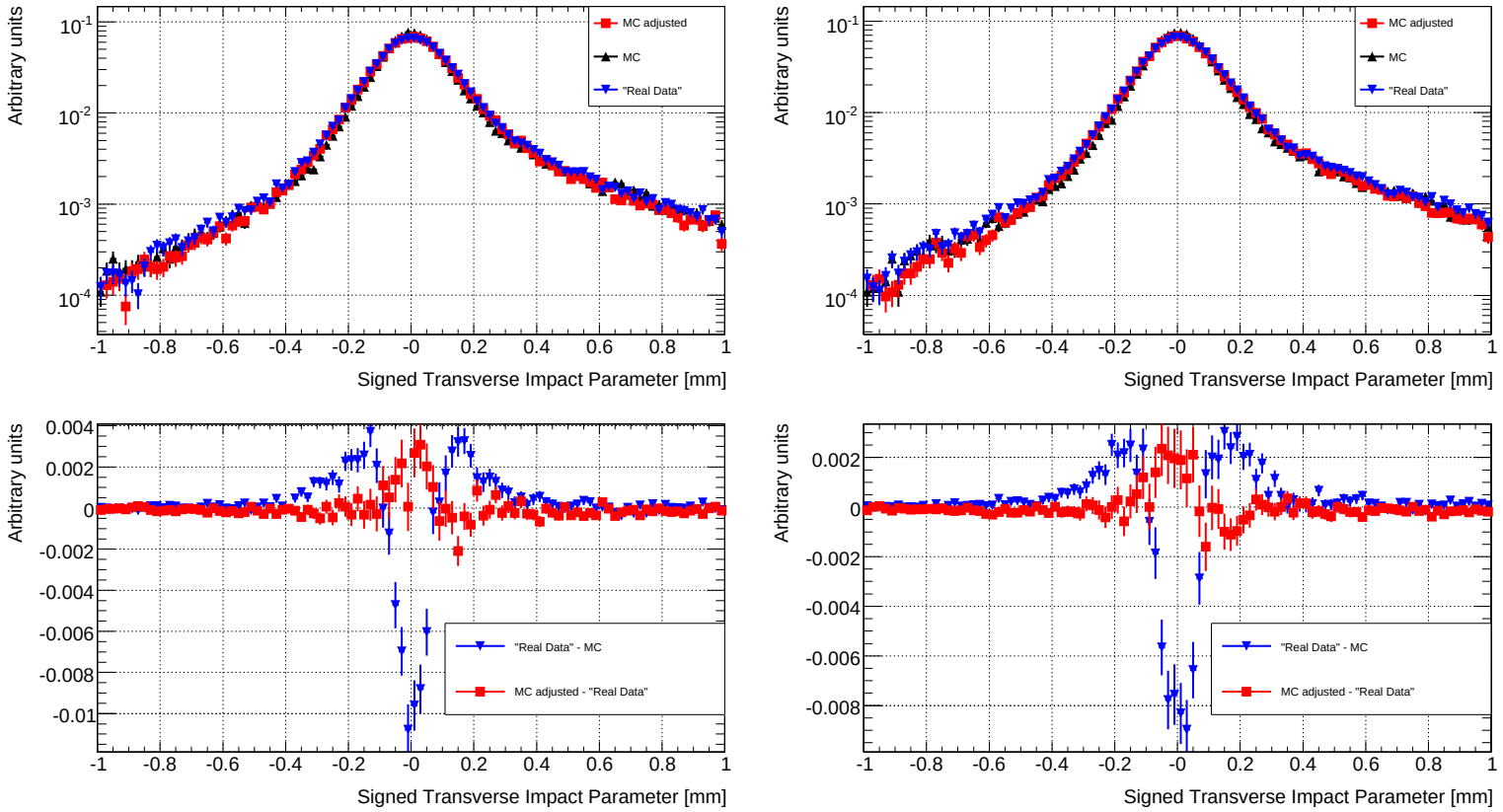

Figure A.4: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $1\,\text{GeV} < p_T < 2\,\text{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

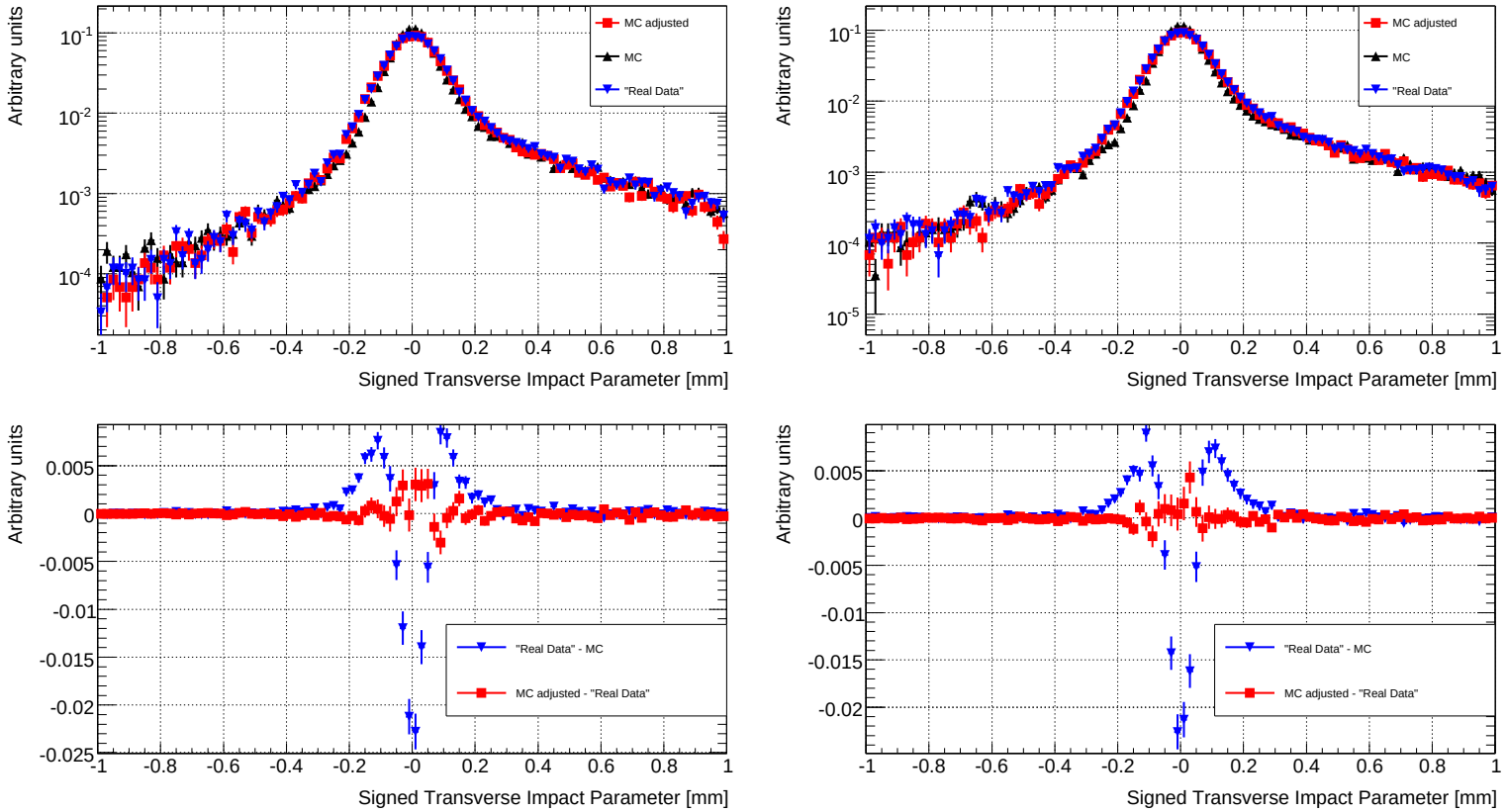

Figure A.5: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $2\,\mathrm{GeV} < p_T < 3\,\mathrm{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

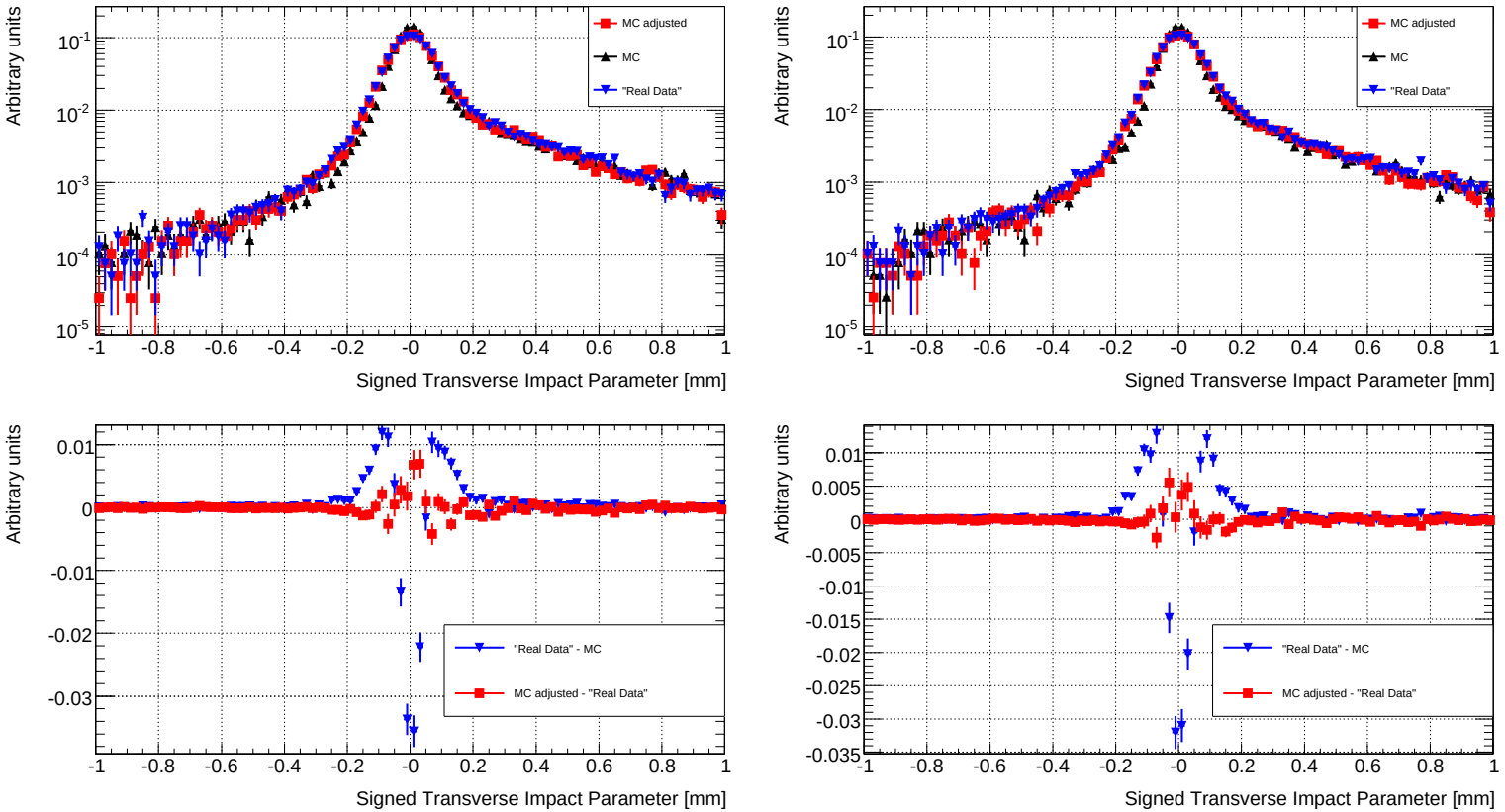

Figure A.6: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $3\,\text{GeV} < p_T < 4\,\text{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

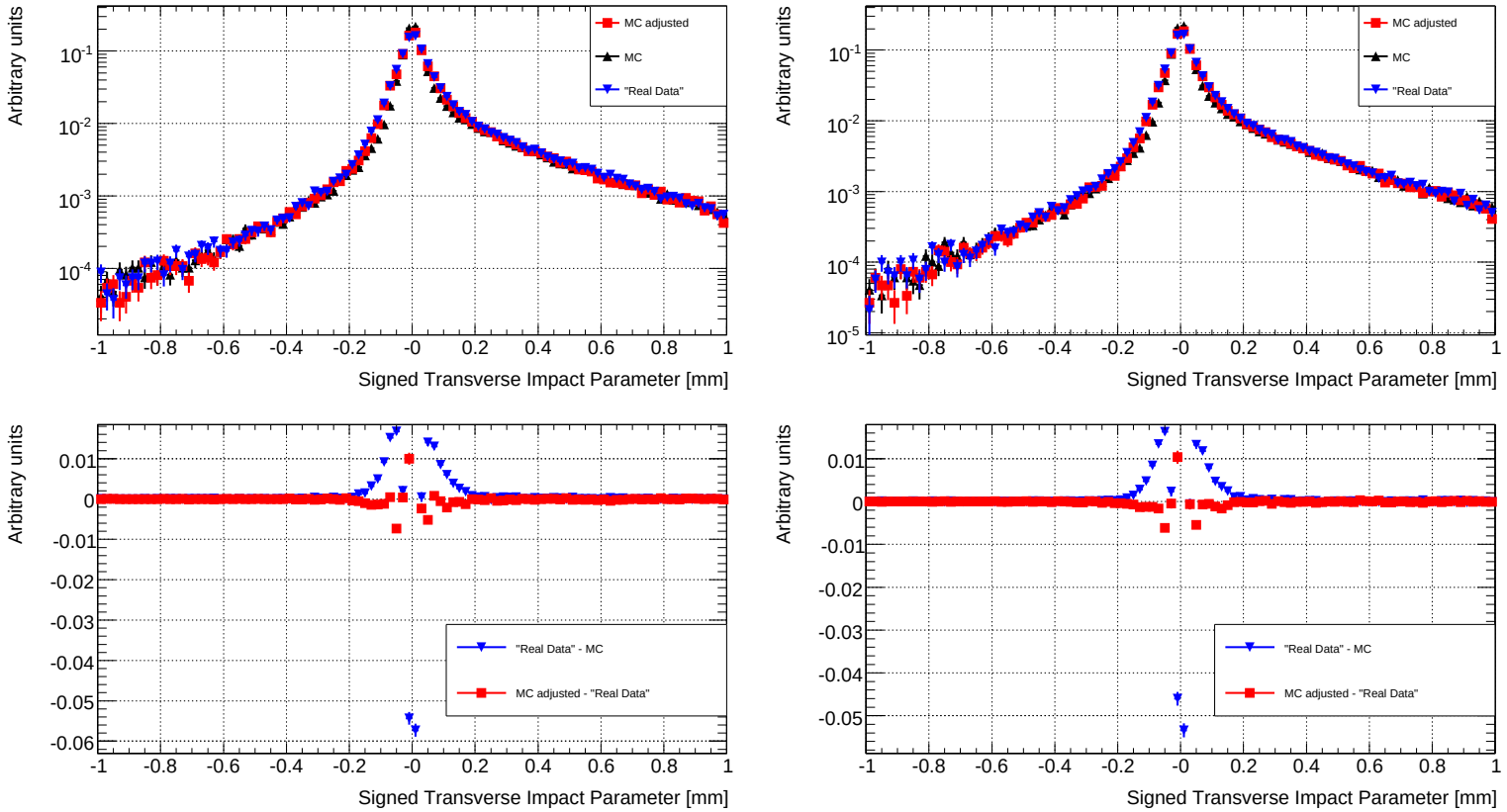

Figure A.7: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $4\,\mathrm{GeV} < p_T$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

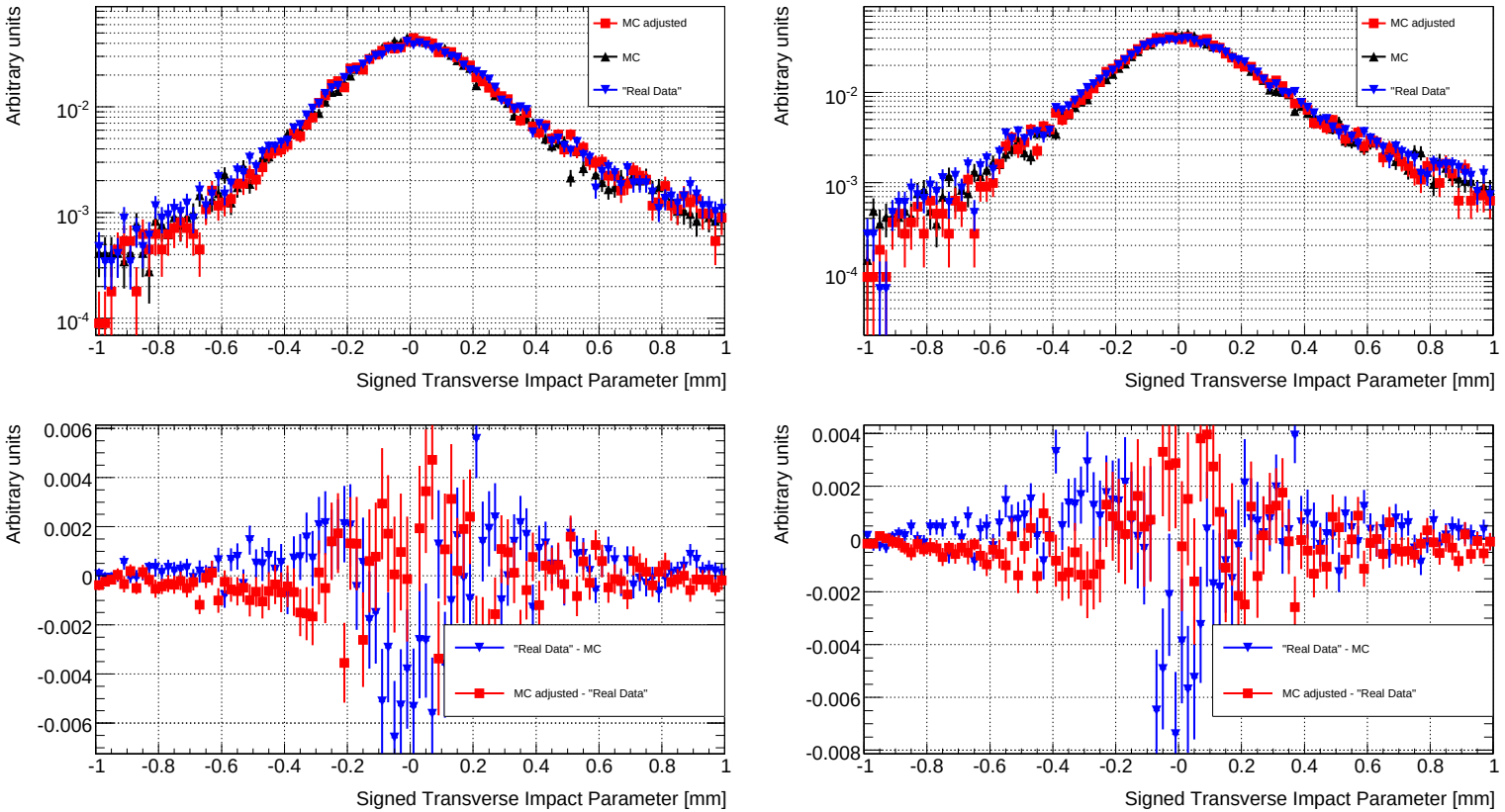

Figure A.8: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $1\,\text{GeV} < p_T < 2\,\text{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

