TECHNISCHE UNIVERSITÄT MÜNCHEN

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Background suppression for a top quark mass measurement in the lepton+jets $t\bar{t}$ decay channel and Alignment of the ATLAS silicon detectors with cosmic rays

Tobias Göttfert

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. A. J. Buras
Prüfer der Dissertation:
1. Hon.-Prof. Dr. S. Bethke
2. Univ.-Prof. Dr. L. Oberauer

Die Dissertation wurde am 22.12.2009 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 21.01.2010 angenommen.

Physik-Department

Background suppression for a top quark mass measurement in the lepton+jets $t\bar{t}$ decay channel and Alignment of the ATLAS silicon detectors with cosmic rays

Dissertation
von
Tobias Göttfert

München
22. Dezember 2009

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Abstract

The investigation of top quark properties will be amongst the first measurements of observables of the Standard Model of particle physics at the Large Hadron Collider. This thesis deals with the suppression of background sources contributing to the event sample used for the determination of the top quark mass. Several techniques to reduce the contamination of the selected sample with events from W+jets production and combinatorial background from wrong jet associations are evaluated. The usage of the jet merging scales of a k_T jet algorithm as event shapes is laid out and a multivariate technique (Fisher discriminant) is applied to discriminate signal from physics background. Several kinematic variables are reviewed upon their capability to suppress wrong jet associations.

The second part presents the achievements on the alignment of the silicon part of the Inner Detector of the ATLAS experiment. A well-aligned tracking detector will be crucial for measurements that involve particle trajectories, e. g. for reliably identifying b-quark jets. Around 700,000 tracks from cosmic ray muons are used to infer the alignment of all silicon modules of ATLAS using the track-based local χ^2 alignment algorithm. Various additions to the method that deal with the peculiarities of alignment with cosmic rays are developed and presented. The achieved alignment precision is evaluated and compared to previous results.

Zusammenfassung

Eine der ersten Messungen von Observablen des Standardmodells der Teilchenphysik am Large Hadron Collider wird die Untersuchung von Topquark-Eigenschaften sein. Diese Arbeit behandelt die Untergrundunterdrückung von Ereignissen, die zur verwendeten Ereignisauswahl für die Topquark-Massenmessung beitragen. Mehrere Verfahren, um die Kontamination der ausgewählten Ereignisse mit Ereignissen aus W+jets-Produktion und kombinatorischem Untergrund zu reduzieren, werden untersucht. Die Jet-Merging-Skalen eines k_{T}-Algorithmus werden als topologische Variablen verwendet und multivariate Techniken (eine Fisher-Diskriminante) werden angewandt, um Signal von Physikuntergrund zu trennen. Weitere kinematische Variablen werden auf ihre Fähigkeit zur Unterdrückung falscher Jetzuordnungen untersucht.

Der zweite Teil dieser Arbeit stellt die Fortschritte beim Alignment der Siliziummodule des inneren Detektors von ATLAS vor. Ein gut alignierter Spurdetektor ist die Voraussetzung für die genaue Vermessung von Teilchenspuren, um zum Beispiel verlässlich b-Quark-induzierte Jets zu identifizieren. Etwa 700 000 Spuren aus kosmischen Myonen werden verwendet, um die Position aller Siliziummodule von ATLAS mittels des Local-χ^2-Alignmentalgorithmus zu bestimmen. Mehrere Erweiterungen der Methode, die sich mit den Besonderheiten des Alignments mit kosmischen Teilchen befassen, werden eingeführt und präsentiert. Die erreichte Alignmentpräzision wird bewertet und mit früheren Ergebnissen verglichen.

Contents

1 Introduction 1

2 Standard Model of particle physics 3

2.1 The Standard Model . . . 3

2.1.1 Quantum chromodynamics . . . 4

2.1.2 Electroweak interactions . . . 4

2.2 Top quark physics . . . 5

2.2.1 Production mechanisms . . . 6

2.2.2 Decay processes . . . 8

3 ATLAS at the LHC 9

3.1 The Large Hadron Collider . . . 9

3.2 The ATLAS detector . . . 11

3.2.1 ATLAS coordinates and terminology . . . 12

3.2.2 Inner Detector . . . 13

3.2.3 Calorimeter system . . . 16

3.2.4 Muon system . . . 17

3.2.5 Magnet system . . . 18

4 Systematic studies for a top quark mass commissioning analysis 19

4.1 Jet reconstruction . . . 19

4.1.1 Input objects and calibration . . . 20

4.1.2 Jet algorithms . . . 20

4.2 Cut-based top quark mass analysis . . . 22

4.2.1 Object and event selection . . . 23

4.2.2 Monte Carlo datasets . . . 24

4.2.3 Mass reconstruction 26

4.3 Usage of d_{Merge} as event shape variables for background rejection 29

4.3.1 Cut-based background rejection 30

4.3.2 Evaluation of multivariate discriminator techniques 34

4.4 Using d_{Merge} variables to suppress wrong combinations 42

4.5 Background discrimination using kinematic variables 45

4.6 Conclusions 54

5 Alignment studies for the silicon part of the ATLAS Inner Detector 57

5.1 The local χ^2 alignment algorithm 57

5.1.1 Track-based alignment 57

5.1.2 The local χ^2 formalism 58

5.1.3 Tracking and the choice of residuals 59

5.1.4 Levels of alignment granularity 60

5.2 Creation of realistic systematic deformations of the ATLAS Inner Detector 61

5.3 Local χ^2 alignment using cosmic ray data 67

5.3.1 Cosmic ray reconstruction and datasets 67

5.3.2 Local χ^2 alignment procedure 70

5.3.3 Quality of the final alignment 98

5.4 Conclusions 104

6 Conclusions 107

A Additional figures 109

B List of Abbreviations 112

List of Figures 113

List of Tables 116

Bibliography 117

Chapter 1

Introduction

The fundamental particles and interactions that make up our universe have been investigated for 55 years at the European Organization for Nuclear Research (CERN) located in Geneva, Switzerland. Since then, high energy physics has reached the region of subatomic particles and provided a wealth of insight on the structure of matter and the fundamental laws of nature. In a collaborative effort, the Large Hadron Collider (LHC) has been designed and built during the last 15 years to achieve an unprecedented collision energy of 14 TeV in its proton-proton collisions. Four main experiments were constructed, which aim to allow for physics analyses that test the Standard Model of particle physics and explore the regions where it is no longer valid. One of these experiments is ATLAS (A Toroidal Lhc ApparatuS).

After the successful beam injection in September 2008 and the following cryogenic incident, the accelerator and the experiments were preparing for a restart this autumn. The major points of work for ATLAS are the successful commissioning of all detector systems to take advantage of, and understand as much as possible, the performance of the experiment, and the preparation of physics analyses to be used for the early data that is now being recorded after the LHC restart.

Top-antitop quark pair production will already set in with a high rate at the early stages of proton-proton collisions. Top quark production and decay is an important topic, since it allows to test the Standard Model to high precision and additionally it appears as background process to many searches for physics beyond the Standard Model.

For a better reduction of backgrounds to top quark decay, the identification of b quark induced jets is of high importance, which can only be achieved with a well-commissioned tracking detector. The alignment of the ATLAS Inner Detector prior to proton-proton collisions is performed using muons which originate from interactions of cosmic ray particles with the atmosphere and that traverse the detector. Track-based alignment methods then can determine the positions in space of all modules to enable decent vertex and tracking resolution. Later, the Inner Detector measurements will be refined to their full resolution by alignment with tracks from proton-proton collisions.

This thesis is structured into the following chapters:

- Chapter 2 briefly reviews the Standard Model of particle physics and the properties of top quark production and decay.

- Chapter 3 introduces the ATLAS detector. The detector design and its physics requirements are illustrated and the individual subdetectors are described.

- Chapter 4 covers the results from systematic investigations on the main backgrounds for top quark mass reconstruction. The cut-based commissioning analysis that determines the top quark mass from the lepton+jets decay channel of top quark pair production is introduced. This is a robust analysis that only relies on basic functioning of the detector components and does not use b-quark identification. Event shape variables using the k_{T} jet algorithm are investigated upon their discrimination against background arising from W+jets physics processes and wrong jet association. A multivariate technique (Fisher discriminant) is introduced and its performance is evaluated. Several other kinematic variables are explored to improve the rejection of W+jets and combinatorial background.