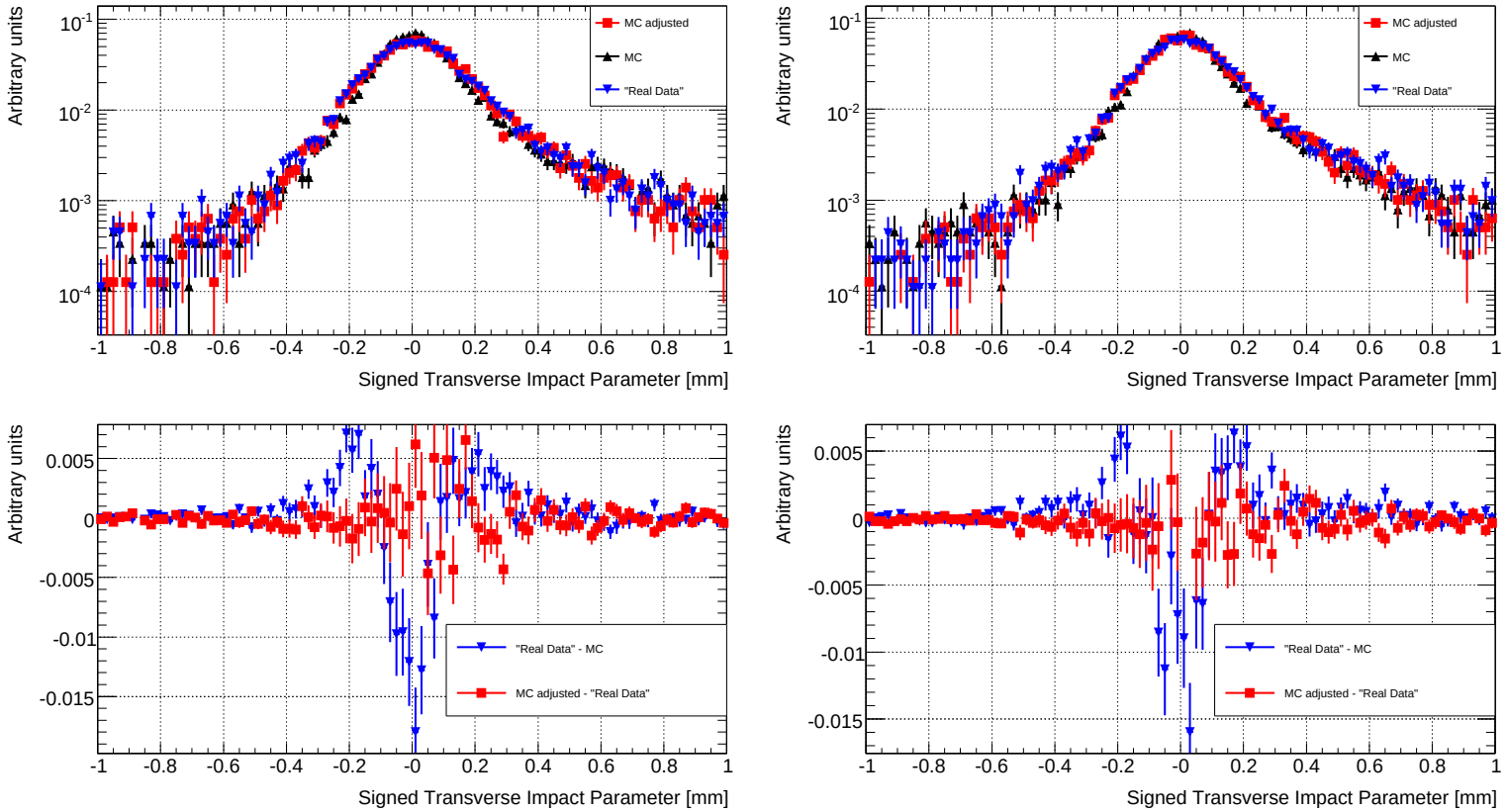

Figure A.9: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $2\,\mathrm{GeV} < p_T < 3\,\mathrm{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

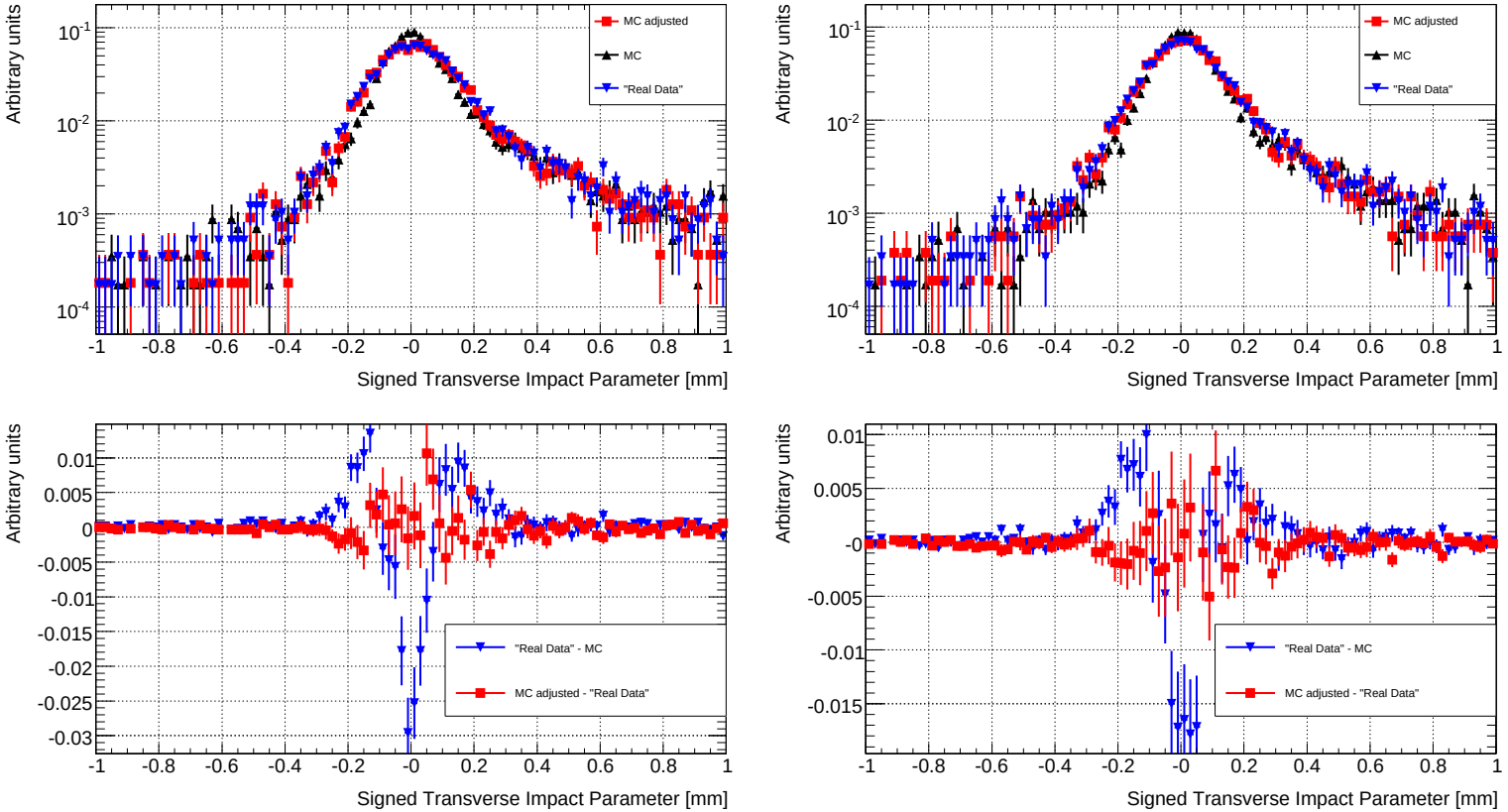

Figure A.10: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $3\,\text{GeV} < p_T < 4\,\text{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

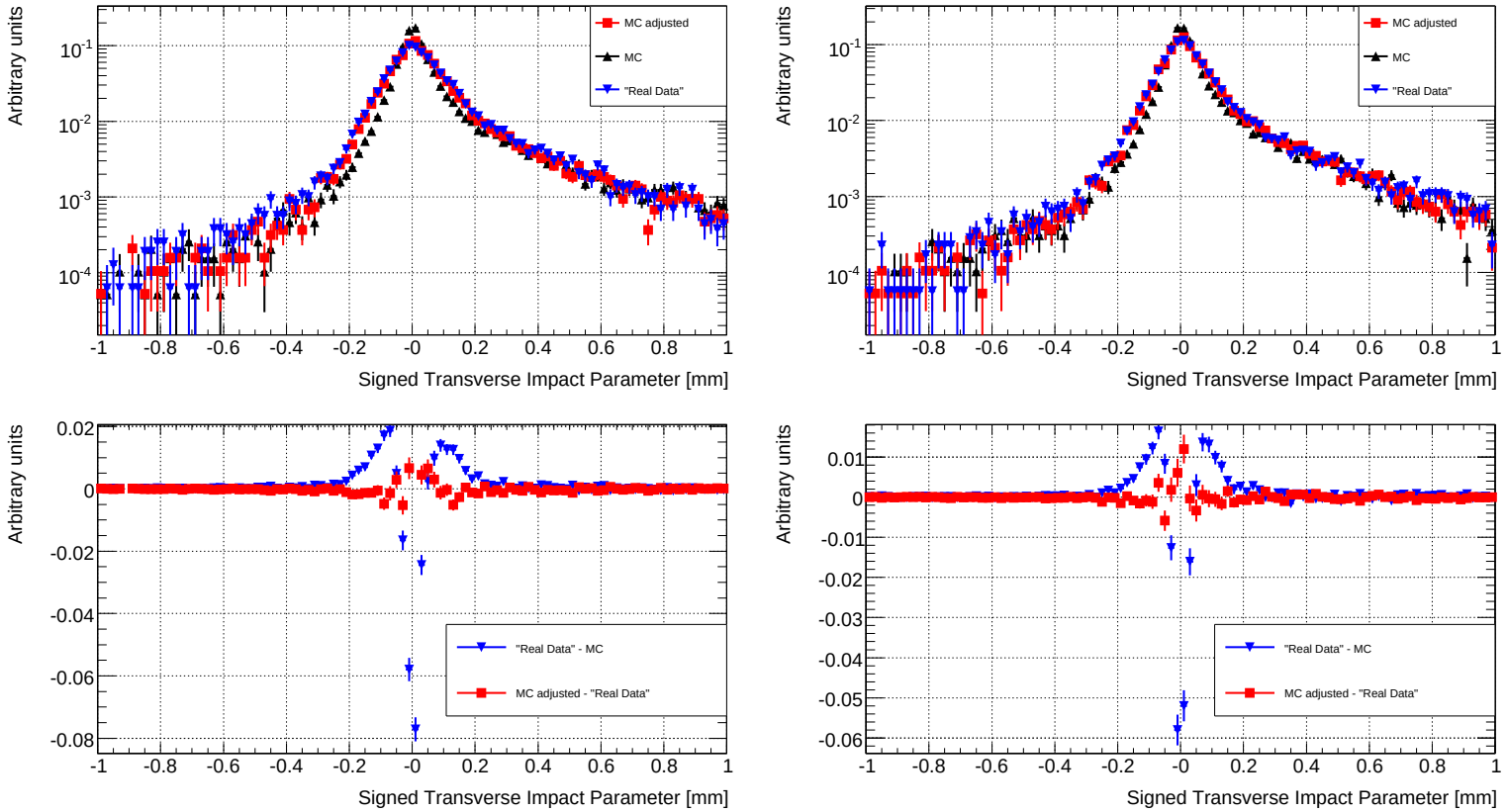

Figure A.11: At the top: a MC distribution (black) of signed transverse track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $4\,\mathrm{GeV} < p_T$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

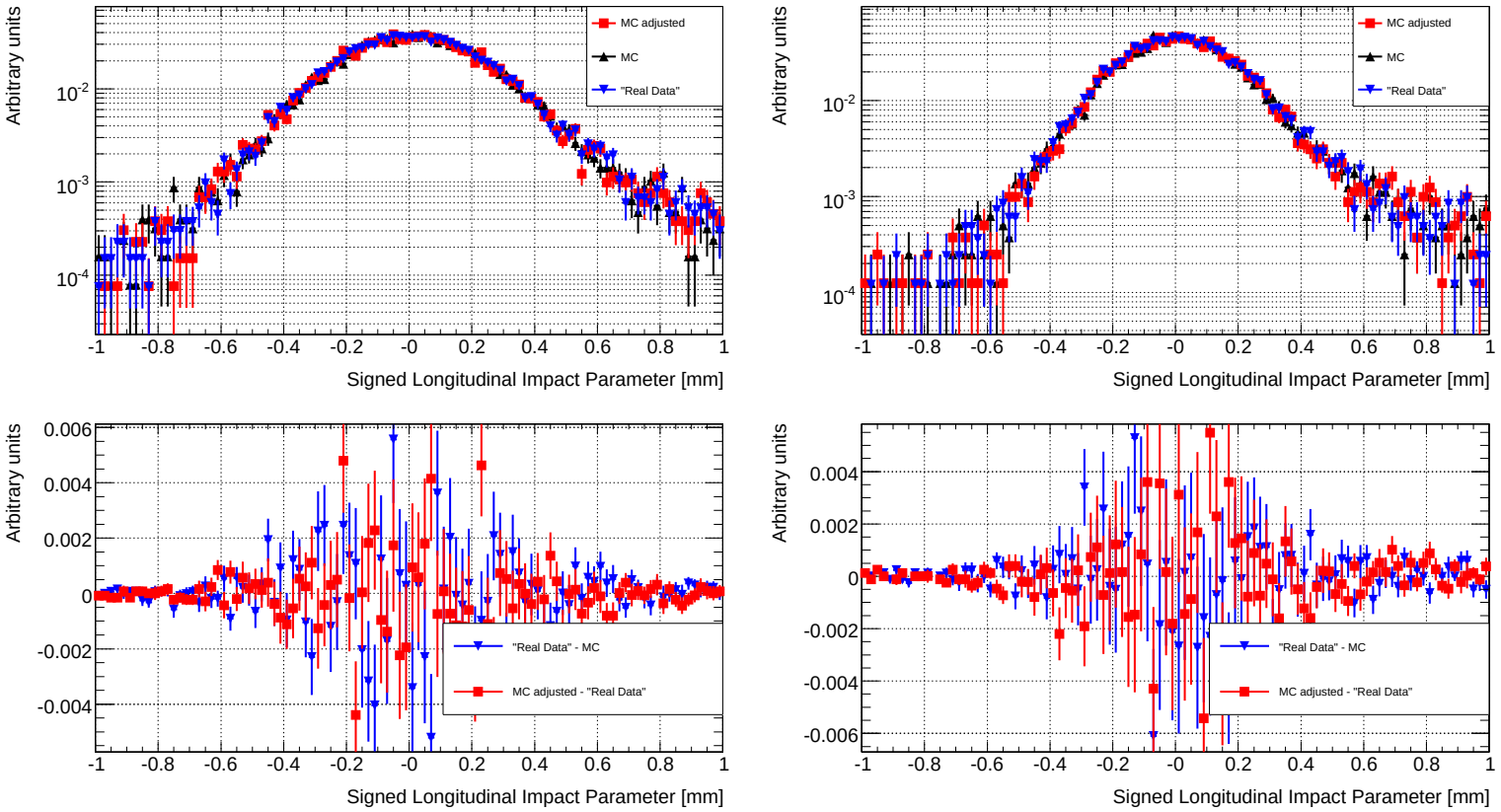

Figure A.12: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $-0.1 < \eta < 0.1$ and $1\,\text{GeV} < p_T < 2\,\text{GeV}$ (at the left) and $2\,\text{GeV} < p_T < 3\,\text{GeV}$ (at the right).

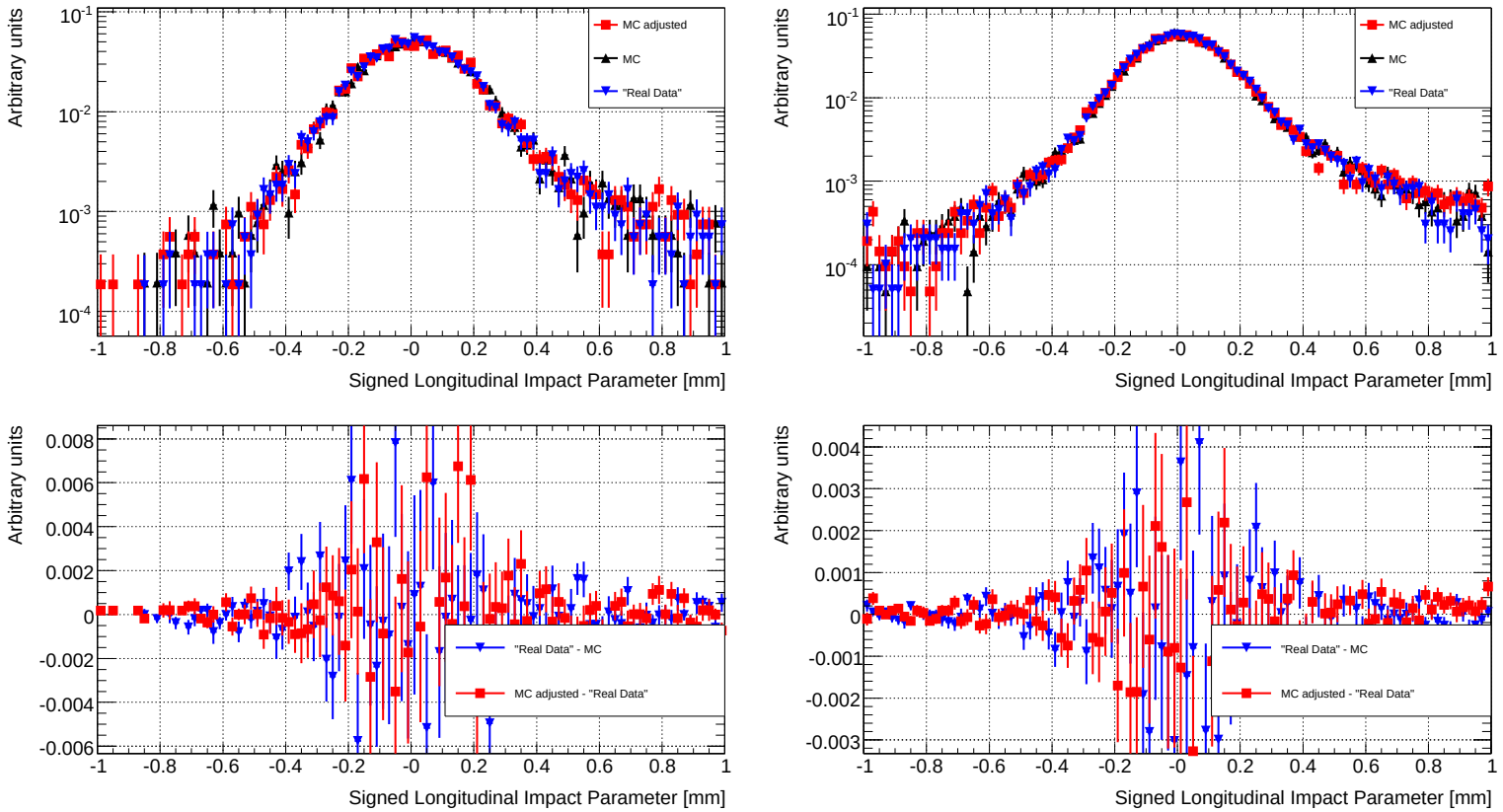

Figure A.13: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $-0.1 < \eta < 0.1$ and $3\,\mathrm{GeV} < p_T < 4\,\mathrm{GeV}$ (at the left) and $4\,\mathrm{GeV} < p_T$ (at the right).