- Chapter 5 presents the results on the alignment of the ATLAS silicon modules within the Inner Detector using cosmic muon data taken in 2008. The local χ^2 alignment algorithm is described and applied to data, together with the improvements that have been made to the software within this thesis. A study on creation of semi-realistic systematic deformations of the Inner Detector is presented. These deformations are used to test the alignment and physics performance of the ATLAS software.

- Chapter 6 presents the summary and conclusions of the thesis, with the important results highlighted and an outlook to further developments given.

Chapter 2

Standard Model of particle physics

2.1 The Standard Model

The Standard Model of particle physics [1] is a set of relativistic quantum field theories. It is at present the most precise description of elementary particle physics at the microscopic scale. It encompasses quantum chromodynamics (QCD) and electroweak theory with quantum electrodynamics (QED) and weak interactions. All components fulfil the requirements of being *renormalisable*, i. e. their divergences can be cancelled and they yield finite results for the physical observables. They are *gauge invariant*, meaning there are local gauge symmetries that make it possible to choose certain parameters without changing the dynamics of the theory. The total symmetry group of the Standard Model is the $SU(3) \otimes SU(2) \otimes U(1)$ group, which describes the symmetries of QCD and electroweak theory.

The particle content of the Standard Model is as follows: the fermions of spin 1/2 form the matter particles, whereas the spin 1 bosons are the exchange quanta of the force fields. In the fermion sector, there are six quark flavours, namely up, down, charm, strange, top, and bottom (u, d, c, s, t, b) quarks, which make three generations (or families) of doublets. Each of the quarks appears in three versions with different colour charge in addition to their electroweak charge. The three generations of leptons, namely, electron (e), muon (μ) and tauon (τ), together with their respective neutrinos, have no colour charge but electroweak charges. The force carrier particles are the massless photons (γ), the massive weak interaction gauge bosons $W^{\pm}$ and Z, and the eight massless gluons (g) that mediate the strong force. In addition, the favoured mechanism for breaking the electroweak symmetry necessitates that there exists at least one spin 0 boson, the Higgs boson (H), which is the only elementary particle of the Standard Model not yet observed. The fourth elementary force known, gravity, has negligible influence at the microscopic scale and is not included in the Standard Model.

All experimental data up to energies of a few hundred GeV strongly support the Standard Model calculations and make it one of the best-tested theories in physics up to now. However, several areas of the theory indicate that it cannot be valid up to energies higher than a few TeV and therefore strongly motivate the search for physics and theories beyond the Standard Model. Amongst the most promising candidates are supersymmetry, theories with extra dimensions and string theory, that strive for solving some or all of the problems of the Standard Model. To observe phenomena predicted by these models, a firm understanding of the Standard Model is needed.

2.1.1 Quantum chromodynamics

QCD [2, 3] is the part of the Standard Model describing strong interactions. It is based on the symmetry group $SU(3)$ with six different quark flavours being colour triplets and eight gluon fields (together called partons). Leptons and the gauge bosons of electroweak theory do not carry colour charges and thus do not participate in strong interactions.

Unlike the photon in QED, gluons carry colour charges themselves, allowing for self-coupling in 3-gluon and 4-gluon vertices. Therefore QCD is a non-Abelian theory, meaning that the generators of the algebra do not commute. Other striking features of QCD are *confinement* and *asymptotic freedom*, i. e. the impossibility of observing free quarks outside of bound hadron states and the asymptotic vanishing of the coupling for interactions with high momentum transfer (deep inelastic processes). The QCD coupling α_s varies as a function of the four momentum transfer Q, as given by (in next-to-leading order, NLO):

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\beta_0 \ln \frac{Q^2}{\mu^2}} \quad . \tag{2.1}$$

This equation gives the evolution of α_s from a known scale μ^2 to a different scale Q^2, with $\beta_0 > 0$ being the first term in the expansion of the β-function of the renormalisation group equation. It can be seen that for growing Q^2, α_s asymptotically vanishes, which describes the asymptotic freedom. For small Q^2, α_s eventually diverges, and the region of confinement is reached, where perturbative expansions cannot be done anymore. The current knowledge of α_s at the Z mass is $\alpha_s(M_{Z^0}) = 0.1184 \pm 0.0007$ [4].

Confinement also results in the observation of *jets* in hadron collisions, which are narrow streams of hadronic particles created in hard parton collisions. Jets arise due to the creation of new colourless quark-antiquark pairs from the vacuum when trying to separate bound quarks in a hard interaction.

2.1.2 Electroweak interactions

Electroweak theory [5–8] unifies weak interactions and electromagnetism. Its Lagrangian obeys the gauge group $SU(2) \otimes U(1)$. It is a chiral theory in the sense that it affects right-handed and left-handed fields differently. All right-handed fermionic fields are electroweak singlets, whereas the left-handed fields are doublets. This forbids mass terms for the fermions, and they are reintroduced into the Standard Model together with $W^\pm$ and Z masses by the mechanism of *electroweak symmetry breaking.* The standard way to introduce the symmetry breaking is via the *Higgs mechanism* [9]. It does so by spontaneously breaking the symmetry group via a doublet of complex scalar fields Φ with a non-vanishing vacuum expectation value ($v \approx 246\,\mathrm{GeV}$). Three of the four degrees of freedom of Φ are absorbed into the longitudinal degree of freedom of massive spin-1 bosons (of the $W^\pm$ and Z), while the photon stays massless. The remaining degree of freedom is physical and should be found as the Higgs boson. A huge experimental effort is undertaken to find this last missing piece in the Standard Model, especially at the various LHC experiments.

Weak interactions are the only Standard Model interactions that change flavour. The weak eigenstates of the fermions, where W bosons couple to, are not the eigenstates of the freely propagating particles, the mass eigenstates. The mixing of the down-type quark mass eigenstates into the weak eigenstates is parameterised in the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix [10, 11]. The mixing of the neutrino mass eigenstates into their weak eigenstates is parameterised in the 3×3 Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [12, 13]. The weak and mass eigenstates for up-type quarks and charged leptons are chosen to be identical.

2.2 Top quark physics

The top quark plays an important role in high energy physics [14]. It is by far the heaviest observed elementary particle. The discovery of its existence was made in 1995 by the CDF and D0 collaborations [15, 16] at the TeVatron collider at Fermilab. The current combined measurements of the two detectors [17] result in a top quark mass of $173.1 \pm 1.3\,\mathrm{GeV}$[1].

The top quark carries a number of interesting properties, which make it special amongst the Standard Model quarks:

- Due to its short lifetime, the top quark is the only quark that decays as a bare quark, i.e. before it can form bound states. The top decay width at NLO is calculable to be $\Gamma_t = 1.36\,\mathrm{GeV}$ [18]. This corresponds to a lifetime of about $0.5 \cdot 10^{-24}\,\mathrm{s}$, which is too short to observe bound states involving top quarks. This is also not enough time for chromomagnetic spin depolarisation, and the top quark passes on its spin to its decay particles.
- The Yukawa coupling of the top quark, i. e. the coupling to the Higgs boson, as given by $\sqrt{2}m_t/v$, is almost unity, which is already interesting on its own. Since the top quark has the largest Yukawa coupling of all quarks, it will play an important role in understanding the mechanism of electroweak symmetry breaking.
- Quantum electroweak theory links the mass of the Higgs boson with the masses of the W boson and the top quark via virtual loop corrections. This gives the possibility to constrain the Higgs boson mass by measuring the two other masses at high precision.

Some properties of the top quark are already known quite precisely, most notably the mass, while many others are still unmeasured or only known vaguely. For example, the electric charge of the top quark has not yet been measured, which leaves the possibility that the observed particle is in fact some exotic type of quark having an electric charge other than 2/3. Recent analyses however exclude this scenario at 92 % CL [19]. At the LHC, the charge of the top quark can be measured with an expected precision of $\mathcal{O}(10\,\%)$ [14].