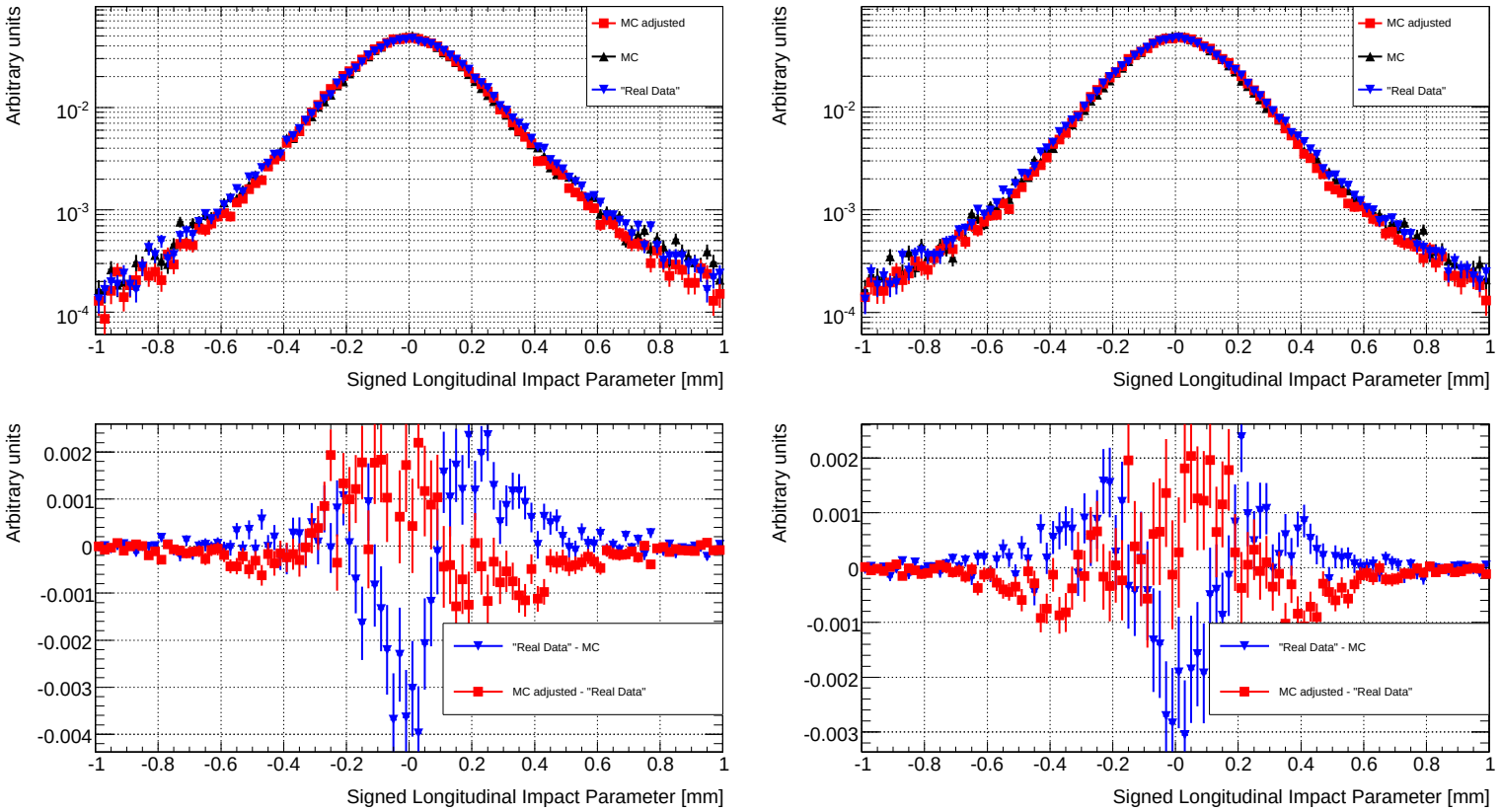

Figure A.14: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $1\,\text{GeV} < p_T < 2\,\text{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

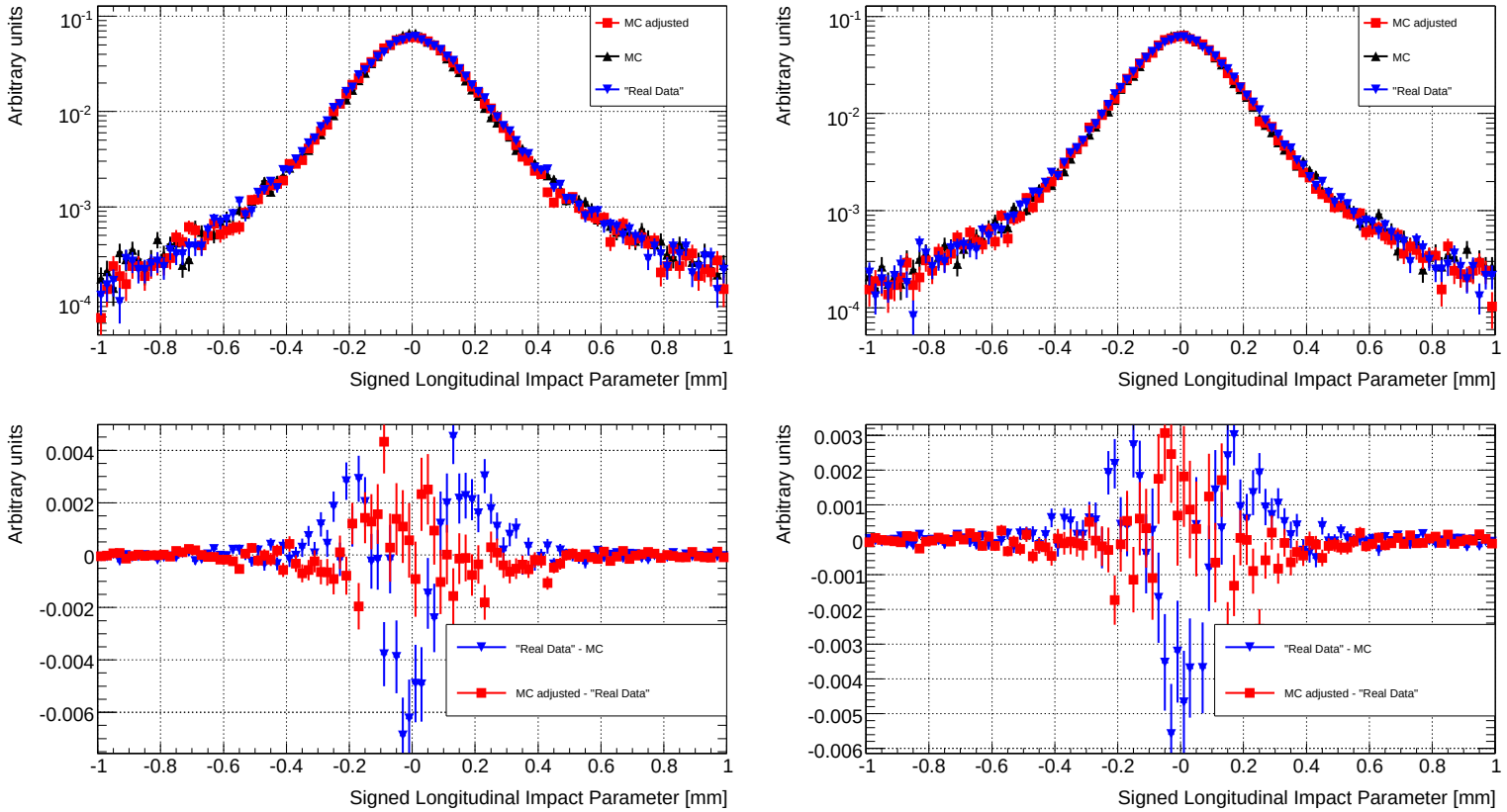

Figure A.15: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $2\,\mathrm{GeV} < p_T < 3\,\mathrm{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

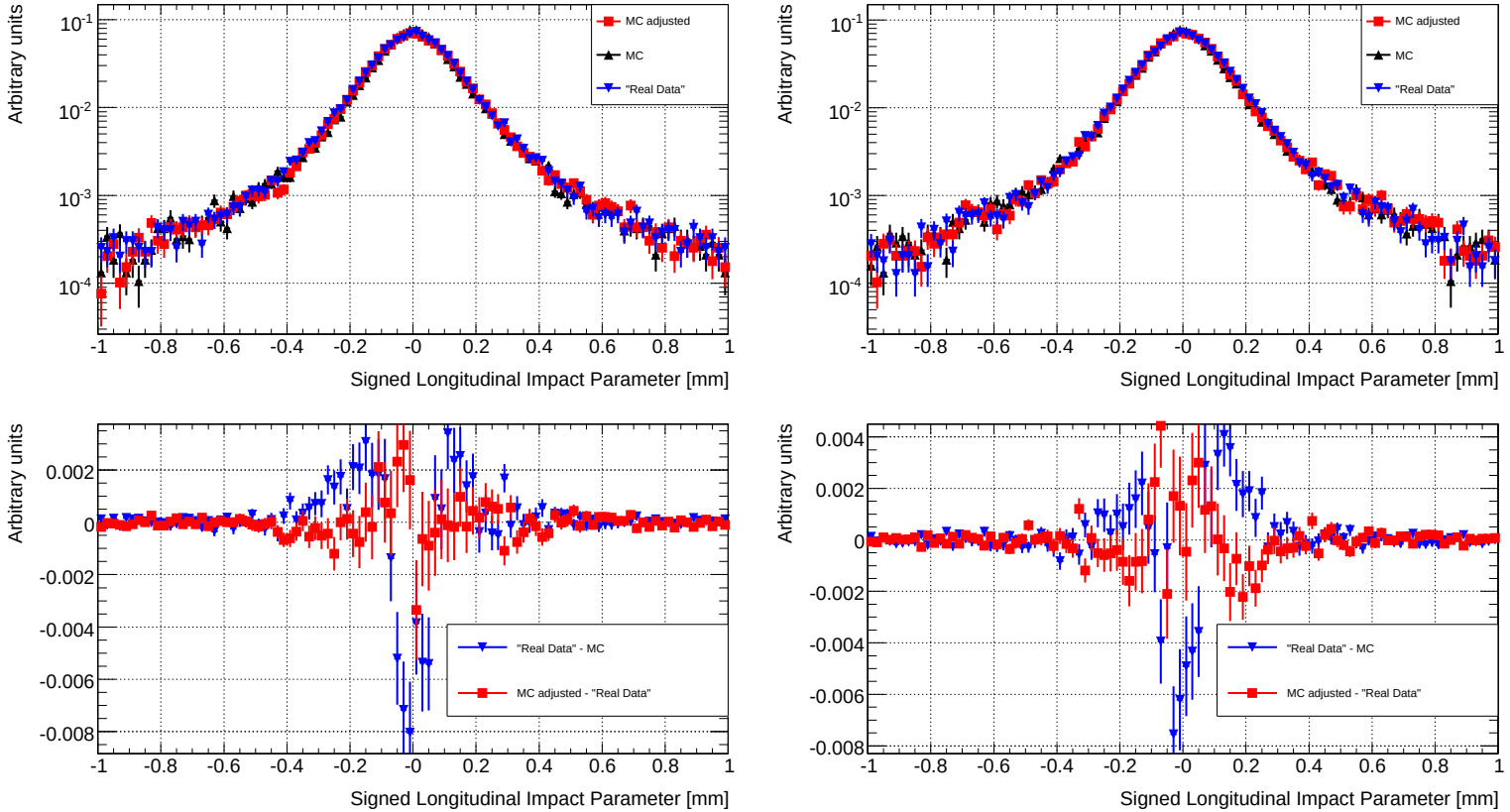

Figure A.16: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $3\,\mathrm{GeV} < p_T < 4\,\mathrm{GeV}$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

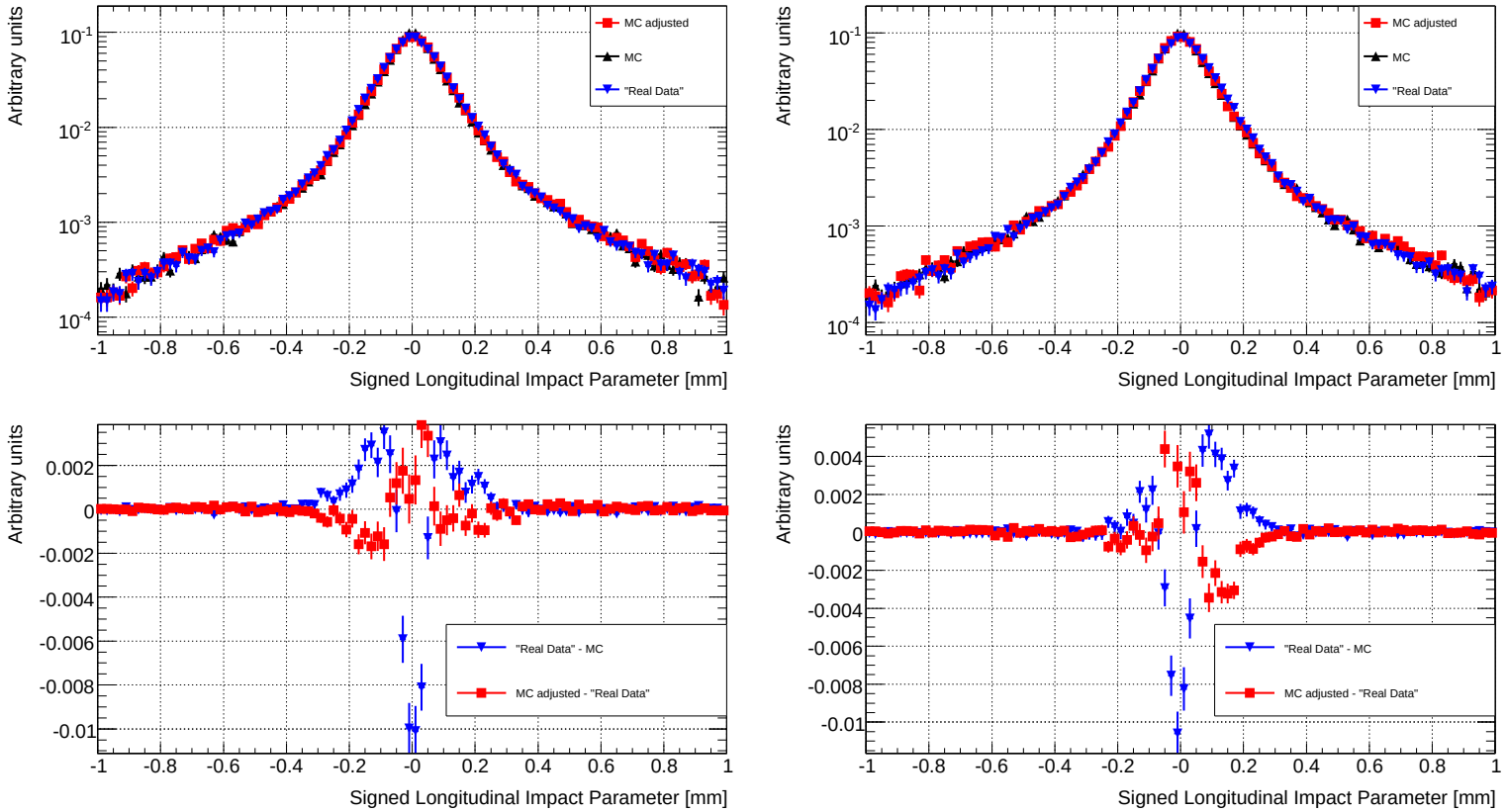

Figure A.17: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $4\,\mathrm{GeV} < p_T$ and $-1.9 < \eta < -0.1$ (at the left) and $0.1 < \eta < 1.9$ (at the right).

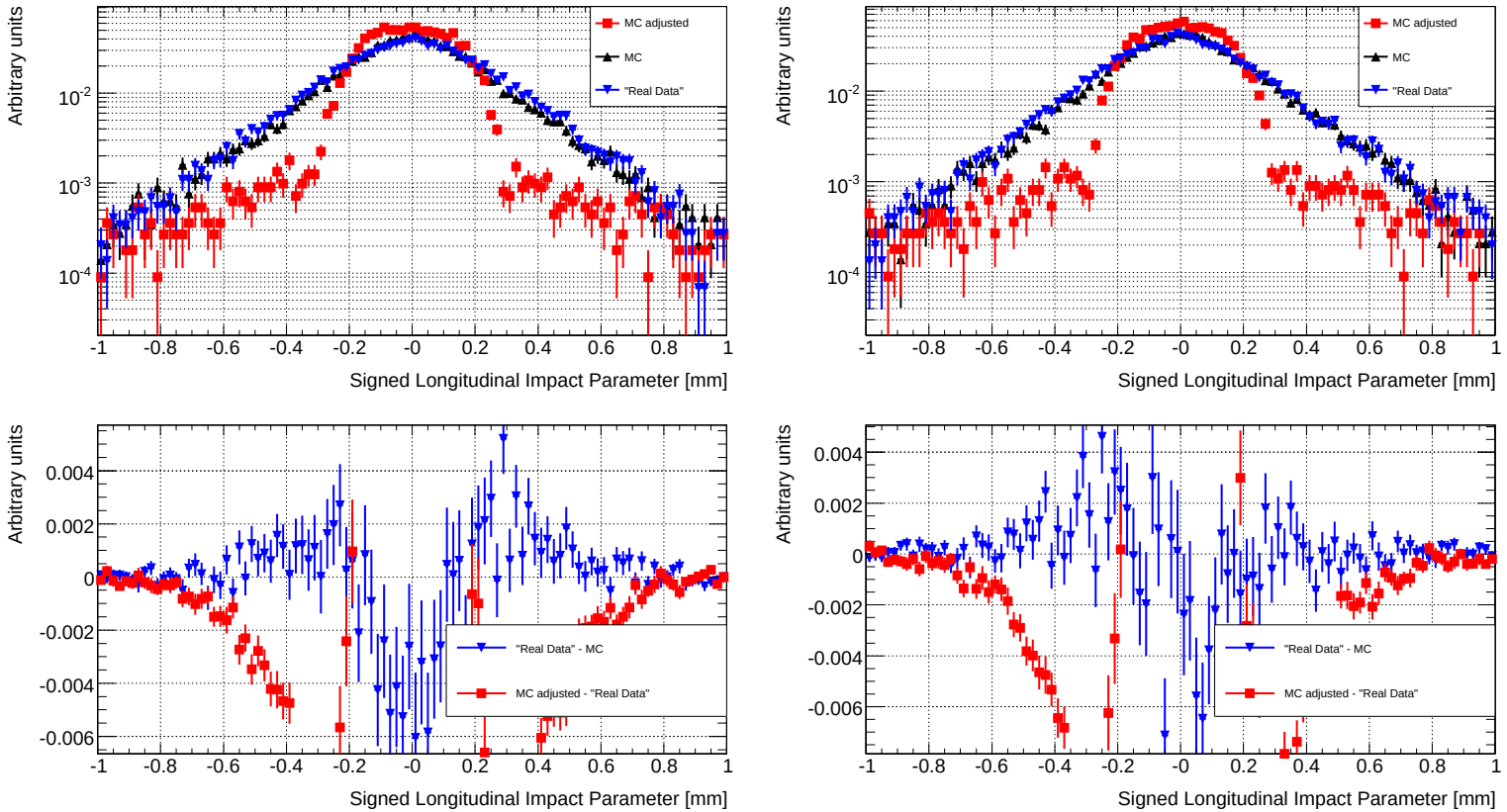

Figure A.18: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $1\,\text{GeV} < p_T < 2\,\text{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

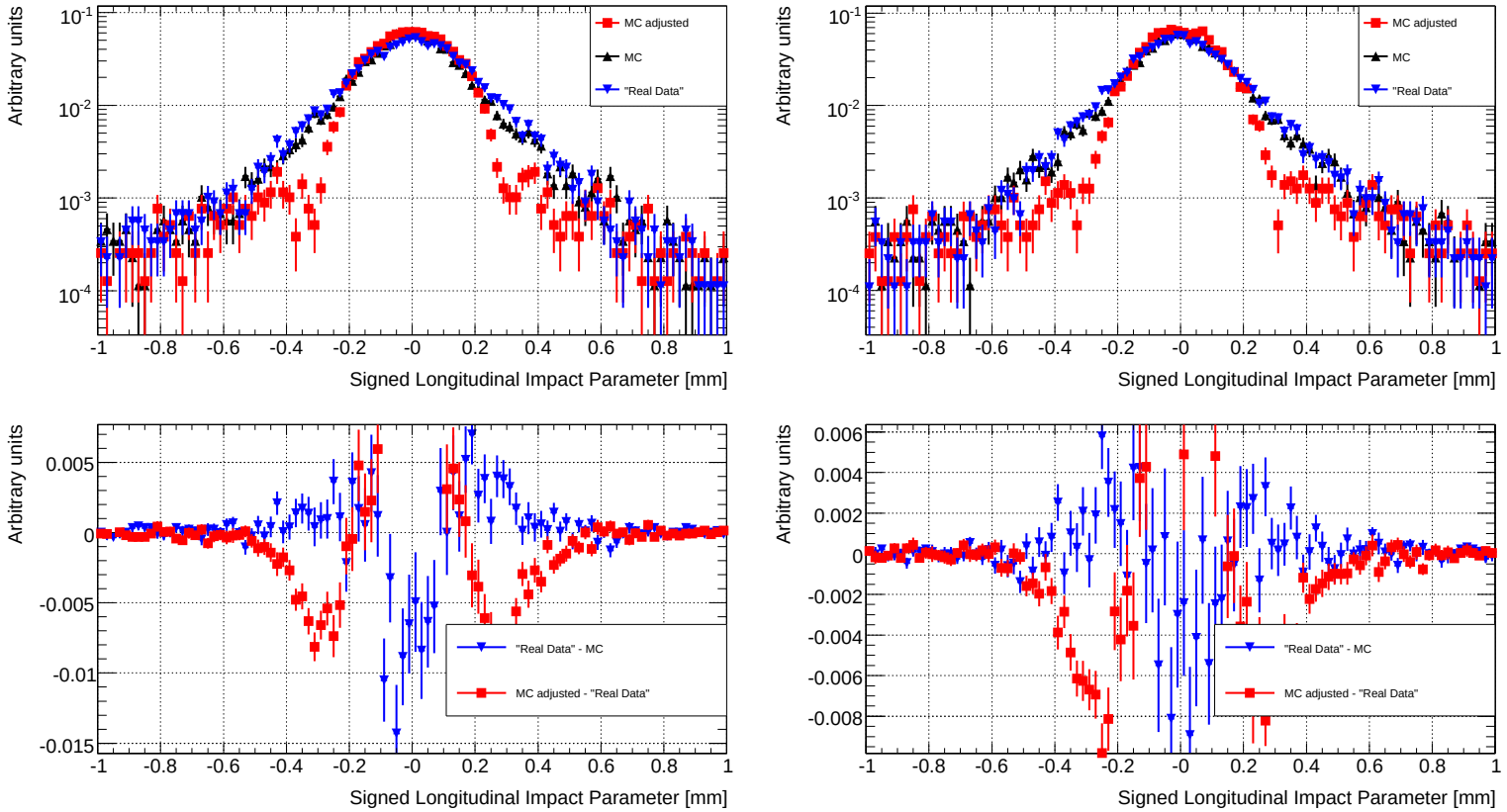

Figure A.19: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $2\,\mathrm{GeV} < p_T < 3\,\mathrm{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

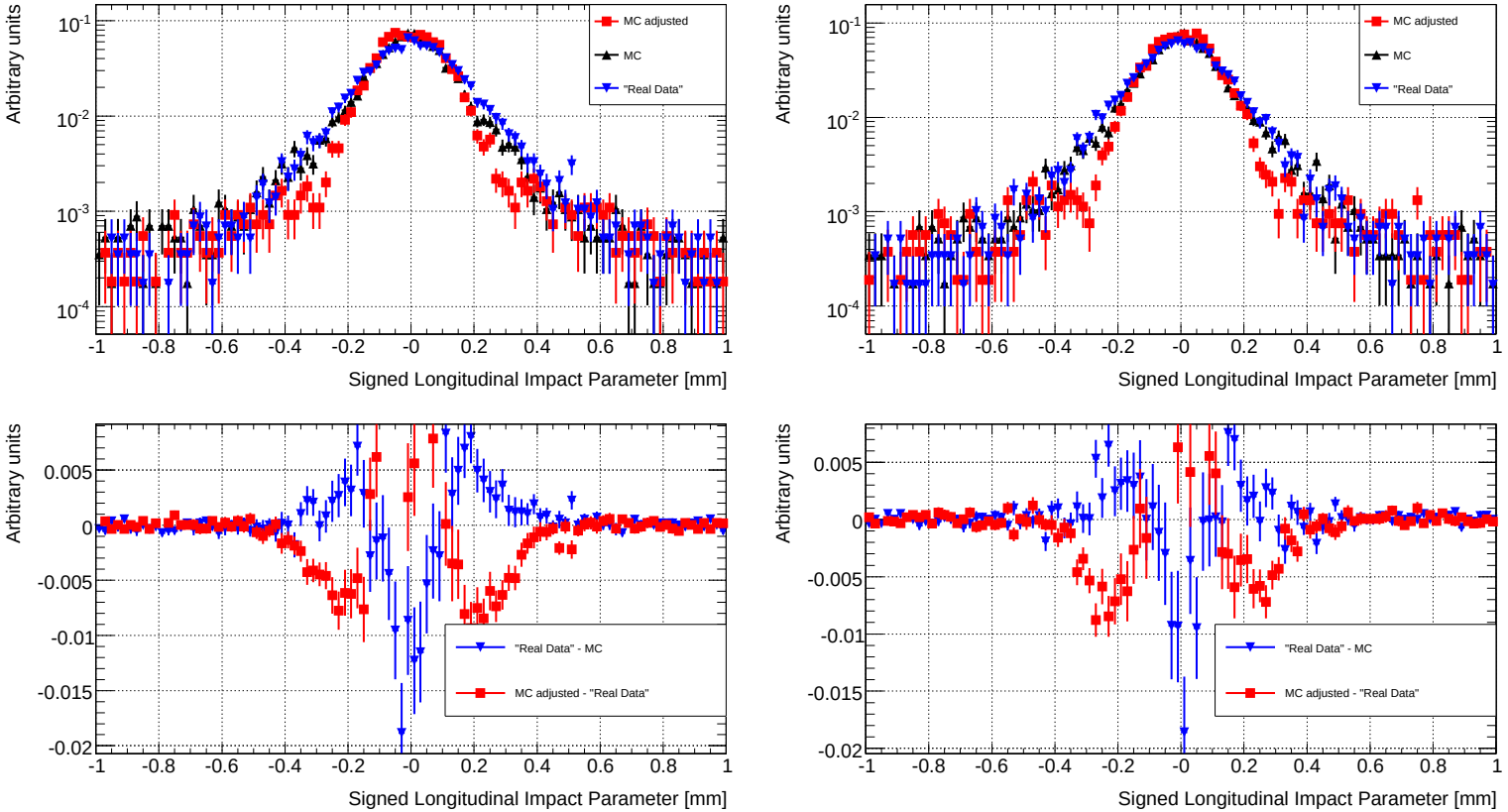

Figure A.20: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $3\,\mathrm{GeV} < p_T < 4\,\mathrm{GeV}$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

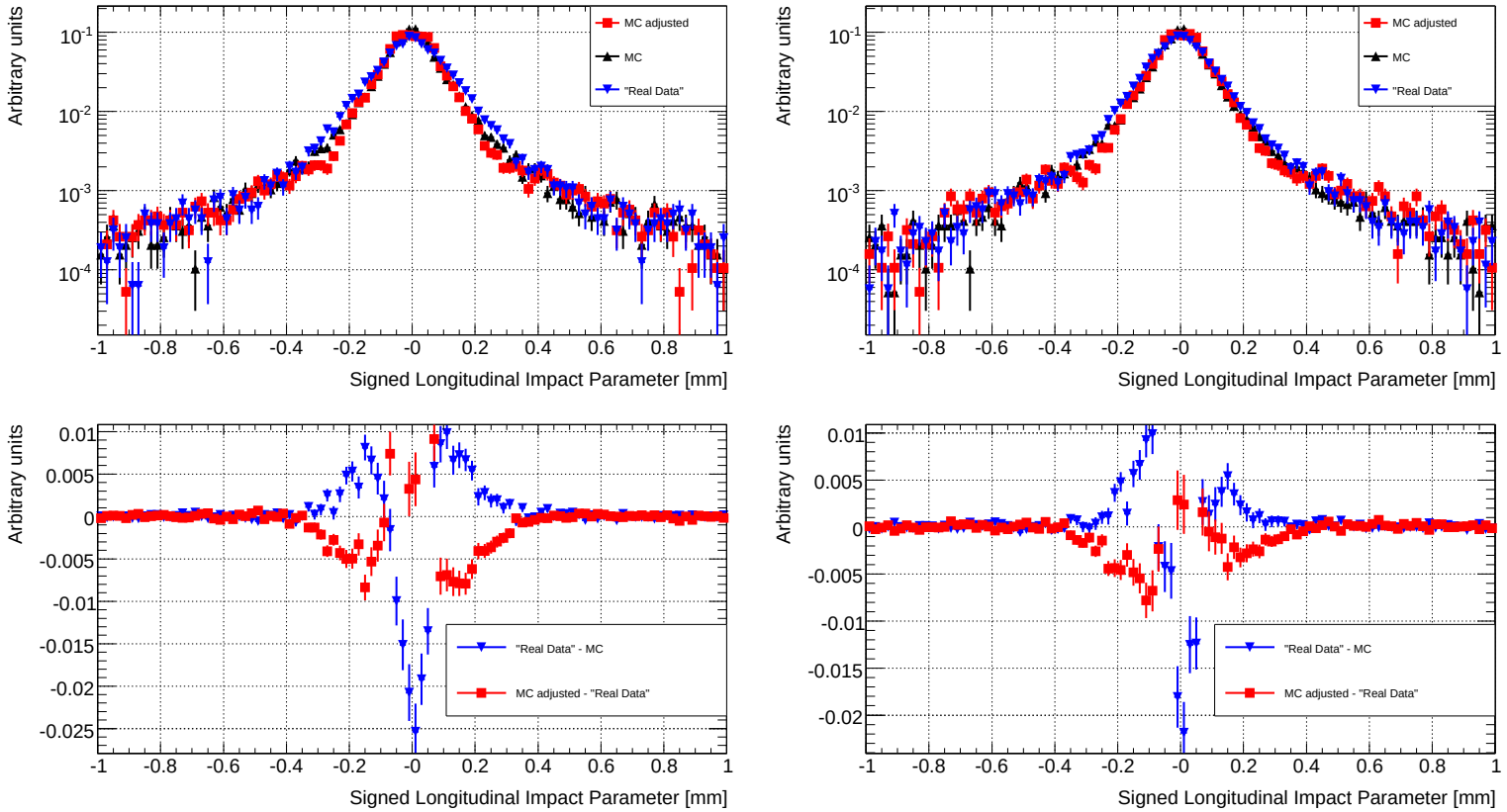

Figure A.21: At the top: a MC distribution (black) of signed longitudinal track impact parameters is mapped (red) to a pseudo real data distribution (blue). At the bottom: the differences of the distributions at the top. The difference between pseudo real data and MC (blue) represents the deviations of the distributions before mapping. The red curve represents the deviations of the distributions after mapping. The mapping is done for tracks with $4\,\mathrm{GeV} < p_T$ and $-2.5 < \eta < -1.9$ (at the left) and $1.9 < \eta < 2.5$ (at the right).

Appendix B

Implementation and Plots of the Reweighting of MC top/antitop Events

B.1 Implementation of the Analysis and of the Reweighting

A new ATLAS offline software package DifferentialCrossSectionTTbar was created to analyze Monte Carlo simulation Analysis Object Data (AOD) files. The package is implemented as a C++ library mostly according to the coding rules [134–137]. It can be downloaded at [140]. The necessary algorithms can be run inside the Athena framework [66] by using a python job option file.

The package is split into parts:

- "EventSelectorTTbar" is an Athena algorithm, which runs over AODs (Monte Carlo or Real Data — so, a comparison of simulated and real data can be easily done) and stores histograms of all relevant observables stepwise according to chosen selection criteria in a ROOT [139] file. The functions for reweighting MC events are provided by a python job option file.

- "MCPlotter" is a standalone program, which runs over the .root files created by the algorithm above, normalizes the histograms and saves them in a .root, .pdf, .ps or .eps file. Fits of distributions for reweighting of MC events are done manually in the first approach.

More detailed documentation of these parts can be found in the doc directory of the package.

B.2 Plots of the Analysis and of the Reweighting

The fits to the ratios between MC and pseudo data (Sherpa arbitrarily chosen) for the transverse momentum of the hadronically decaying top quark (as defined by formula 7.6) are provided for four common MC generators. Furthermore, these ratios are plotted before and after reweighting of the MC events by using those fit functions. The plots are given with detector simulation and cuts or without detector simulation and without any cuts. Finally, the distributions of the observables used for cuts are shown before and after reweighting, both with detector simulation and cuts or without detector simulation and without any cuts: the transverse momentum of electrons, muon, jets (and the leading 4 jets), the missing transverse energy, the pseudorapidity of electrons, muons, jets.

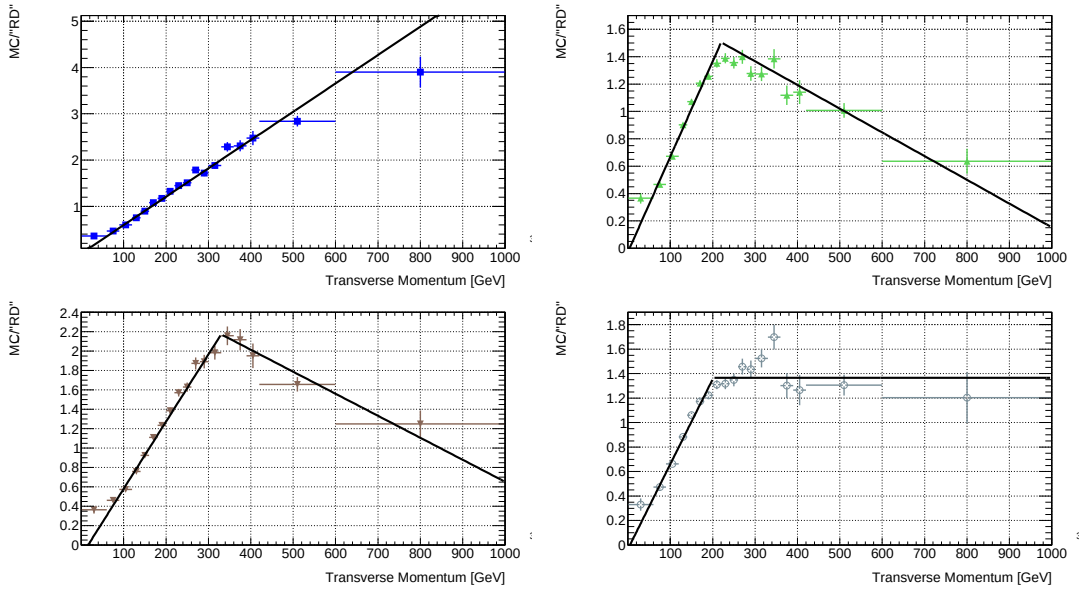

Figure B.1: The ratio between MC and pseudo data (Sherpa arbitrarily chosen) for the transverse momentum of the hadronically decaying top quark in semileptonic $t\bar{t}$ events for four common Monte Carlo generators. Blue: Pythia, green: Herwig, brown: AcerMC, grey: MC@NLO. The obtained smooth fit functions are drawn as well.