Finally, a good understanding of top quark properties and production and decay mechanisms is also essential for other physics at the ATLAS detector: top quark production and decay processes can serve as important data to calibrate and commission the detector. Moreover, almost all analyses that search for physics beyond the Standard Model have to cope with top quark production and decay as background processes.

---

[1] For the whole of this thesis *natural units* will be used, i. e. $\hbar = c = 1$

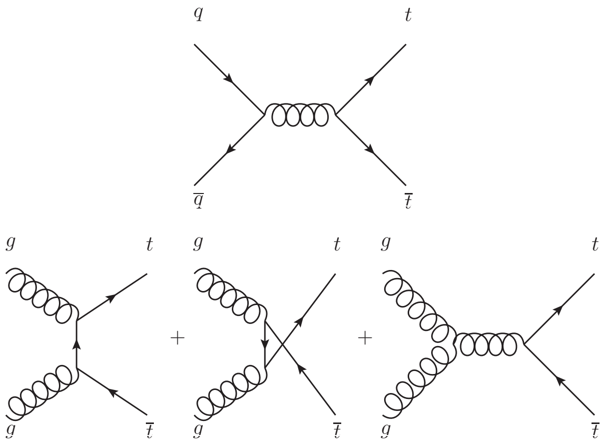


Figure 2.1: *The lowest-order Feynman diagrams contributing to $t\bar{t}$ production. These are $q\bar{q}$-annihilation (upper) and gluon-gluon-fusion (lower).*

2.2.1 Production mechanisms

The production mechanisms of top quarks at hadron colliders are twofold: There are top-antitop pair production and single top quark production mechanisms. The Feynman diagrams contributing to $t\bar{t}$ production at lowest order are shown in figure 2.1. These are $q\bar{q}$-annihilation and gluon-gluon-fusion. The Feynman diagrams contributing to single top quark production at lowest order are shown in figure 2.2. These are s-channel production, two t-channel processes and the associated production of a top quark with a W boson.

The cross sections for these processes can be calculated by factorizing them into the hard scattering process of the two partons and the parton longitudinal momentum distribution functions (PDFs) in the incoming protons. The matrix element of the parton-parton interaction can be calculated and is denoted with $\hat{\sigma}^{i,j}$ for two partons of type i and j. The PDFs $f_i(x_i, \mu_F^2)$ are not calculable from the theory, but rather are the results of parameterised fits to data from deep-inelastic scattering and other experiments. The total cross section for top quark pair production is then given by

$$\sigma_{t\bar{t}}(\sqrt{s}, m_t^2, \mu_r^2, \mu_F^2) = \sum_{i,j=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \cdot \hat{\sigma}^{ij \to t\bar{t}}(\sqrt{s}, m_t^2, x_i, x_j, \mu_r^2, \mu_F^2) \quad , \tag{2.2}$$

where $\sqrt{s}$ denotes the centre-of-mass energy of the partonic collision, x_i and x_j are the momentum fractions of the respective protons that the partons i and j carry. The symbol μ_r denotes the renormalisation scale at which the matrix element calculation is performed, and μ_F is the factorisation scale at which the parton density functions are evaluated. A usual choice is to set $\mu_r = \mu_F = m_t$.

Single top quark production is calculated to have about 50 % of the cross section of top quark pair production. A comparison of calculated production cross sections at LHC energies is given in table 2.1. From the cross sections one can see that at an integrated luminosity of 100 fb^{-1}, which is the design luminosity per year expected for the LHC experiments, top quarks will be produced in high abundance (about 100 million pairs). This means that every measurement of top quark properties will soon be dominated by systematic uncertainties.

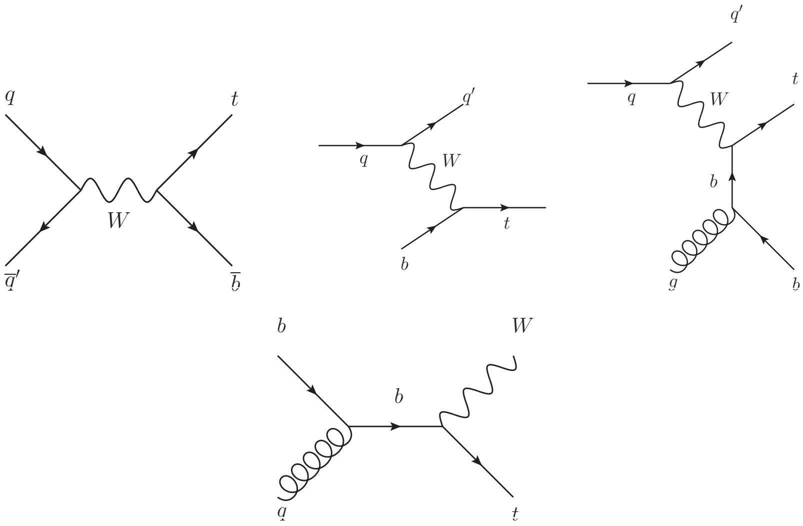


Figure 2.2: *The lowest-order Feynman diagrams contributing to single top quark production. These are s-channel (upper left), t-channel (upper middle and right) and Wt-channel (lower).*

Table 2.1: *Calculated production cross sections for top quark single and pair production at various LHC energies. The $t\bar{t}$ calculations are done in approximate NNLO with the PDF parametrisation CTEQ6.6 and an assumed top quark mass of $m_t = 172.5$ GeV [20]. For the single top quark processes at $\sqrt{s} = 10$ TeV, the program MCFM was used to calculate the cross sections to NLO using a top quark mass of $m_t = 172.5$ GeV [21]. The single top quark cross sections at $\sqrt{s} = 14$ TeV are calculated to NLO for the s- and t-channel [22] and for the Wt-channel [21], assuming $m_t = 175$ GeV. These cross sections are also used as reference for the ATLAS Monte Carlo production of top quark samples.*

top quark pair production

$\sqrt{s}$	$\sigma_{t\bar{t}}$
10 TeV	401.60 pb
14 TeV	883.90 pb

single top quark production

$\sqrt{s}$	$\sigma_{s-channel}$	$\sigma_{t-channel}$	$\sigma_{Wt-channel}$
10 TeV	6.627 pb	124.51 pb	32.66 pb
14 TeV	10.65 pb	246.6 pb	66 pb

After the discovery of top quark pair production in 1995, single top quark production was also discovered recently at the TeVatron detectors [23, 24].

2.2.2 Decay processes

The top quark decays almost exclusively to a W boson and a b quark, as reflected by the large V_{tb} matrix element in the CKM matrix of almost unity. The current best estimates of the CKM matrix elements involving top quarks, imposing the unitarity of the matrix, are given in the following [18]. These lead to the predicted branching fractions $Br(t \to Wq)$ that favour $t \to Wb$ over all other decays:

$$\begin{aligned}
|V_{td}| &= 0.0087^{+3.0\,\%}_{-4.2\,\%} & \Rightarrow \quad & Br(t \to Wd) \approx 0.01\% \\
|V_{ts}| &= 0.0407 \pm 2.5\,\% & \Rightarrow \quad & Br(t \to Ws) \approx 0.17\% \\
|V_{tb}| &= 0.9991 \pm 0.004\,\% & \Rightarrow \quad & Br(t \to Wb) \approx 99.82\%
\end{aligned}$$

The W boson then further decays either leptonically into $e\nu_e$, $\mu\nu_\mu$ or $\tau\nu_\tau$ or hadronically into $q\bar{q}'$ with $q = u, d, s, c$. This leaves three main decay channels for $t\bar{t}$ pairs, characterised by the number of leptons and jets they produce. They are called di-lepton channel (ll), all-hadronic channel, and lepton+jets channel (l+jets), and their respective branching fractions are given in table 2.2.