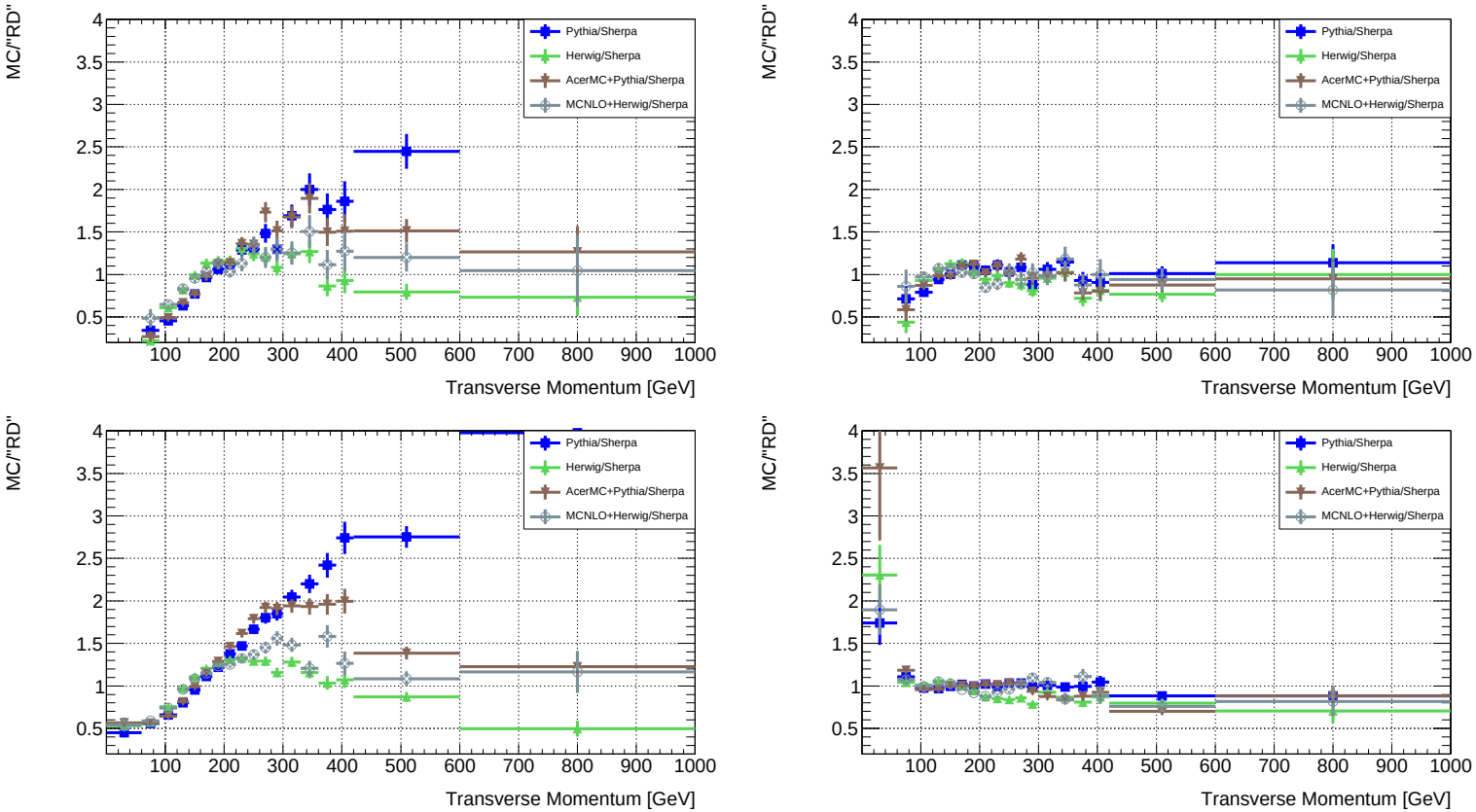

Figure B.2: Transverse momentum of the hadronically decaying top quark in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

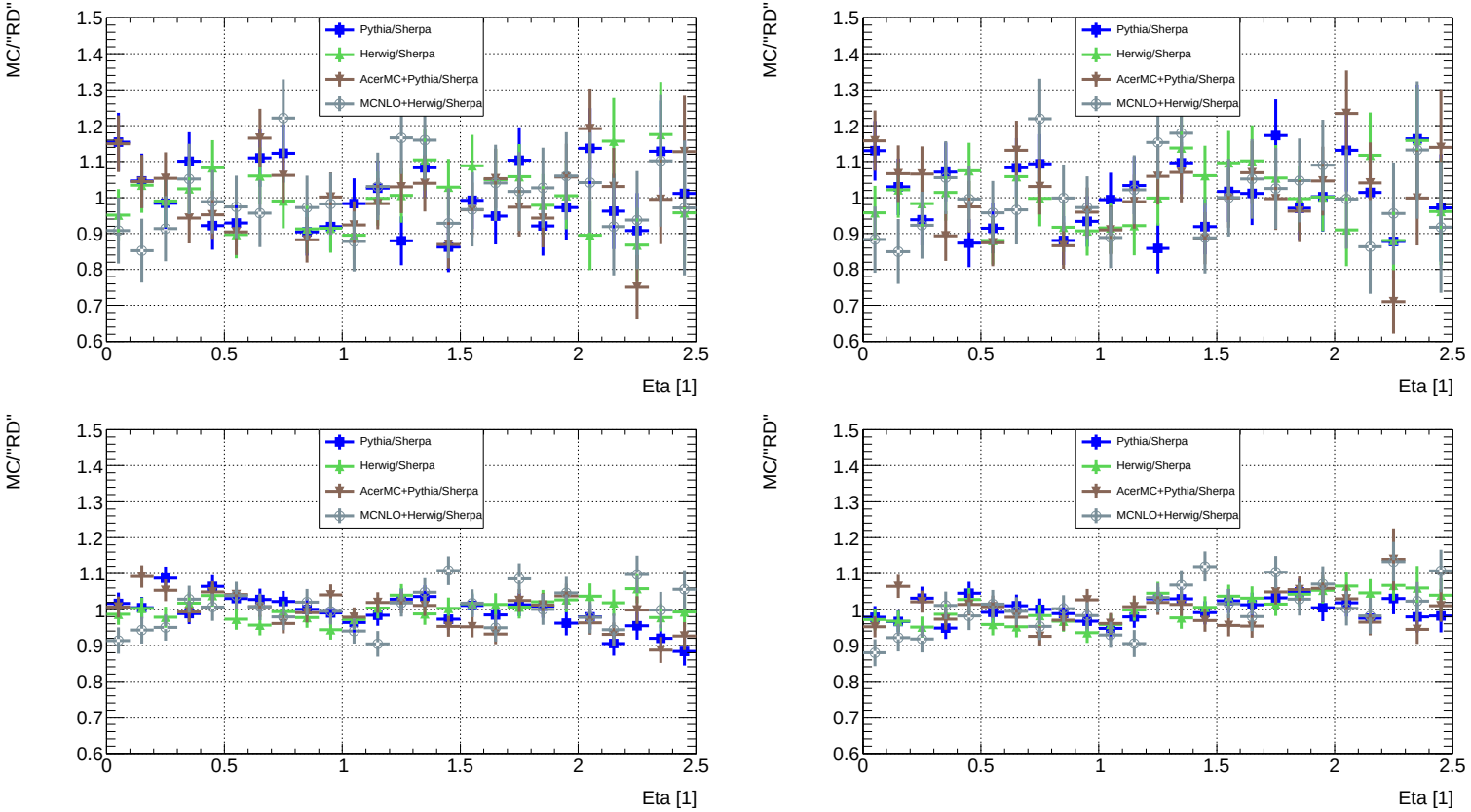

Figure B.3: Pseudorapidity of the hadronically decaying top quark in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

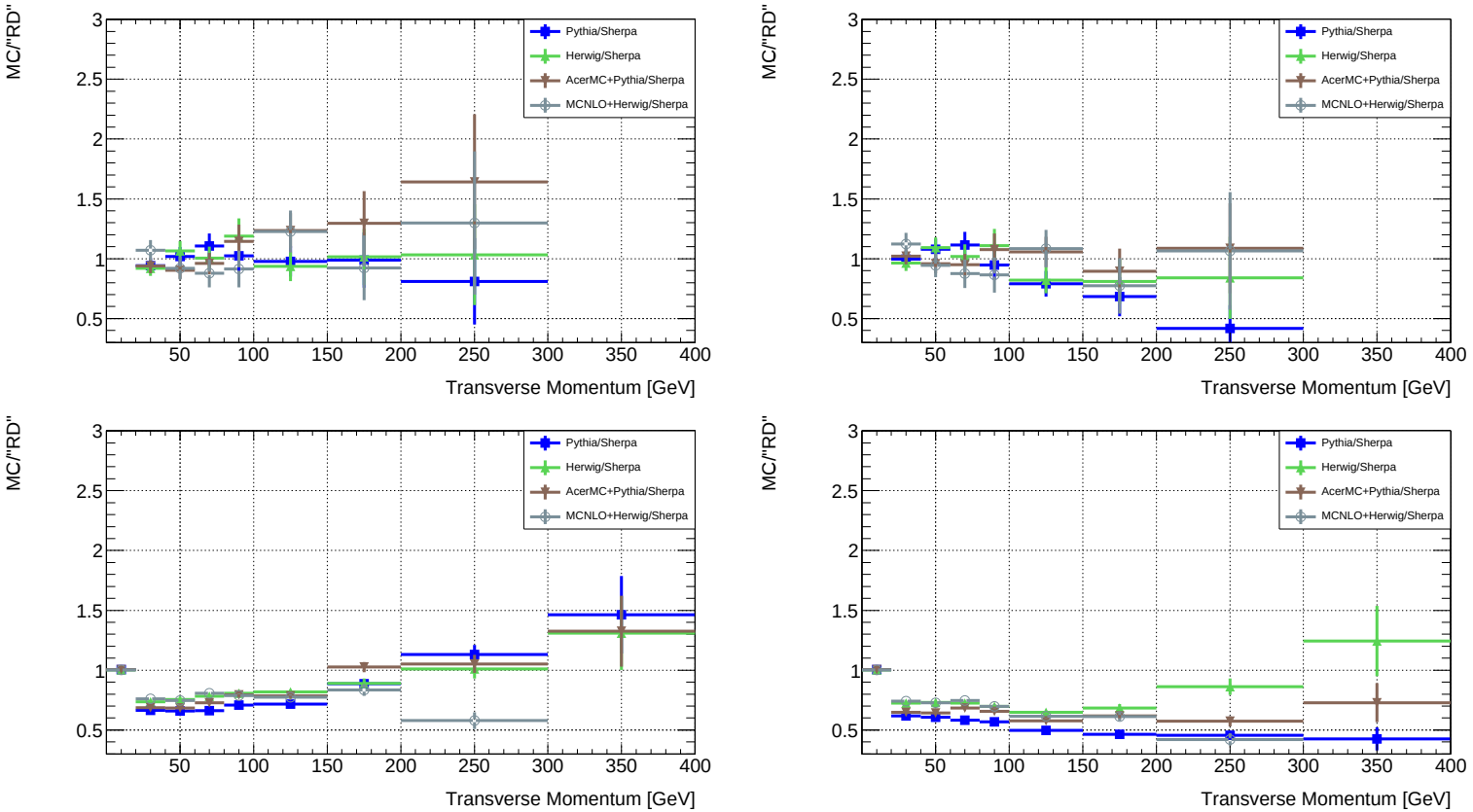

Figure B.4: Transverse momentum of electrons in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

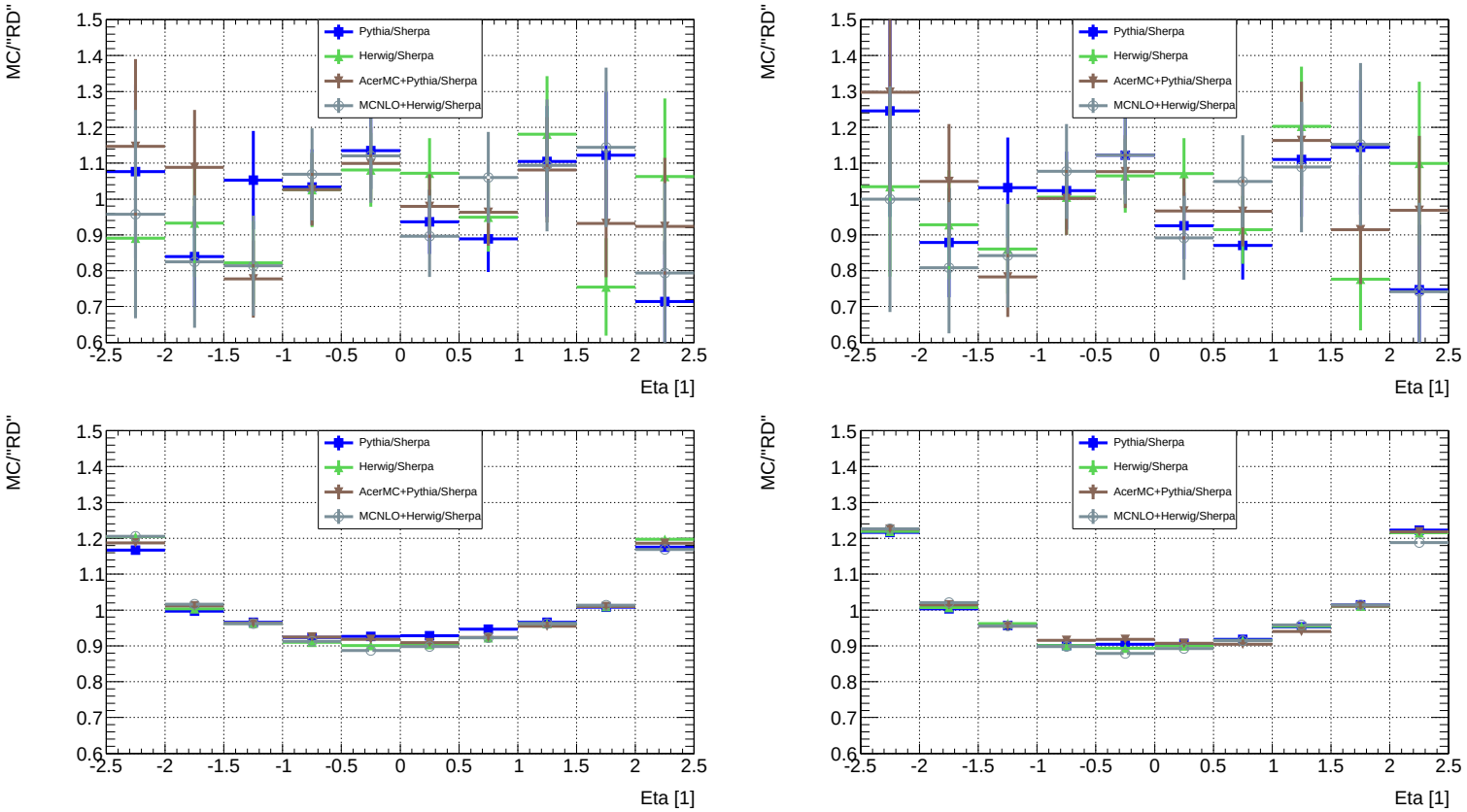

Figure B.5: Pseudrapidity of electrons in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

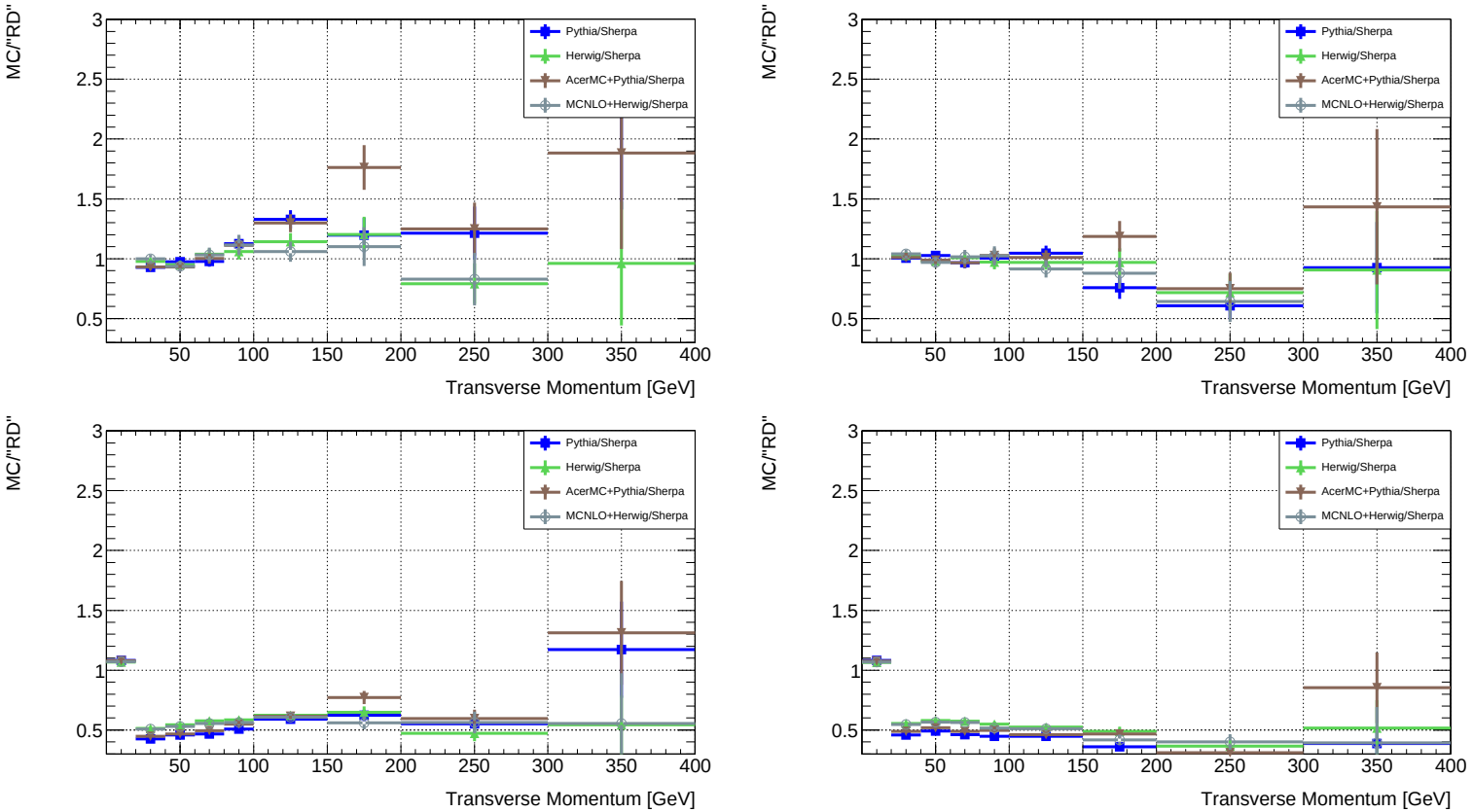

Figure B.6: Transverse momentum of muons in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