Table 2.2: *The expected branching fractions for the various decay channels of top quark pairs, calculated from the measured W decay fractions [18]. In this table, $\hat{l}$ stands for e, μ.*

	decay mode	Br
di-lepton	$t\bar{t} \to \hat{l}\nu_{\hat{l}} b \hat{l}' \nu_{\hat{l}'} b$	4.5%
	$t\bar{t} \to \hat{l}\nu_{\hat{l}} b \tau \nu_\tau b$	4.8%
	$t\bar{t} \to \tau\nu_\tau b \tau' \nu_{\tau'} b$	1.3%
all-hadronic	$t\bar{t} \to q q' b q'' q''' b$	45.7%
lepton+jets	$t\bar{t} \to \hat{l}\nu_{\hat{l}} b q q' b$	28.8%
	$t\bar{t} \to \tau\nu_\tau b q q' b$	15.2%

The analysis described in this thesis focuses on the l+jets channel, since this channel offers a good compromise between high branching fraction and good experimental reconstruction properties. The ll channel only offers 10.6 % of all decays, while possessing the additional complication of two neutrinos in the final state. This makes kinematic reconstruction of the decay products ambiguous. The all-hadronic channel offers 45.7 % branching fraction, while making it necessary to correctly assign the six jets in the final state to the two decaying top quarks. This leaves at least 10 combinations per event to consider (assuming no b-tagging). The l+jets channel possesses 44.0 % branching fraction and has only one neutrino in the final state. The leptonically decaying top quark can be used for tagging the event.The mass of the top quark can then be extracted from a full reconstruction of the hadronically decaying top. Since τ leptons are hard to reconstruct, one often restricts the analysis to electrons and muons for leptonically decaying top quarks. In this case, the l+jets channel still offers a branching fraction of 28.8 %.

Chapter 3

ATLAS at the LHC

3.1 The Large Hadron Collider

The Large Hadron Collider [25] is the large proton-proton ring collider located at the CERN site. It is designed to deliver unprecedented collision energy and luminosity to the four main experiments: ALICE [26], CMS [27], LHCb [28] and ATLAS [29]. The design centre-of-mass energy of the LHC is $\sqrt{s}$ =14 TeV with two beams of 7 TeV protons colliding. After the magnet accident in September 2008 [30], it was decided that the LHC will initially operate at 3.5 TeV proton energy at the startup in autumn 2009. This lowers the demands on the magnetic field strengths of the dipole magnets, of which not all are commissioned to the final 7 TeV proton energy yet.

The event rate for a given process is directly proportional to the luminosity of the machine and the cross section of the process:

$$N = \mathscr{L}\sigma(\sqrt{s}) \quad . \tag{3.1}$$

The luminosity depends only on machine parameters and can be calculated in the assumption of Gaussian beam profiles as:

$$\mathscr{L} = \frac{fn^2}{4\pi\sigma_x\sigma_y} \quad , \tag{3.2}$$

with f being the collision frequency, n the number of protons per bunch and σ_x, σ_y the transverse beam widths.

At the design luminosity of $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the LHC ring will contain 2×2808 proton bunches with $11.5 \cdot 10^{10}$ protons per bunch at a collision frequency of $f = 40\,\mathrm{MHz}$. During the startup phase, the luminosity will be considerably lower, in the range of $(10^{30} - 10^{32})\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The expected integrated luminosity within the first year of LHC running is around $100\,\mathrm{pb}^{-1}$. The startup phase will be followed by a low-luminosity phase, which will provide $\mathscr{L} \approx 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In this phase, one can expect around $10\,\mathrm{fb}^{-1}$ per year. Finally, the design luminosity will deliver $100\,\mathrm{fb}^{-1}$ annually.

The LHC can also accelerate heavy ions. It is planned to have lead-lead collision runs with up to $\mathscr{L} = 10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which will be investigated mainly by ALICE, but also by the other experiments.

A sketch of the LHC ring with the interaction points can be seen in figure 3.1.

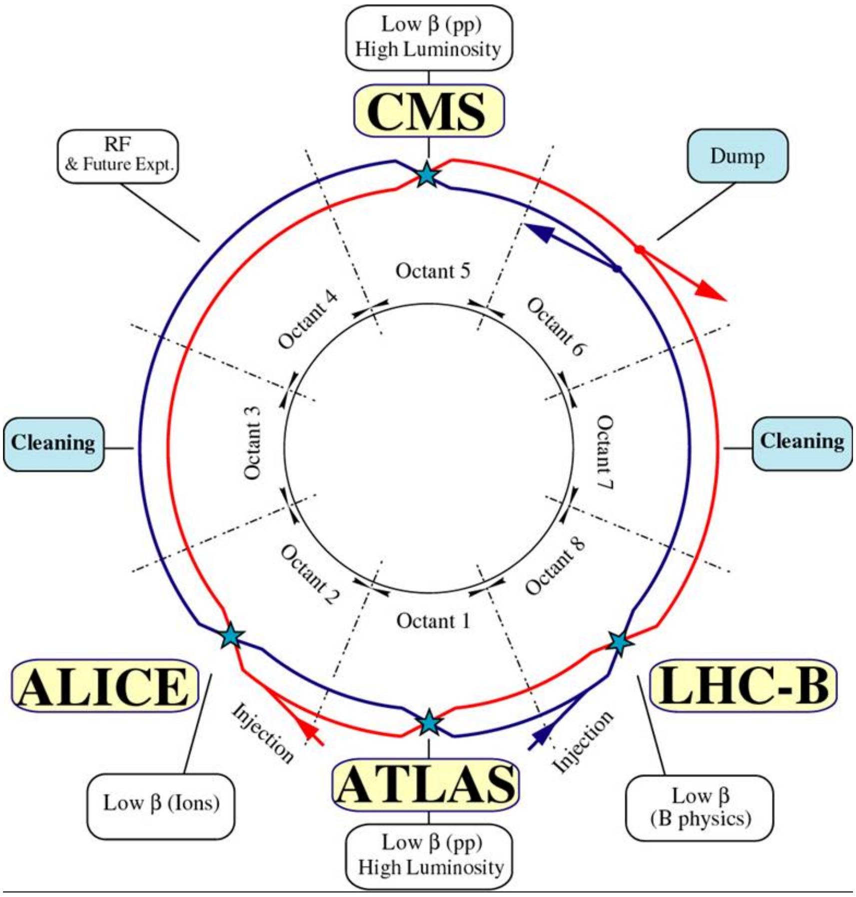


Figure 3.1: *A sketch of the LHC ring with the different collision points and the experiments.*

3.2 The ATLAS detector

The ATLAS detector [29] is one of the two big multi-purpose detectors for the LHC particle collisions. Its design has to fulfil several conditions to make optimal use of the physics potential that the LHC accelerator offers. Amongst the primary requirements are:

- a radiation hard design, especially for the Inner Detector, and the ability to cope with the high occupancies and trigger rates when operated at the high luminosity and bunch crossing rate that the LHC delivers,
- an acceptance as close as possible to a full 4π solid angle to detect all collision products, even when the centre-of-mass system of the hard interaction is highly boosted, and to reliably deduce missing transverse energy,
- very good momentum resolution and secondary vertex determination in the Inner Detector for a precise track reconstruction and a good b-quark tagging efficiency,
- calorimeters with very good energy resolution and coverage plus small energy leakage to measure jets and missing energy precisely,
- a good identification and momentum resolution for electrons and muons and the ability to identify the sign of charge for very high momentum particles,
- a fast and efficient trigger system to get good sensitivity for many of the interesting physics processes.