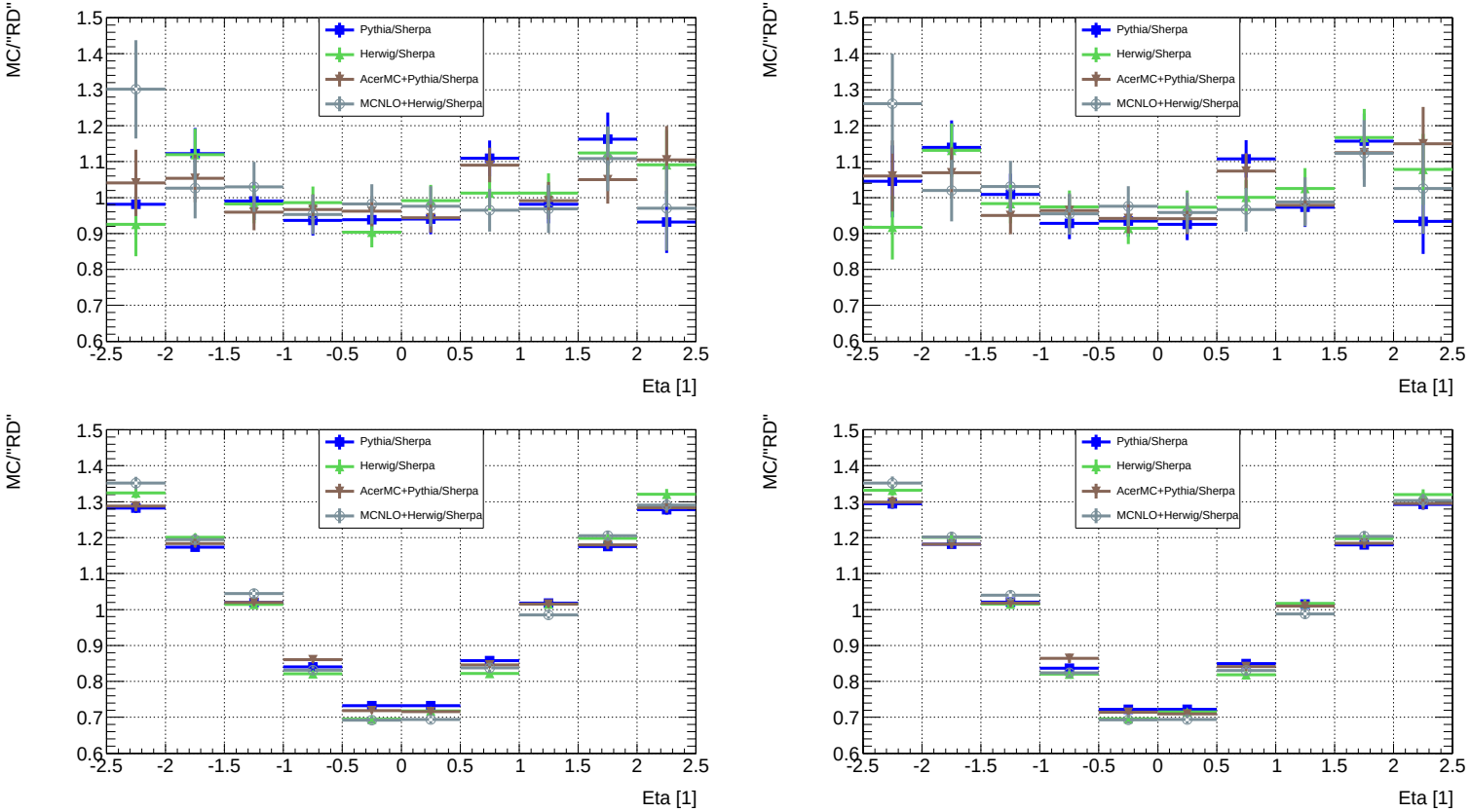

Figure B.7: Pseudrapidity of muons in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

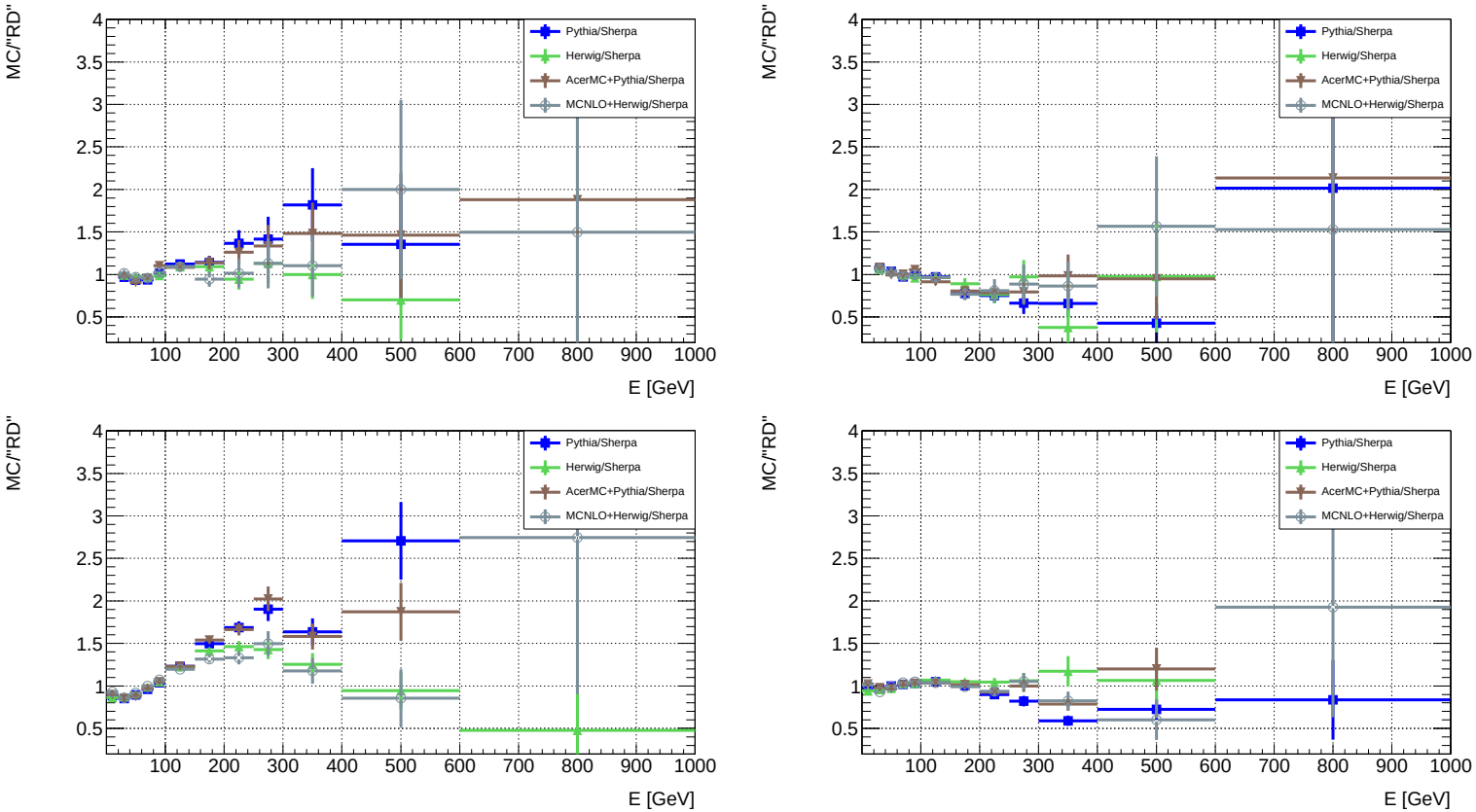

Figure B.8: Missing transverse energy in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

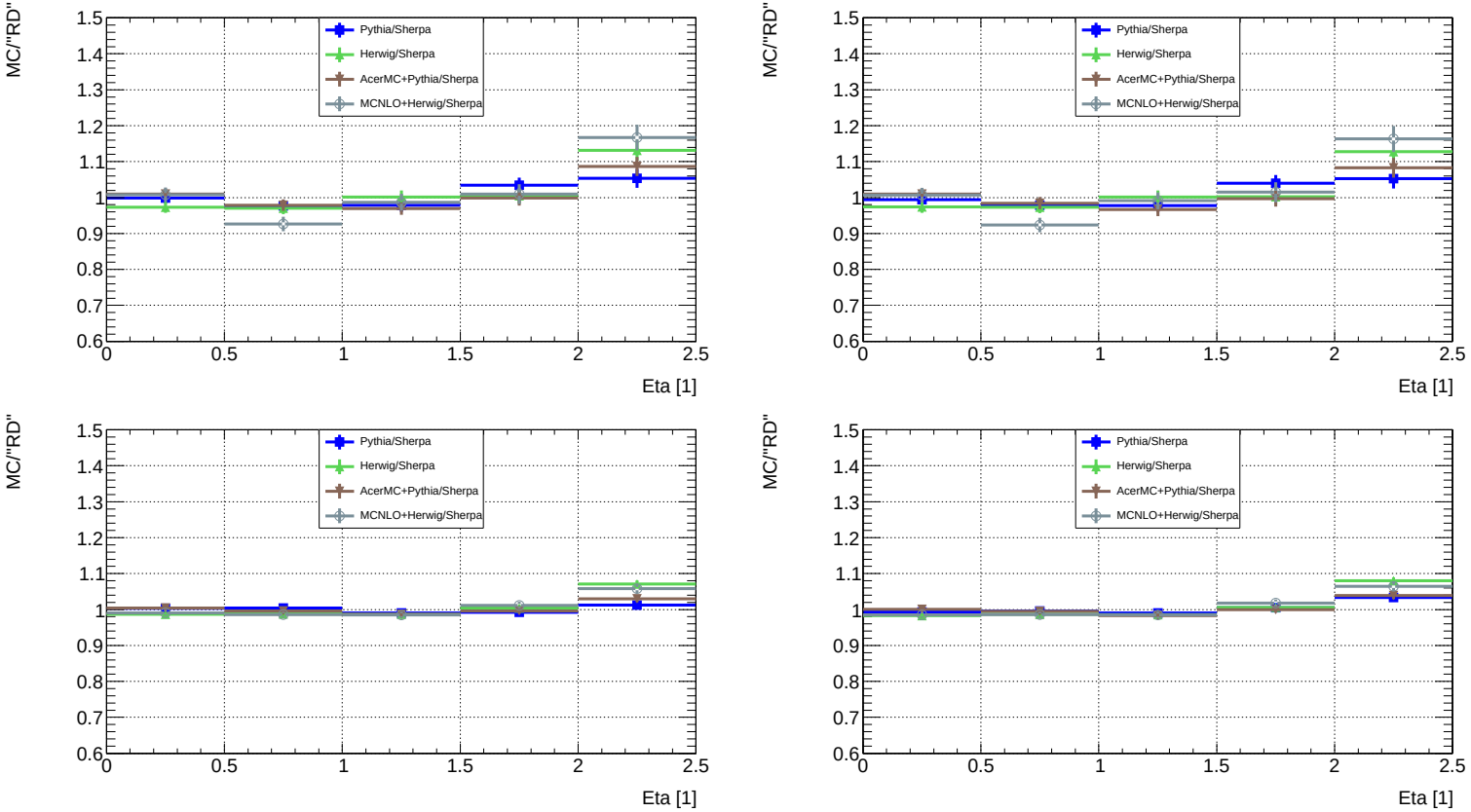

Figure B.9: Pseudrapidity of jets in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

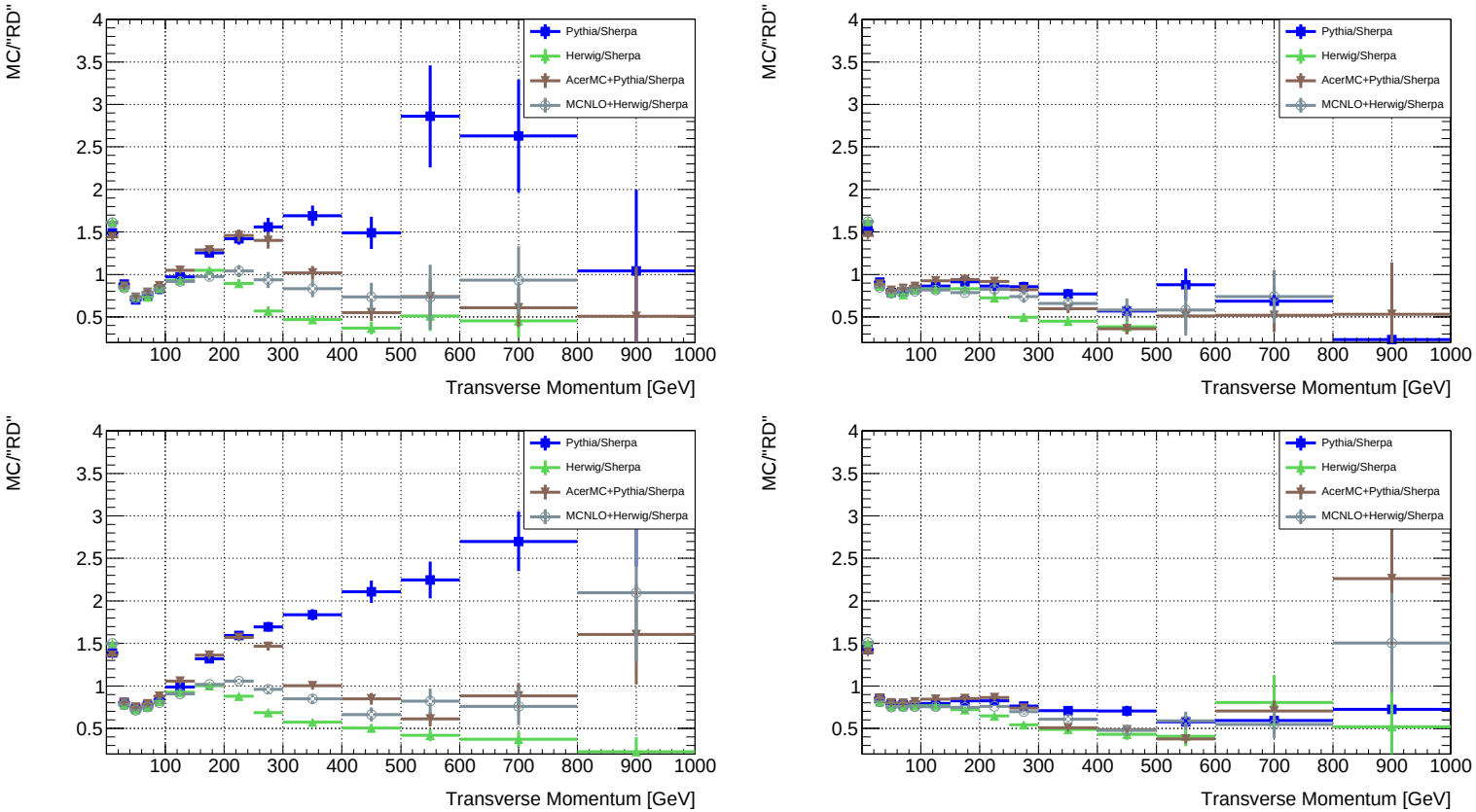

Figure B.10: Transverse momentum of jets in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

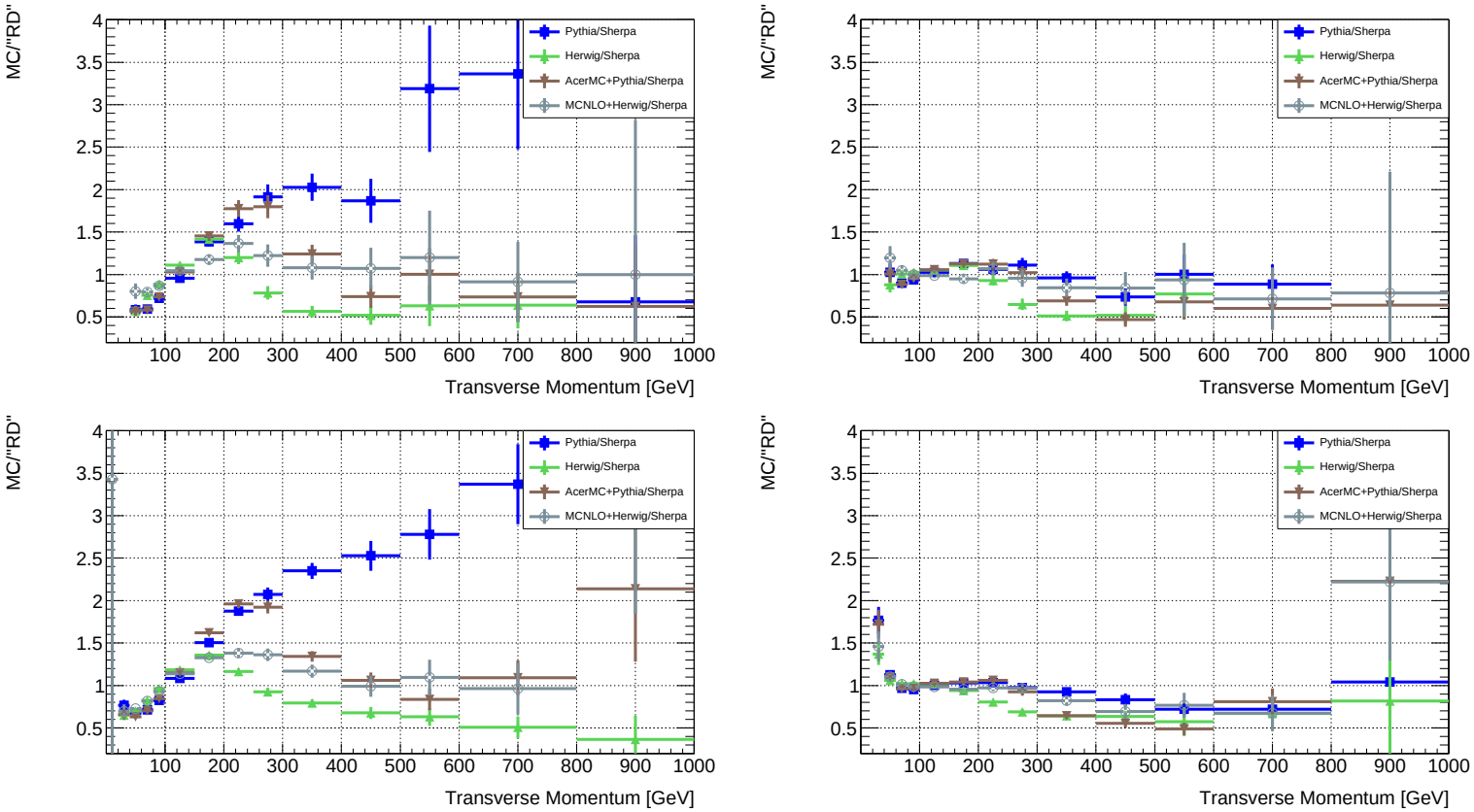

Figure B.11: Transverse momentum of the first leading jet in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