In order to fulfil all these requirements, the following overall design has been chosen, which will be described in more detail in the following sections:

- an Inner Detector comprising three subsystems with different technologies:
  - The Pixel detector employs radiation hard silicon pixel sensors to reach high granularity and good vertex resolution around the interaction point.
  - The SemiConductor Tracker (SCT) employs silicon strip sensors in double layered modules to provide precise spacepoint measurements of the tracks outside the Pixel detector volume.
  - The Transition Radiation Tracker (TRT) as the outermost of the three uses many layers of kapton tubes with gold wires, interspersed with plastic foils and fibres. It is designed to provide high momentum resolution. In addition, particle identification via the detection of transition radiation photons is made possible.
- a calorimeter system with high-granularity sampling calorimeters made of different absorbers and active materials
- a muon system extending to large radii within a toroidal magnetic field
- a magnet system with two different orientations for the magnetic field:
  - a superconducting solenoidal magnet enclosing the Inner Detector and producing a field of about 2 T

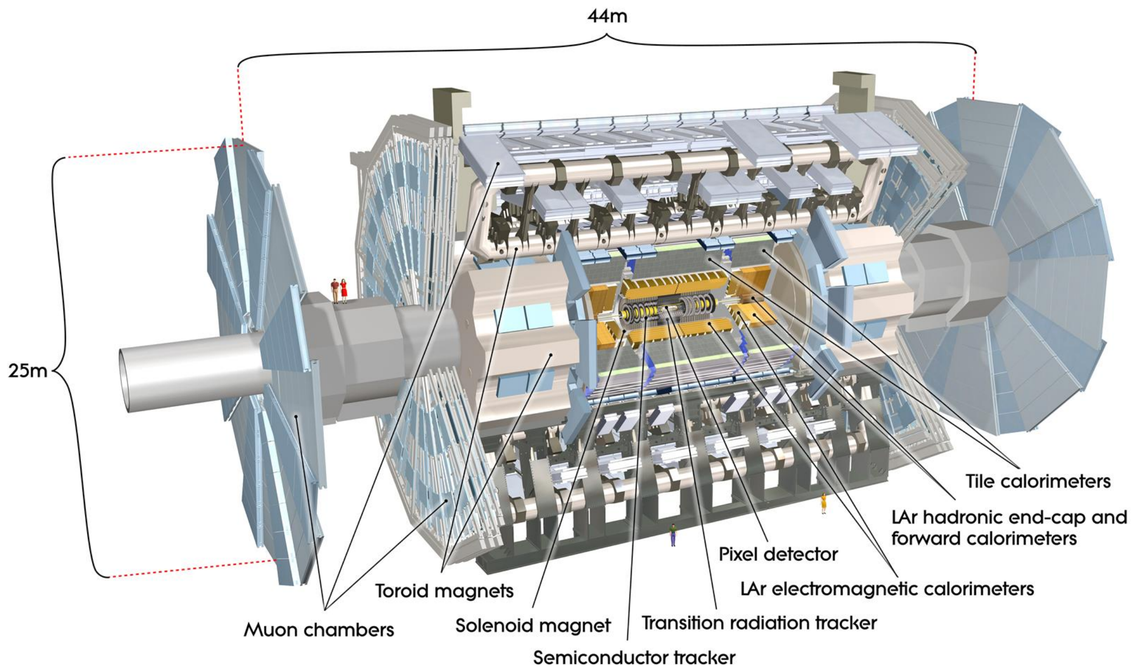


Figure 3.2: *Computer generated sketch of ATLAS with its subsystems and dimensions labelled [31].*

  - eight superconducting large toroidal coils with air core technology complemented by two endcap toroidal coils providing bending power within the huge volume of the muon system.

Figure 3.2 shows a computer generated sketch of ATLAS with the different subsystems indicated. The persons on the figure give an impression of the large size of the detector.

3.2.1 ATLAS coordinates and terminology

Throughout this thesis, the following sets of coordinates will be used:

- The ATLAS *global coordinate system* is a right-handed orthogonal system, where the Z-axis follows the beamline and the X-Y-plane is perpendicular to it. The X-axis is defined to point towards the centre of the LHC ring, the Y-axis points upwards. The side of the detector with positive Z is defined to contain endcap A; endcap C lies in the negative Z side.

- Often cylindrical coordinates are used, which have the form (R, ϕ, θ), where R is the distance from the Z-axis, ϕ is the azimuthal angle and θ the polar angle measured from the positive Z-axis. The pseudorapidity η is defined as $\eta = -\ln\tan(\theta/2)$. Distances in angular space are usually given in terms of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, which is invariant under Lorentz boosts along the Z-axis.

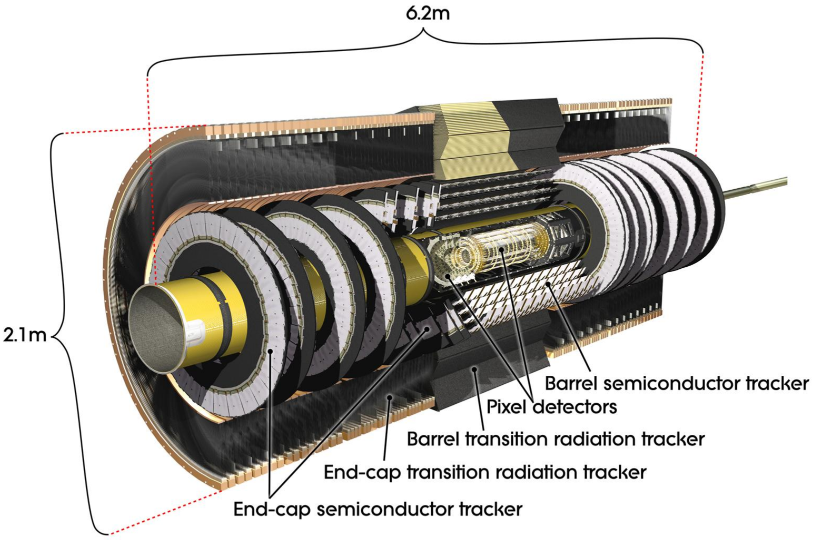


Figure 3.3: *A computer generated sketch of the ATLAS Inner detector [31].*

- Every detector element within ATLAS has a per-module *local coordinate system*, which is also righthanded and orthogonal. These coordinates will be denoted with (x, y, z). Additionally, for alignment rotations around these axes will frequently be determined. The rotations around the local x, y, z axes are called α, β, γ, respectively.

3.2.2 Inner Detector

The Inner Detector is used to accurately reconstruct tracks of charged particles and their momenta and charge sign, as well as the primary and possible secondary vertices in the collision. A sketch of the three subsystems can be seen in figure 3.3. They are described in the following, starting from the inside:

Pixel detector

The Pixel detector consists of 1744 rectangular modules of size $19 \times 63\,\mathrm{mm}^2$, of which 1456 are located in the three barrel layers and 2×144 in the 2×3 endcap disks. The pixel detector dimensions are 1.3 m in length and 150 mm in radius, with the innermost layer at a radius of 50.5 mm away from the beamline.

The barrel is composed of *staves* of pixel modules, where 13 modules are mounted onto a support structure. The middle module lies parallel to the beamline, whereas the other modules are shingled slightly to create overlaps in the Z direction. The staves are mounted as bi-staves with a common cooling unit to the barrel such that the modules are inclined by 20° with respect to the tangent of the layer in the R-Φ plane (turbine arrangement). This creates overlap and therefore full coverage in R-Φ and compensates for the mean of the Lorentz drift angle in the magnetic field of the Inner Detector. The endcaps possess one ring of pixel modules each, where the modules are parallel to the R-Φ plane and to each other and overlap by alternating frontside and backside mounting on the support disk.

All pixel modules are identical and possess 47232 rectangular pixels with the usual size of $50 \times 400\,\mu\mathrm{m}^2$. They are read out by 16 front-end (FE) chips in two rows and eight columns, which need a small gap between each other. Thus, 4 pixels in each FE chip column are ganged together, i. e. are read out by the same channel on the FE chip. This leads to 46080 readout channels in total on a pixel module. The sensor is made of $250\,\mu\mathrm{m}$ thick n^+-in-n silicon, where the pixels are bump-bonded to the respective readout channels of the FE chips.

The use of highly oxygenated silicon and the n^+-in-n design allow the pixel modules to be operated at the very high radiation fluences close to the beam. To slow down the radiation damage and control the leakage currents, the pixel detector will be operated at a temperature of $-5\,^\circ\mathrm{C}$ to $-10\,^\circ\mathrm{C}$. The n-type silicon bulk will type-invert into p-type material after a fluence of $\Phi_{\mathrm{eq}}(1\,\mathrm{MeV}$ neutron equivalent$) \approx 2\cdot 10^{13}\,\mathrm{cm}^{-2}$; from then on the depletion zone will grow starting from the n^+ pixel implants and maintain a good charge collection efficiency at low bias voltages. However, the performance of the innermost pixel layer is expected to strongly deteriorate after roughly three years of running at design luminosity.

The pixel detector delivers three space points per track and is the most important detector for a precise vertex, secondary vertex and pileup vertices determination. The design goals put stringent requirements on the spatial resolution and thus the allowed alignment tolerances, which can eventually only be reached by track-based alignment. The required precisions are listed in table 3.1. They are derived from the demand that the detector misalignment must not reduce the track parameter resolutions by more than 20 % [32]. In addition, the estimated as-built accuracy as well as the intrinsic resolution of the sensors are listed. The high as-built accuracy for pixel endcap modules comes from a precise survey of the assembled endcap disks. However, it only measures relative module-to-module positions and not absolute positions in space so that a precise alignment is still needed. The other as-built precisions stem from the mounting precisions of the modules on the support structures and (much smaller) the module fabrication tolerances. These numbers can only give estimates of the positioning precision, since the module locations might well change by assembling the individual subdetectors to the full Inner Detector, lowering them into the experimental cavern and cooling the detector to its operating temperature.

The local coordinate system for every pixel module is as follows: The origin of the coordinate system is in the centre of the rectangular module, the x-coordinate runs along the short side of the pixels, the y-coordinate runs along the long side of the pixels. The local z-coordinate is normal to the plane of the module and pointing away from the interaction point.

SemiConductor Tracker

The SCT consists of 4088 modules with single-sided silicon strip sensors and two sensor layers per module. The 2112 barrel modules have rectangular shape with a size of $12.8 \times 6.3\,\mathrm{cm}^2$ and are arranged in 4 layers of modules such that a particle on average crosses

Table 3.1: *The required [29] and the as-built [29, 33, 34] precisions of the various Inner Detector module types. In addition, the intrinsic accuracy of the sensors is listed [29].*

Pixel

	barrel		endcap		
coordinate	required	as-built	required	as-built	intrinsic
Radial (R)	20 µm	50 µm	20 µm	12.7 µm	115 µm (endcap)
Axial (z)	20 µm	20 µm	100 µm	4.7 µm	115 µm (barrel)
Azimuth (ϕ)	7 µm	50 µm	7 µm	4.6 µm	10 µm

SCT

	barrel		endcap		
coordinate	required	as-built	required	as-built	intrinsic
Radial (R)	100 µm		50 µm		580 µm (endcap)
Axial (z)	50 µm	$\mathcal{O}(100\,\mu\text{m})$	200 µm	$\mathcal{O}(100\,\mu\text{m})$	580 µm (barrel)
Azimuth (ϕ)	12 µm		12 µm		17 µm

TRT

	barrel		endcap		
coordinate	required	as-built	required	as-built	intrinsic
Azimuth (ϕ)	30 µm	≈ 50 µm	30 µm	≈ 50 µm	130 µm

8 strip layers and yields 4 space points. The strips have a strip pitch of 80 µm. On one side of the module, they are parallel to the beam direction and each other; on the other side they are rotated by a stereo angle of 40 mrad with respect to the first side. The modules are mounted individually on the barrel layer structures with a inclination angle of 10°. The barrel consists of 12 rings of modules along the Z-direction, where the modules are parallel to the Z-axis and overlap by mounting them at alternating heights on the support.

In the endcap, 1976 wedge-shaped modules with radially running strips on one side of the module are mounted onto 2 × 9 endcap disks. Their other side's wafers are also rotated by a stereo angle of 40 mrad. Due to their radial orientation, the strips have a fan structure; their average pitch is as well 80 µm. The modules are mounted onto the disk in two or three rings (inner, middle, outer), where the middle ring is attached to the other side of the support disk compared to inner and outer ring. By this, the modules overlap in the R-direction. Overlap in the Φ-direction is created by mounting at alternate heights on the support.

The modules are made of 285 µm thick p-in-n silicon wafer material. Usually two wafers per side are daisy-chained with wire bonds and glued to a graphite support structure that provides mechanical stability and heat transport. In the endcaps, some modules only have one wafer per side to ensure that particles from the interaction point cross 4 SCT layers while avoiding the regions of too high radiation at high η. Each wafer possesses 768 AC-coupled strips. The 1536 readout channels per module are handled by 12 ASIC chips, which perform readout of the strips and signal transmission to the data acquisition system.

As for the pixel modules, the provided and required space point resolutions are listed in table 3.1. The centre of the local coordinate system in the SCT modules lies in the geometrical centre of the module, i. e. in the support structure between the two wafer planes. It has its y-axis along the strips of one wafer side and the x-axis perpendicular to them in the wafer plane. The z-coordinate is normal to the module, pointing away from the interaction point.

Transition Radiation Tracker

The TRT consists of about 351000 drift tubes (straws) of radius 2 mm filled with a xenon-based gas mixture and an anode wire made of gold plated tungsten. It is designed to deliver a high number of track hits (on average 36 per track) and thus a precise momentum estimation in the outer regions of the Inner Detector. The total dimensions of the TRT barrel are 2.2 m in diameter and 1.6 m in length; the two endcaps have a diameter of 2.2 m and a length of 1.9 m each. The TRT straws are arranged parallel to the beam direction in the barrel part and radially on wheels for the endcap. Thus, only R-ϕ-information can be obtained by the TRT. Additionally, plastic foils and fibres are interspersed between the straw layers to provoke the emission of transition radiation of high-β particles and thus make it possible to distinguish between electrons and pions. The spatial precision which is required and provided is listed in table 3.1. For the TRT, the determination of the wire positions has to be accompanied also by a precise calibration of the R-t-relations of the ion drift in the filling gas.

ATLAS readout identifier scheme for the Inner Detector

In the ATLAS offline software, a scheme of identifiers for the different readout structures allows to describe and access specific parts of the detector. The identifier is a set of numbers separated by "/" and is structured as:

Identifier = Subsystem / Subdetector / Barrel or Endcap / Layer or Disk / Phi / Eta

The identifier is not restricted to the Inner Detector, but will in this thesis only be used in conjunction with the alignment of silicon Inner Detector structures (c. f. chapter 5). The subsystem value is 2 for the Inner Detector. The subdetector value can be 1 (Pixel) or 2 (SCT). The third position can contain 0 for the barrel or +2 and −2 for endcap A and C, respectively. For the fourth position, the three pixel layers and three pixel endcap disks are labeled from 0 to 2, the four SCT layers from 0 to 3. The SCT endcap disks are labeled from 0 to 8. The range of the fifth and sixth number (azimuthal and polar identifier) depends on the particular layer or disk addressed. This form of identifier describes an Inner Detector readout module and can be extended by two more positions to describe individual channels on a module.

3.2.3 Calorimeter system

The calorimeter system consists of various sampling calorimeters and covers pseudorapidities of $|\eta| < 4.9$. The calorimeters are split according to the different techniques and granularity required for the different physics processes of interest and the different radiation environments.

The electromagnetic calorimeter consists of a barrel part, two endcap calorimeters (EMEC) and two very forward electromagnetic calorimeters (FCal1). The barrel part and

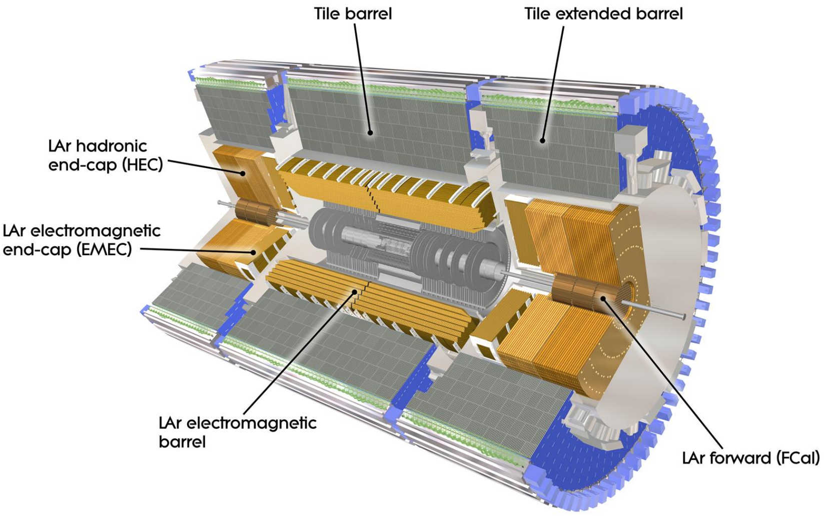


Figure 3.4: *A computer generated sketch of the ATLAS calorimeter system [31].*

the EMECs have a finely segmented accordeon structure with lead absorbers and liquid argon (LAr) as active material. Over the pseudorapidity range of the Inner Detector ($|\eta| < 2.5$), they have a high granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ or better to ensure excellent electron and photon reconstruction. The EMECs extend up to $|\eta| = 3.2$ with coarser granularity ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$). The FCal1 calorimeters cover $3.1 < |\eta| < 4.8$.