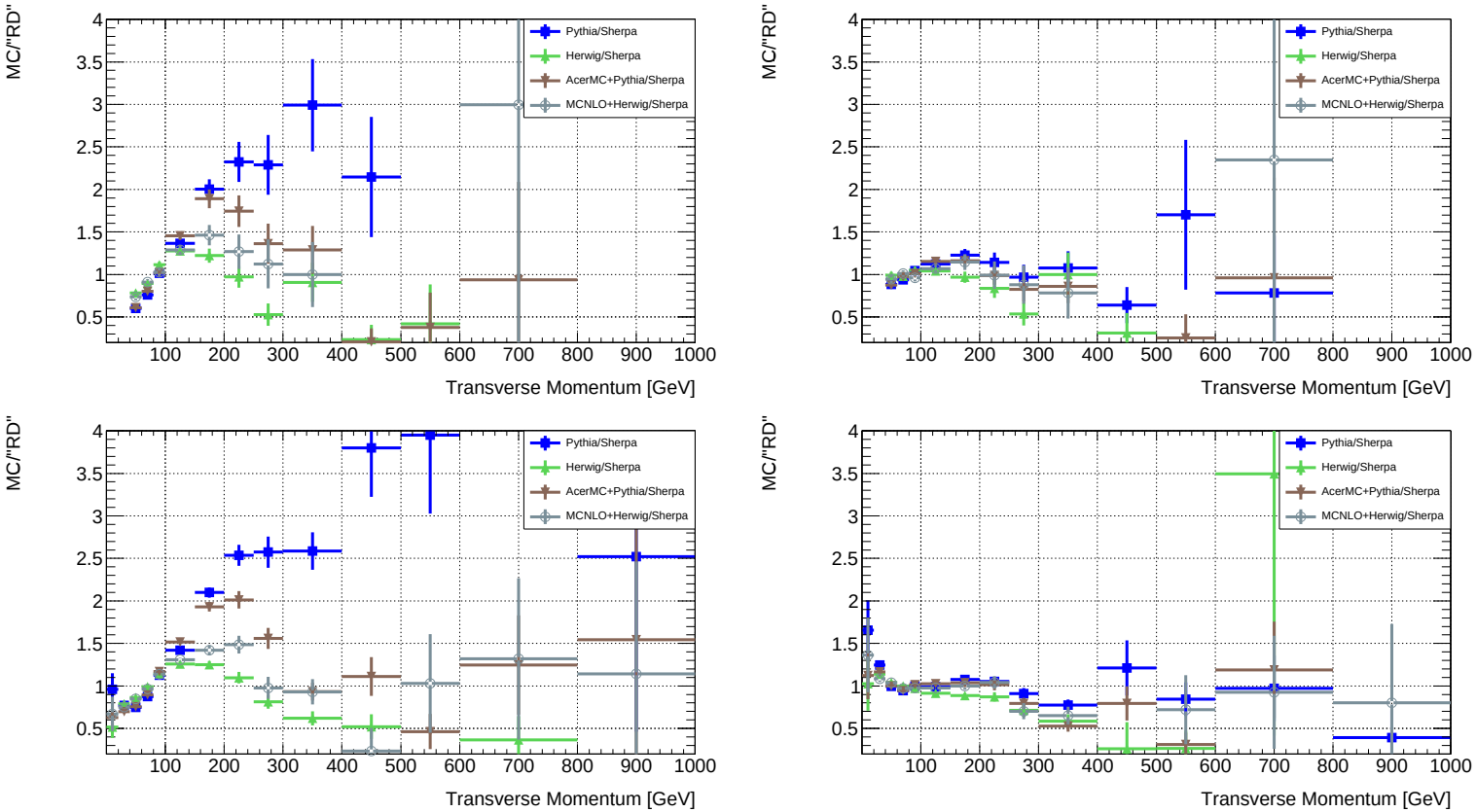

Figure B.12: Transverse momentum of the second leading jet in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

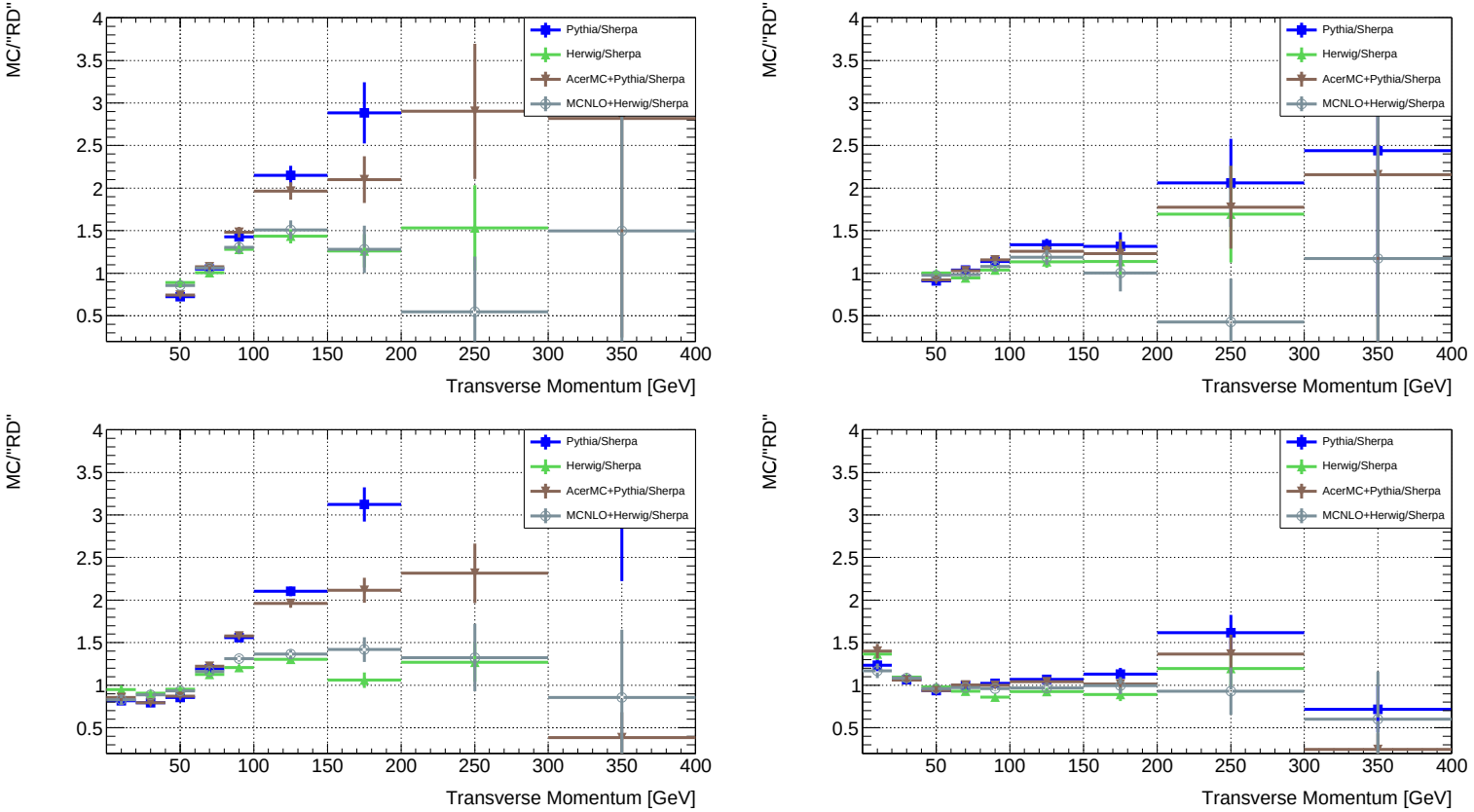

Figure B.13: Transverse momentum of the third leading jet in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

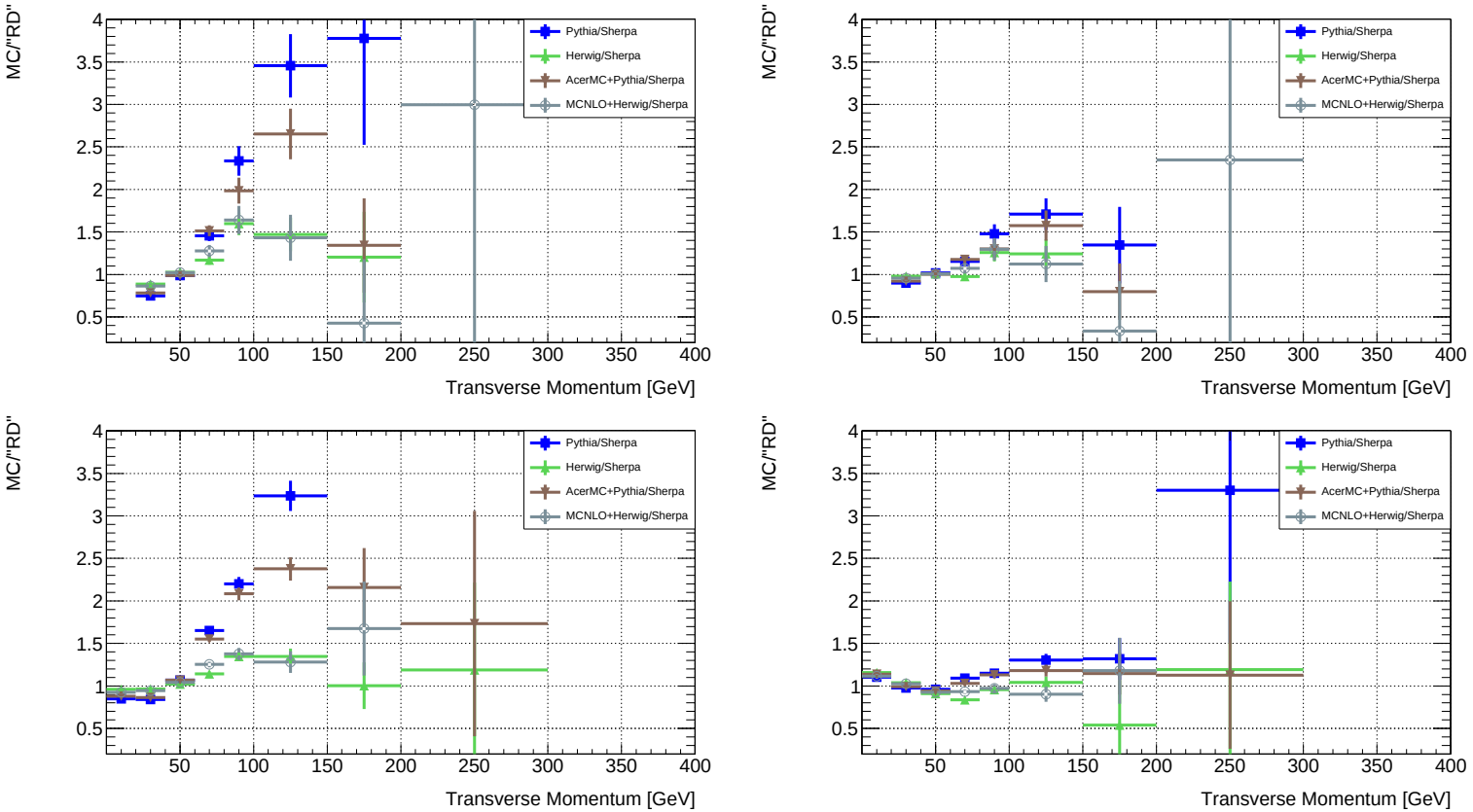

Figure B.14: Transverse momentum of the fourth leading jet in semileptonic $t\bar{t}$ events for four common Monte Carlo generators relative to pseudo data (Shepra arbitrarily chosen). At the top: reconstructed events with cuts, at the bottom: generated events without cuts. At the left: before reweighting, at the right: after reweighting.

List of Figures

List of Tables

Glossary

Notation	Description	
ALICE	A Lhc Ion Collider Experiment	3, 17
AOD	Analysis Object Data	23, 51, 79, 99, 123
ASIC	Application Specific Integrated Circuit	33
ATLAS	A Toroidal Lhc ApparatuS experiment	3, 17
BE	Back-End	23
BIS	Beam Interlock System	22
BOC	Back Of Crate card	34
BPM	Bi-Phase Mark	35
CAN	Controller Area Network	23
CERN	European Organization for Nuclear Research	3, 17
CKM	Cabibbo-Kobayashi-Maskawa	9
CMS	Compact Muon Solenoid experiment	3, 17
CPLD	Complex Programmable Logic Device	37, 38
CSC	Cathode Strip Chamber	25
DAQ	Data AcQuisition	22
DCS	Detector Control System	22
DPD	Derived Physics Data	23, 51, 79
DSS	Detector Safety System	22
EB	Event Builder	23
ECAL	Electromagnetic CALorimeter	26
EF	Event Filter	23
ELMB	Embedded Local Monitor Board	23
ESD	ElectroStatic Discharge	41
ESD	Event Summary Data	23, 51, 79
FCAL	Forward CALorimeter	26
FE	Front-End chip of pixel detector	33
FE	Front-End of DCS	23

Notation	Description	
FPIAA	Finding Persons Inside Atlas Area	22
GCS	Global Control Station	23
GEANT4	GEometry ANd Tracking version 4	77
GRIN	GRadient INdex	42, 48
HCAL	Hadronic CALorimeter	26
HEC	Hadronic Endcap Calorimeter	27
HLT	High-Level Trigger	23
HOLA	High speed Optical Link for Atlas	37
HV	High Voltage	33
ID	Inner Detector	27
IP	Impact Parameter	61
L1A	Level-1 Accept	23
LAr	Liquid Argon	27
LCG	Lhc Computing Grid	23
LCS	Local Control Stations	23
LHC	Large Hadron Collider	3, 9, 17
LHCb	LHC beauty experiment	3, 17
LLA	Leading Logarithmic Accuracy	13
LO	Leading Order	12
LVL1	LeVeL-1 trigger	22
LVL2	LeVeL-2 trigger	23
MC	Monte Carlo	4, 5, 51
MCC	Module Control Chip	33
MDT	Monitored Drift-Tube chamber	25
MM	Multi Mode	47
MT	Mechanical Transfer	43
NLO	Next to Leading Order	12
NTC	Negative Temperature Coefficient	33
OB	OptoBoard	33, 34
OTDR	Optical Time Domain Reflectometer	44, 47
PC	Personal Computer	44
PDF	Parton Distribution Function	12
PIN	P-type, Intrinsic, N-type	37
PP	Patch Panel	39

Notation	Description	
PS	Proton Synchrotron	17
PV	Primary Vertex	61
QCD	Quantum Chromo-Dynamics	7
QED	Quantum Electro-Dynamics	8
RCC	Readout Crate Controller	34
RD	Real Data	66
RDO	Raw Data Object	23, 51, 79
ROB	Read Out Buffer	23, 35
ROD	Read Out Driver	23, 34
RoI	Regions of Interest	23
RoIB	Regions of Interest Builder	23
ROOT	"ROOT is an object-oriented framework aimed at solving the data analysis challenges of high-energy physics"	99, 123
RPC	Resistive Plate Chamber	25
RX	Receiver	37, 39
SBC	Single Board Computer	34
SCS	Sub-detector Control Station	23
SCT	SemiConductor Tracker	27, 35
SIMM	Stepped Index Multi Mode	42, 48
SM	Single Mode	44
SM	Standard Model	5
SMC	Sub Miniature C	43
SPS	Super Proton Synchrotron	17
ST	Stab and Twist	44
SV	Secondary Vertex	61
TDAQ	Trigger and Data AcQuisition	22
TGC	Thin Gap Chamber	25
TIM	Ttc Interface Module	34
TMT	Thermal Management Tile	33
TRT	Transition Radiation Tracker	27
TTC	Timing, Trigger and Control	22, 34
TX	Transmitter	37, 39
USB	Universal Serial Bus	44
VCSEL	Vertical Cavity Surface Emitting Laser	37
VME	VersaModule Eurocard	34

Bibliography

[1] R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley 1998.

[2] R. P. Feynman, *The Meaning of It All*, Perseus Books 1998.

[3] J. Rohlf, *Modern Physics from α to Z^0*, John Wiley & Sons 1994.

[4] ALEPH, DELPHI, L3 and OPAL Collaborations, Phys. Lett. **B565** (2003) 61.

[5] *Combined CDF and DZero Upper Limits on Standard Model Higgs-Boson Production with up to 4.2 fb-1 of Data*, FERMILAB-PUB-09-060-E 2009
`http://arxiv.org/abs/0903.4001`

[6] *Design report Tevatron 1 Project*. FERMILAB-DESIGN-1982-01, 1982
`http://lss.fnal.gov/archive/design/fermilab-design-1982-01.pdf`
`http://www-bdnew.fnal.gov/tevatron/`

[7] CDF Collaboration, Phys. Rev. **D50** (1994) 2966 and Phys. Rev. **D52** (1995) 4784.
`http://www-cdf.fnal.gov/`

[8] DØ Collaboration, Phys. Rev. Lett. **77** (1996) 3309.
`http://www-d0.fnal.gov/`

[9] *Combination of CDF and DØ Results on the Mass of the Top Quark*, FERMILAB-TM-2427-E 2009
`http://arxiv.org/abs/0903.2503`
`http://tevewwg.fnal.gov/top/`

[10] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, "NLL resummation of the heavy-quark hadroproduction cross-section", *Nucl. Phys.* **B529** (1998) 424–450.
`http://arxiv.org/abs/hep-ph/9801375`

[11] T. Flick, *Studies on the Optical Readout for the ATLAS Pixel Detector*, PhD thesis, Universität Wuppertal, 2006.

[12] Particle Data Group, The European Physical Journal B **667,1** (2008) http://pdg.lbl.gov

[13] E. Thomson, *Top Quark Pair Production Cross Section Combination* http://www.physics.upenn.edu/~thomsone/2006_1030_dpf_v5-1.pdf

[14] Dobbs, Frixione, Laenen, Tollefson: *Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics*. 2004 http://arxiv.org/abs/hep-ph/0403045

[15] F. Krauss: Sketch of a proton-proton collision event simulated by Monte Carlo.

[16] *Parton Fragmentation and String Dynamics*, Bo Andersson, G. Gustafson, G. Ingelman, T. Sjostrand (Lund U., Dept. Theor. Phys.), LU-TP-83-10, Jul 1983. Phys.Rept.97:31-145,1983.

[17] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, B. R. Webber: *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010. http://arxiv.org/abs/hep-ph/0011363 http://hepwww.rl.ac.uk/theory/seymour/herwig/

[18] T. Sjöstrand, S. Mrenna, P. Skands, *Pythia 6.4 Physics and Manual*, LU TP 06-13, FERMILAB-PUB-06-052-CD-T, 2006. http://home.thep.lu.se/~torbjorn/Pythia.html

[19] B. Kersevan, E. Richter-Was: *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, Comput. Phys. Commun. **149** (2003) 142. http://arxiv.org/abs/hep-ph/0405247 http://borut.home.cern.ch/borut/

[20] T. Gleisberg et al.: *Event generation with SHERPA 1.1*, (2008), SLAC-PUB-13420 http://arxiv.org/pdf/0811.4622 http://www.sherpa-mc.de/

[21] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029. http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/

[22] UA5 Collaboration, G. J. Alner et al., Nucl. Phys. **B 291**, (1987) 445.

[23] J. M. Butterworth, J. R. Forshaw, M. H. Seymour, Z. Phys. **C 72** (1996) 637.
http://arxiv.org/abs/hep-ph/9601371

[24] *LHC Design Report*, CERN-2004-003-V-1
http://cdsweb.cern.ch/record/782076
http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html
http://lhc.web.cern.ch/lhc/

[25] *ALICE physics performance: Technical Design Report*, ALICE-TDR-13, CERN-LHCC-2005-030
http://cdsweb.cern.ch/record/879894
http://aliceinfo.cern.ch

[26] *ATLAS detector and physics performance: Technical Design Report 1+2*, ATLAS-TDR-014, CERN-LHCC-99-014; ATLAS-TDR-015, CERN-LHCC-99-015
http://cdsweb.cern.ch/record/391176
http://cdsweb.cern.ch/record/391177
http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/TDR/access.html
http://www.atlas.ch/

[27] *CMS physics: Technical Design Report 1+2*, CERN-LHCC-2006-001, CMS-TDR-008-1; CERN-LHCC-2006-021, CMS-TDR-008-2
http://cdsweb.cern.ch/record/922757
http://cdsweb.cern.ch/record/942733
http://cms.cern.ch

[28] *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, LHCb-TDR-9
http://cdsweb.cern.ch/record/630827
http://lhcb.web.cern.ch/

[29] Proton Synchrotron at CERN webpage:
http://www.cern.ch/CERN/Divisions/PS/Welcome.html

[30] Super Synchrotron at CERN webpage:
http://ab-dep-op-sps.web.cern.ch/ab-dep-op-sps/

[31] Overall view of LHC experiments
http://cdsweb.cern.ch/record/841555?ln=en

[32] Cross section of the vacuum pipes inside a dipole magnet
https://edms.cern.ch/document/101865/1.1

[33] LHC luminosity webpage:
http://ab-dep-bi-pm.web.cern.ch/ab-dep-bi-pm/?n=Activities.BRAN

[34] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
https://twiki.cern.ch/twiki/pub/Atlas/AtlasTechnicalPaper/Published_version_jinst8_08_s08003.pdf
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTechnicalPaperListOfFigures

[35] Computer generated image of the whole ATLAS detector
http://mediaarchive.cern.ch/MediaArchive/Photo/Public/2008/0803012/0803012_01/0803012_01-Icon.jpg

[36] *ATLAS inner detector: Technical Design Report 1+2*, ATLAS-TDR-004, CERN-LHCC-97-016; ATLAS-TDR-005, CERN-LHCC-97-017
http://cdsweb.cern.ch/record/331063
http://cdsweb.cern.ch/record/331064
https://twiki.cern.ch/twiki/bin/view/Atlas/InnerDetector

[37] *ATLAS calorimeter performance: Technical Design Report*, ATLAS-TDR-001, CERN-LHCC-96-040
http://cdsweb.cern.ch/record/331059
http://atlas.web.cern.ch/Atlas/TDR/caloperf/caloperf.html

[38] *ATLAS muon spectrometer: Technical Design Report*, ATLAS-TDR-010, CERN-LHCC-97-022
http://cdsweb.cern.ch/record/331068
http://atlas.web.cern.ch/Atlas/GROUPS/MUON/TDR/Web/TDR.html

[39] *Event Cross Section in a computer generated image of the ATLAS detector*, CERN-GE-0803022
http://cdsweb.cern.ch/record/1096081

[40] ATLAS magnet system: Technical Design Report, ATLAS-TDR-006, CERN-LHCC-97-018
http://cdsweb.cern.ch/record/338080

[41] *ATLAS central solenoid: Technical Design Report*, ATLAS-TDR-009, CERN-LHCC-97-021
http://cdsweb.cern.ch/record/331067

[42] *ATLAS barrel toroid: Technical Design Report*, ATLAS-TDR-007, CERN-LHCC-97-019
http://cdsweb.cern.ch/record/331065

[43] *ATLAS endcap toroids: Technical Design Report*, ATLAS-TDR-008, CERN-LHCC-97-020
http://cdsweb.cern.ch/record/331066

[44] S. Baranov, M. Bosman, I. Dawson, V. Hedberg, A. Nisati, M. Shupe, *Estimation of Radiation Background, Impact on Detectors, Activation and Shielding Optimization in ATLAS*, ATL-GEN-2005-001, ATL-COM-GEN-2005-001, CERN-ATL-GEN-2005-001
http://cdsweb.cern.ch/record/814823

[45] *Hierarchical Control of the ATLAS Experiment*, CERN-THESIS-2007-037
http://cdsweb.cern.ch/record/1034400

[46] Burckhart, H. J., Cook, J., Filiminov, V., Franz, T., Gutzwiller, O., Khomutnikov, V., Schlenker, S., *The Common Infrastructure Control of the ATLAS experiment*, Topical Workshop on Electronics for Particle Physics, Naxos, Greece, 15–19 Sep 2008, pp.428–431
http://cdsweb.cern.ch/record/1159528

[47] *Machine Protection for the LHC: Architecture of the Beam and Powering Interlock Systems*, LHC-Project-Report-521, CERN-LHC-Project-Report-521
http://cdsweb.cern.ch/record/531820

[48] B.G. Taylor: *Timing distribution at the LHC*. Workshop on Electronics for LHC Experiments, 2002.
http://cdsweb.cern.ch/record/592719

[49] *The ATLAS central level-1 trigger logic and TTC system*, ATL-COM-DAQ-2008-006
http://cdsweb.cern.ch/record/1119310

[50] C. Cardeira, O. Beltramello, and H. J. Burckhart et al., *Communication Architecture of a system to find persons inside ATLAS, proceedings of the 6th IEEE International Workshop on Factory Communication Systems*, Torino (2006) 405–409

[51] O. Beltramello, H. Burckhart, S. Franz, M. Jaekel, M. Jeckel, S. Lueders, G. Morpurgo, F. dos Santos Pedrosa, K. Pommes, H. Sandaker, *The ATLAS Detector Safety System*

[52] A. Barriuso Poy et al., *The detector control system of the ATLAS experiment*, JINST **3** (2008) P05006
http://www.iop.org/EJ/article/1748-0221/3/05/P05006/jinst8_05_p05006.pdf
http://atlas-dcs.web.cern.ch/atlas-dcs/

[53] *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, ATLAS-TDR-016, CERN-LHCC-2003-022
http://cdsweb.cern.ch/record/616089

[54] *ATLAS DAQ and Controls*, ATL-DAQ-SLIDE-2009-094, ATL-COM-DAQ-2009-032
http://cdsweb.cern.ch/record/1174769

[55] *ATLAS level-1 trigger: Technical Design Report*, ATLAS-TDR-012, CERN-LHCC-98-014
http://cdsweb.cern.ch/record/381429

[56] *ATLAS computing: Technical Design Report*, ATLAS-TDR-017, CERN-LHCC-2005-022
http://cdsweb.cern.ch/record/837738
http://atlas-proj-computing-tdr.web.cern.ch/atlas-proj-computing-tdr/Html/
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasComputing

[57] *The ATLAS Analysis Architecture*, ATL-SLIDE-2007-028, CERN-ATL-SLIDE-2007-028
http://cdsweb.cern.ch/record/1047631

[58] *LHC computing Grid: Technical Design Report*, CERN-LHCC-2005-024, LCG-TDR-001
http://cdsweb.cern.ch/record/840543
http://lcg.web.cern.ch/LCG/

[59] J. Shiers, *The Worldwide LHC Computing Grid (LCG)*, Comput. Phys. Commun. **177, 1–2** (2007) 219–223.
http://cdsweb.cern.ch/record/1063847

[60] *The ATLAS Computing Model*, ATL-SOFT-2004-007, ATL-COM-SOFT-2004-009, CERN-ATL-COM-SOFT-2004-009, CERN-LHCC-2004-037-G-085
http://cdsweb.cern.ch/record/811058

[61] The C++ Standards Committee webpage:
http://www.open-std.org/jtc1/sc22/wg21/

[62] B. Stroustrup, *The C++ Programming Language*. Addison-Wesley 2000.

[63] The Fortran Standards webpage:
http://www.nag.co.uk/sc22wg5/

[64] The Java Community Process Program
http://www.jcp.org/en/home/index

[65] Python Programming Language webpage:
http://www.python.org/

[66] ATLAS Athena Framework wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/AthenaFramework

[67] *ATLAS tile calorimeter: Technical Design Report*, ATLAS-TDR-003, CERN-LHCC-96-042
http://cdsweb.cern.ch/record/331062

[68] *ATLAS liquid-argon calorimeter: Technical Design Report*, ATLAS-TDR-002, CERN-LHCC-96-041
http://cdsweb.cern.ch/record/331061

[69] M. Capéans-Garrido, *The transition radiation tracker of the ATLAS experiment*, IEEE Trans. Nucl. Sci. **51** (2004) 994–1000.
http://cdsweb.cern.ch/record/816797

[70] J. N. Jackson, *The ATLAS semiconductor tracker (SCT)*, Nucl. Instrum. Methods Phys. Res., **A 541** (2005) 89–95.
http://cdsweb.cern.ch/record/908854?ln=en
ATLAS SCT detector webpage:
http://atlas.web.cern.ch/Atlas/GROUPS/INNER_DETECTOR/SCT/
ATLAS SCT detector wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/SctWiki

[71] *ATLAS pixel detector: Technical Design Report*, ATLAS-TDR-011, CERN-LHCC-98-013
http://cdsweb.cern.ch/record/381263
ATLAS Pixel detector webpage:
http://atlas.web.cern.ch/Atlas/GROUPS/INNER_DETECTOR/PIXELS/pixel.html
ATLAS Pixel wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/PixelWiki

[72] A 3D model of the pixel detector and it's framework
http://hep.phys.sfu.ca/openhouse_2008/kiosk/gallery/atlas_photos/selected-photos/inner-detector/pixel/Pixels_Complete_Opened.jpg

[73] F. Hugging et al., *Design and test of pixel sensors for operation in severe radiation environments*, Nucl. Instrum. Meth., **A439** (2000) 529.

[74] T. Stockmanns, *Multi-Chip-Modul-Entwicklung für den ATLAS Pixeldetektor: Analyse der Front-End-Chip-Elektronik in strahlenharter 0,25-µm-Technologie und Entwicklung und Realisierung eines Serial-Powering-Konzeptes*, PhD thesis, Universität Bonn, July 2004.

[75] E. Mandelli et al., *Digital Column Readout Architecture for the ATLAS Pixel* 0.25 *µm Front End IC*, IEEE, 0-7803-7324-3/02 2002.
`http://ieeexplore.ieee.org/iel5/7884/21725/01008527.pdf`

[76] R. Beccherle et al., *The Module Controller Chip for the ATLAS Pixel Detector*, NIM-A, **492** (2002) 117–133.

[77] K. E. Arms et al., *ATLAS pixel opto-electronics*, NIM-A, **554** (2005) 458–468.
`http://www.physics.ohio-state.edu/~gan/optoboard/NIM.pdf`

[78] BOC card documentation webpage:
`http://www.atlas.uni-wuppertal.de/~bocprod/boc_docu.html`

[79] ROD card documentation webpage:
`http://wisconsin.cern.ch/ROD/ATLAS_SiRod.shtml`

[80] TIM card documentation webpage:
`http://www.hep.ucl.ac.uk/atlas/sct/tim/`

[81] I.-M. Gregor, *Optical Links for the ATLAS pixel detector*, PhD thesis, Universität Wuppertal, 2001.

[82] Cavendish Laboratory Cambridge webpage:
`http://www.phy.cam.ac.uk/`

[83] M. L. Chu et al., *The off-detector opto-electronics for the optical links of the ATLAS semiconductor tracker and pixel detector*, Nucl. Instrum. Meth., **A530** (2004) 293–310.
`http://dialnet.unirioja.es/servlet/articulo?codigo=972163`

[84] E. van der Bij, R. A. McLaren, O. Boyle and G. Rubin: *S-LINK, a data link interface specification for the LHC era*, IEEE Trans. Nucl. Sci., **44** (1997) 398–402.

[85] The CERN S-Link webpage:
`http://hsi.web.cern.ch/HSI/s-link/`

[86] Turck Elektronik/Fertigung webpage:
http://electronic-production.eu/

[87] M. Goodrick: *BOC - The Back of Crate interface for the SCT ROD*. Technical report, Cavendish-Laboratory, Cambridge, 2003.

[88] BOC card production webpage:
http://www.atlas.uni-wuppertal.de/~bocprod/boc_production.php

[89] Pixel Production Database:
http://www.ge.infn.it/~atlas/pdb/

[90] ATLAS pixel connectivity table:
https://edms.cern.ch/file/447936/3/ATL-IP-ES-0102_AppendixA_v3.91.xls

[91] BOC/ROD installation tables:
https://twiki.cern.ch/twiki/pub/Atlas/CountingRoomsInstallation/AtlasPixelUSA15BOCRODInstallationVer1.xls

[92] Andreas Korn's talk about DAQ crate signoff:
http://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=23435

[93] ATLAS pixel TX-plugin failures wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/PixelTxPlugins

[94] G. Lenzen, *Pixel Opto Services*, ATL-IP-ES-0070, January 2004.

[95] Arvidsson, Björk, Pearce, Troska, Vasey, Zanet: *A Dense Multi-Ribbon Cable For Installation in a Harsh Environment at CERN*, 2000.

[96] Luciol ν-OTDR
http://www.luciol.com/

[97] Wiki of ATLAS pixel connectivity test in SR1-building:
https://twiki.cern.ch/twiki/bin/view/Atlas/ConnectivityTestResults

[98] ATLAS pixel connectivity table of optical cables:
https://twiki.cern.ch/twiki/pub/Atlas/PixelConnectivity/ATL-IP-ES-0102_ver4.1_Stephan_Georg20071024.xls
https://twiki.cern.ch/twiki/bin/viewfile/Atlas/PixelConnection?rev=1;filename=ATL-IP-ES-0102_ver4.1_Stephan_Georg20071024_swaps_V1_CorrectDB_AfterTestFinal.xls

[99] Test results of the ATLAS pixel optical cables:
https://twiki.cern.ch/twiki/pub/Atlas/PixelConnectivity/Connectivity.doc

[100] Data of the OTDR test of the optical cables of the ATLAS pixel detector:
/castor/cern.ch/user/s/sandvoss/atlas/hardware/pixel/OptoCablesRibbonsFibres/pit/USA15/afterribboninstall

[101] ATLAS pixel connectivity test wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/PixelConnectivityTest

[102] Optical fiber connection test results:
http://indico.cern.ch/getFile.py/access?contribId=0&resId=0&materialId=slides&confId=31220

[103] ATLAS Reconstruction wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/ReconstructionDocumentation

[104] T. Cornelissen et al., *Concepts, Design and Implementation of the ATLAS New Tracking*, ATLAS Note ATL-SOFT-PUB-2007-007 (2007).
http://cdsweb.cern.ch/record/1020106

[105] V. Kartvelishvili, *Electron bremsstrahlung recovery in ATLAS*, Nucl. Phys. **172** (Proc. Suppl.) (2007) 208.

[106] R. Frühwirth, *Track fitting with non-Gaussian noise*, Comput. Phys. Comm. **100** (1997) 1.

[107] R. Frühwirth and A. Strandlie, *Track fitting with ambituities and noise: a study of elastic tracking and nonlinear filters*, Comput. Phys. Commun. **120** (1999) 197.

[108] F. Bergsma, *Calibration of Hall sensors in three dimensions*, in *Proceedings of the 13th international magnetic measurement workshop*, Stanford USA (2003).
http://cdsweb.cern.ch/record/1072471

[109] M. Aleksa et al., *Measurement of the ATLAS solenoid magnetic field*, 2007 JINST **3** P04003.

[110] A. Salzburger: *A Parametrization for Fast Simulation of Muon Tracks in the ATLAS Inner Detector and Muon System*, Diploma Thesis Innsbruck 2003
http://cdsweb.cern.ch/record/813003

[111] P. Billoir, S. Quan, *Fast vertex fitting with a local parametrization of tracks*, Nucl. Instrum. Meth. **A311** (1992) 139–150.

[112] R. Frühwirth, *Application of Kalman Filtering to Track and Vertex Fitting*, Nucl. Inistrum. and Methods **225** (1984) 352.

[113] I. Abt et al., *The tracking, calorimeter and muon detectors of the H1 experiment at HERA*, Nucl. Instrum. Meth. **A 386** (1997) 348.

[114] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment; Detector, Trigger and Physics*, arXiv:0901.0512v3. April 2009.

[115] H. Bachacou: *Mistags at CDF*. ATLAS Flavour Tagging Workshop 2007. `http://indico.cern.ch/getFile.py/access?contribId=64&sessionId=12&resId=0&materialId=slides&confId=14475`

[116] C. Gwilliam, S. Burdin, *Track Impact Parameter Tuning* `http://indico.cern.ch/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=41360`

[117] MC AOD dataset 5200 ($t\bar{t}$ simulated with MC@NLO, Jimmy and Herwig) with $6 \cdot 10^5$ events, reconstructed with athena 13.0.30.4 with misaligned detector and trigger information: `trig1_misal1_mc12.005200.T1_McAtNlo_Jimmy.recon.AOD.v13003004_tid018416`

[118] MC AOD dataset 5200 ($t\bar{t}$ simulated with MC@NLO, Jimmy and Herwig) with $5 \cdot 10^4$ events, reconstructed with athena 13.0.30.4 with heavily misaligned detector (special V6) and trigger information: `special1_misal1_mc12_V6.005200.T1_McAtNlo_Jimmy.recon.AOD.v13003004`

[119] G. Borisov, C. Mariotti, *Fine tuning of track impact parameter resolution of the DELPHI detector*, Nucl. Inst. Meth **A372** (1996) 181.

[120] Burdin, Anzelc, *Decay Length Resolution Studies*, Application to B_s^0 Mixing and p17 Calibration. DØ Note 5336, 2007

[121] G. Cowan, *Statistical data analysis*, Oxford Science Publications 2002.

[122] ATLAS Generation wiki: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MonteCarloWorkingGroup`

[123] J. Allison et al., *Geant4 Developments and Applications*, IEEE Transactions on Nuclear Science **53 No. 1** (2006) 270–278.
S. Agostinelli et al., *Geant4 - A Simulation Toolkit*, Nuclear Instruments

and Methods **A 506** (2003) 250–303.
http://geant4.web.cern.ch/geant4/

[124] *ATLFAST 2.0 a fast simulation package for ATLAS*, ATL-PHYS-98-131
http://cdsweb.cern.ch/record/683751
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlfastII
http://indico.cern.ch/getFile.py/access?contribId=82&
sessionId=22&resId=0&materialId=slides&confId=52623

[125] ATLAS Digitization wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasDigitization

[126] ATLAS Simulation wiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasSimulation

[127] *ATLFAST 1.0 A package for particle-level analysis*, ATL-PHYS-96-079, ATL-GE-PN-79
http://cdsweb.cern.ch/record/682460
https://twiki.cern.ch/twiki/bin/view/Atlas/
AtlfastDocumentation

[128] *The Fast ATLAS Track Simulation (FATRAS)*, ATL-SOFT-PUB-2008-001, ATL-COM-SOFT-2008-002
http://cdsweb.cern.ch/record/1091969
https://twiki.cern.ch/twiki/bin/view/Atlas/FatRas

[129] J. Pumplin et al., JHEP **07** (2002) 012.

[130] J. Pumplin, A. Belyaev, J. Huston, D. Stump and W. K. Tung, JHEP **02** (2006) 032.

[131] The 5 MC top/antitop datasets:
grid-ui.physik.uni-wuppertal.de:/sfsscratch/atlas/sandvoss/
data/atlas/software/mcgenerators/

[132] The scripts for the production of the 5 MC top/antitop datasets:
grid-ui.physik.uni-wuppertal.de:/sfsscratch/atlas/sandvoss/
data/atlas/software/mcgenerators/

[133] Object selection of the ATLAS top physics group:
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/
TopGroupCSCObjectSelection

[134] ATLAS Quality Assurance Group, *ATLAS C++ Coding Standard*, 2003
http://atlas-computing.web.cern.ch/atlas-computing/projects/
qa/AtlasCCS1_2_1.pdf

[135] ATLAS Improving Software wiki:
`https://twiki.cern.ch/twiki/bin/view/Atlas/ImprovingSoftware`

[136] Linux Kernel Coding Style:
`/usr/src/linux/Documentation/CodingStyle`

[137] Kernighan, Pike, *The Practice of Programming*, Addison-Wesley 1999

[138] The Athena package ImpactParameterTuning:
`http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/offline/PhysicsAnalysis/JetTagging/JetTagCorrections/ImpactParameterTuning`

[139] CERN ROOT webpage:
http://root.cern.ch/drupal/

[140] The Athena package DifferentialCrossSectionTTbar:
`http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/users/sandvoss/DifferentialCrossSectionTTbar`