The hadronic calorimeter consists of the tile barrel, two tile extended barrels, the two hadronic endcaps (HEC) and the four forward (FCal2, FCal3) calorimeters. The tile barrel and extended barrels cover the regions $|\eta| < 1.0$ and $0.8 < |\eta| < 1.7$, respectively, with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. They use steel as absorber and scintillating tiles made of polystyrene as active material. The HEC calorimeters consist of two independent wheels per endcap, made of copper absorbers with LAr as active material. They span pseudorapidity ranges of $1.5 < |\eta| < 3.1$. The FCal2 and FCal3 calorimeters are integrated within the same cryostat as the HEC and consist of two tungsten modules with LAr as active material. They span ranges of $3.2 < |\eta| < 4.8$.

In total, the calorimeter system provides about 192000 readout channels. A sketch of the calorimeters can be seen in figure 3.4.

3.2.4 Muon system

The ATLAS muon system covers the pseudorapidity region of $|\eta| < 2.7$ with four different types of muon detectors. The Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC) are used for precise measurements of muon tracks and their momenta. The MDTs are aluminium tubes with a diameter of 30 mm. They are filled with an $\mathrm{Ar/CO_2}$ gas mixture and possess an anode wire made of tungsten-rhenium. The CSCs are multiwire proportional chambers with higher granularity to withstand high muon rates and background conditions. They are used in the innermost plane for large pseudorapidities ($2.0 < |\eta| < 2.7$). The Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC)

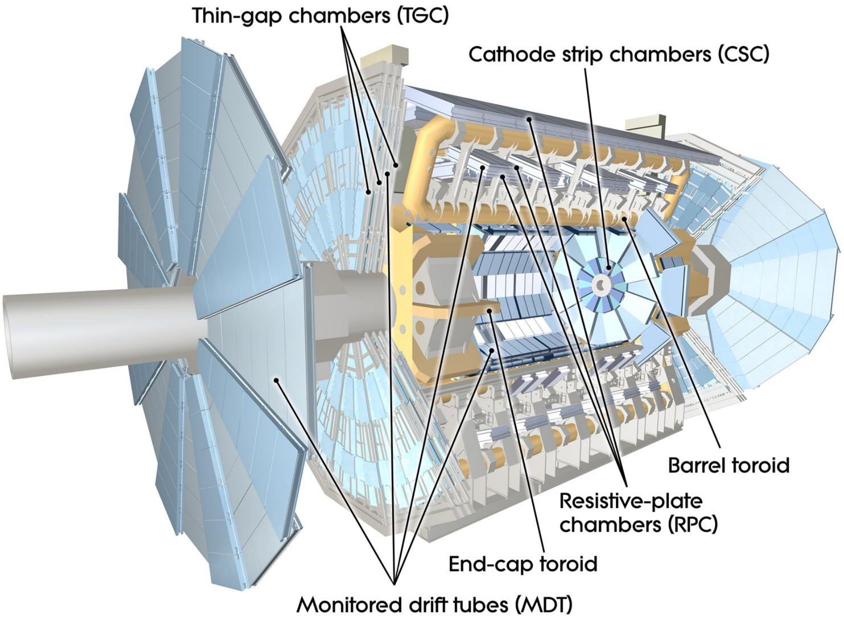


Figure 3.5: *A computer generated sketch of the ATLAS muon system [31].*

are designed for fast triggering of muons and are located between the barrel chambers or endcap chambers, respectively. Additionally, they measure muon coordinates perpendicular to the direction determined by the precision chambers. The muon trigger system covers an η-range of $|\eta| < 2.4$.

A sketch of the different types of muon detectors can be found in figure 3.5.

3.2.5 Magnet system

The magnet system is providing the necessary bending power to reconstruct the transverse momenta of charged particles in the tracking detectors. The central solenoid [35] provides a 2 T axial field in the Inner Detector volume. It covers an inner diameter of 2.5 m and a length of 5.8 m and consists of a steel barrel equipped with Al-NbTi superconducting coil windings. The magnetic flux is returned through the steel of the hadronic calorimeter.

The toroid magnets are split into 8 superconducting coils for the barrel air core toroid [36] and two endcap toroids [37] in separate cryostats. The Al-NbTiCu coil windings are housed in racetrack-shaped stainless steel cryostats. The barrel toroid makes up the impressive size of ATLAS, spanning 20.1 m in outer diameter and 25.3 m in length. The magnets deliver magnetic field strengths of up to 3.5 T in the volume covered by the muon chambers. The highly non-uniform field configuration that the toroid magnets produce has to be mapped precisely to allow for a proper reconstruction of muon trajectories.

Chapter 4

Systematic studies for a top quark mass commissioning analysis

The main part of this thesis covers the application of an early top quark analysis on simulated ATLAS data, and the search for and investigation of additional requirements that could further enhance the signal purity in those samples. Since at the beginning the calibration of the detector will be not final, robust methods are needed that do not rely on the ultimate detector resolution. For the early data scenario, a cut-based mass measurement without using *b*-quark identification methods is foreseen and termed "commissioning analysis".

First, some aspects of jet reconstruction at ATLAS are highlighted, which play an important role in the following analyses. The exclusive k_T jet algorithm has promising features and its jet merging scales can be used as event shape variables. These are observables which quantify the topology of one event by incorporating several kinematic variables into a single quantity. Afterwards, the current cut-based top quark mass analysis is presented. The usefulness of the event shape variables from the k_T algorithm to discriminate between signal and background events is examined. A multivariate analysis technique, the Fisher discriminant method, is applied and the effects on the top quark mass analysis are discussed. Then possibilities to handle combinatorial background, i. e. misreconstruction resulting from wrong jet association, with these variables is discussed. Finally, other kinematic variables and their discrimination power against physics and combinatorial background are investigated.

4.1 Jet reconstruction

Jet algorithms are used to group the objects stemming from a hard particle interaction into jets. They are supposed to provide a mapping between the hard partons emerging from the interaction, the hadrons evolving from those, and their energy deposits that are measured in the calorimeter. A jet algorithm therefore needs to be applicable to all these inputs from theory and experiment. During event simulation, one can reconstruct jets using the partons that evolved from a hard interaction matrix element, or the stable hadrons that appear after parton showering and hadronisation. After detector simulation or in real data taking, the energy deposits in calorimeter cells and clusters can be grouped to reconstructed jets. In practice, a jet algorithm takes a list of four-vectors as input to the jet making procedure.

4.1.1 Input objects and calibration

For reconstructing jets from detector measurements, four-vectors derived from calorimeter information are used. In ATLAS, calorimeter clusters are either reconstructed as *projective towers* with constant widths in η-ϕ or as *topological clusters* as the result of a clustering procedure of the individual calorimeter cells [38]. In the latter case, a cluster is extended with neighbouring calorimeter cells around a seed cell with large signal. This procedure is stopped once the neighbouring clusters fall below an energy threshold.

There are also different ways to achieve a proper hadronic energy calibration for the calorimeter clusters. In the case of a purely electromagnetic shower, e. g. from an electron, the calorimeter response stays proportional to the particle energy. Thus, a fixed proportionality constant can be derived for an electromagnetic calibration [39]. In the case of showers with a hadronic component, several effects contribute to a reduction of reconstructable energy in the calorimeter, most notably hadronic leakage and invisible energy. Hadronic leakage happens when the hadronic shower is not fully contained in the calorimeter and some particles escape into the muon system. The invisible energy is composed of energy deposited in the insensitive absorber layers, energy used for breaking of nuclear bindings, and a fraction of neutron energy that cannot be recovered. These effects make more involved calibration techniques necessary.

The two methods of hadronic calibration which are used in ATLAS are *global calibration* in the way of “H1-style cell weighting” and *local calibration.* The former tries to calibrate the final reconstructed jets to the energy of the constituent particles by comparing to jets on stable particle level from simulation. The jet energies are obtained as the sum of the calorimeter cells involved and weight factors for each cell are derived in a global χ^2 minimisation [40]. Since this procedure directly compares reconstructed jet energies with the particle energies before the detector simulation, it has to be repeated for each jet algorithm configuration one wants to use. The local calibration [41] tries to calibrate each calorimeter cluster to the hadronic energy scale and to remain jet algorithm agnostic. It classifies clusters into electromagnetic and hadronic clusters and tries to consider all effects in the hadronic shower in a bottom-up approach. These are e. g. the sampling fraction for hadronic showers, out-of-cluster corrections and dead material corrections. Within this thesis, locally calibrated topological clusters will be exclusively used.

4.1.2 Jet algorithms

For the actual jet finding, there are two common types of jet algorithms: the *cone type* and *successive combination type* algorithms. One example of each type will be described.

Cone algorithm

Cone type algorithms treat jet reconstruction as finding cones of fixed size in η-ϕ-space, whose energy barycentre aligns with the axis of the cone. In ATLAS, the seeded fixed-size cone algorithm used to be the standard for jet reconstruction. It starts with cones of fixed radius around seed clusters and calculates a new jet axis using the four-vector sum of all clusters directions in the cone. The calculation of a new axis is iterated until the new and current jet axis only differ by less than a given cut value [42]. Commonly used values for the cone size are $R = 0.4$ or $R = 0.7$.

In general, simple cone algorithms face some problems. First, it is possible that neighbouring cones overlap and one has to introduce a split and merge procedure between jets to ensure that double counting of energy is avoided. Second, the infrared and collinear safety of such an algorithm is not guaranteed per se. However, it is important for theoretical calculations that a jet algorithm be infrared safe, i. e. yields the same jet configuration upon emission of infinitely soft particles, and collinear safe, i. e. yields the same jet configuration upon particle splittings with zero opening angle. For the cone type algorithms, one can make modifications to ensure these properties at the expense of increased algorithm complexity [43].

k_T algorithm

Successive combination type jet algorithms [44–46] do not share the overlap problem by construction. They advance in the clustering procedure by merging two input objects at a time and iterating this procedure. Thus, any input object unambiguously ends up in one jet. The problem of infrared and collinear safety is overcome with a proper choice of distance measure between two objects. This measure decides which input objects to merge first.

The k_T jet algorithm is a successive combination type algorithm to reconstruct jets in hadron collider experiments. Its main feature is the use of relative transverse momentum k_T as the distance variable, which leads to its safety against collinear and soft radiation. This makes it a more preferred choice compared to (seeded) cone-type algorithms for correct calculation of observables both theoretically and experimentally. The origins of the algorithm are a modification of the JADE [47] and Durham [48] algorithms that are in use to reconstruct jets in e^+e^- collisions. The adaptation to hadron-hadron collisions has to take the beam remnants that are present in those collisions into account. The algorithm also puts an emphasis on the boost invariance of the event along the beam axis rather than on rotational invariance, as it was the case for e^+e^- collisions [49].

The mode of operation of the k_T algorithm mostly used is the *inclusive* mode. In this thesis, in addition the *exclusive* mode will be used, since it delivers event shape variables that can potentially be used for event classification. In the exclusive mode, one tries to reconstruct the jet evolution of an event and separate the hard jets from proton remnants and from objects caused by soft interactions. The algorithm takes the following steps:

1. Define a resolution scale d_cut.
2. Take all input objects to the algorithm into the *list of protojets.*
3. Determine the transverse momenta $d_i = p^2_{\mathrm{T},i}$ for all protojets i,
   and the relative transverse momenta $d_{ij} = \min(p^2_{\mathrm{T},i}, p^2_{\mathrm{T},j})\Delta R^2$ for all pairs of protojets i and j.

4. Find the smallest of all d_i and d_{ij}, and call it d_{min}.
5. If $d_{\mathrm{min}} > d_{\mathrm{cut}}$, the jet making procedure is stopped.
6. If $d_{\mathrm{min}} \in d_{ij}$, merge the two protojets i and j according to $p^{\mu}_{ij} = p^{\mu}_i + p^{\mu}_j$;
   if $d_{\mathrm{min}} \in d_i$, discard protojet i from the list of protojets and take it as a *beam jet*.
7. Iterate from step 3.

The choices of distance parameters d in step 3 and recombination prescription in step 6 are called ΔR distance scheme and E recombination scheme and are used throughout this thesis. Other choices would e. g. be a pure angular distance scheme and E_{T}- or p_{T}-conserving recombination prescriptions. All protojets remaining in the list after step 7 are the final reconstructed jets of the k_{T} exclusive algorithm. An alternative stopping criterion would be to cluster until a fixed number of jets is left over in the list of protojets and to declare these as final jets.

Since recently, there are also improved implementations of the k_{T} algorithm, that do not recalculate all distances in step 3, but only the ones which change in step 6 [50]. This makes it feasible to use k_{T} algorithms for jet reconstruction in dense environments (even ion collisions) in a reasonable amount of computing time and is now the ATLAS standard for k_{T} jet reconstruction.

At every step in the clustering procedure, the quantity d_{min} is a measure for the smallest separation of all protojets in relative transverse momentum and can be used as an event shape variable, as is traditionally done in e^+e^- experiments [51]. The inclusive mode of operation is the one closer to the original formulation of the k_{T} algorithm for application in e^+e^- collisions. There one does not need to separate the hard interaction from the underlying event and it is much easier to reconstruct inclusive cross sections than for hadronic collisions. In the inclusive mode of operation, jets having the smallest d_i are not classified as beam jets but as final jets. The resolution scale d_{cut} is replaced by an R parameter that controls the balance between merging and finalising jets. The procedure is iterated until all jets are marked final. Thus, one gets a large number of jets (also low-energetic ones) that are suitable for measuring inclusive cross sections. In this thesis, calorimeter reconstructed jets made with locally calibrated topological clusters are used, reconstructed by an inclusive k_{T} jet algorithm with ΔR distance scheme and E recombination scheme. Additionally, event shape variables are reconstructed from the merging scales of a separate k_{T} algorithm in the exclusive mode.

4.2 Cut-based top quark mass analysis

The cut-based analysis investigated by the top quark analysis group at MPP is aimed at delivering an early and robust measurement of the top quark mass. As such, it cannot rely on perfect detector calibration that is necessary e. g. for b-tagging. The top quark mass is reconstructed from the jets in the hadronic branch of l+jets-decaying top quark pairs (more specifically from e+jets and μ+jets-decaying top quark pairs). This decay channel provides acceptable branching fractions together with good tagging and reconstruction possibilities (c. f. chapter 2). This makes it possible to apply strict requirements to reduce background contaminations, while still retaining enough events to reconstruct a resonance peak in the invariant mass spectrum. The analysis is based on previous work [52, 53] and is presented here in brief:

4.2.1 Object and event selection

The input objects (jets, electrons, muons, $E_{\mathrm{T}}^{\mathrm{miss}}$) have to fulfil certain quality criteria to be regarded in the reconstruction:

- Electrons have to be reconstructed by the *egamma* algorithm with tight cut definitions [52]. They need at least 20 GeV of transverse momentum, and an energy in a cone of $\Delta R = 0.2$ around them of less than 6 GeV to be isolated.
- Muons have to be reconstructed by the *Staco* algorithm, which takes a combination of Inner Detector and Muon system tracks for muon identification [54]. They need at least 20 GeV of transverse momentum, and an energy in a cone of $\Delta R = 0.2$ around them of less than 6 GeV to be isolated.
- Jets need to have at least 20 GeV (40 GeV) of transverse momentum to pass the low-p_{T} (high-p_{T}) selection, respectively. Jets that are closer in ΔR than 0.15 to a selected electron or muon are not considered. Jets are reconstructed using the k_{T} inclusive algorithm with $R = 0.4$. The input to the jet algorithms consists of locally calibrated topological clusters.

Tauons are not considered in this analysis, since they are difficult to reconstruct and mainly decay into a low number of hadrons. Therefore it is not expected that a tauon reconstruction with high efficiency is available from the start of ATLAS. However, a τ+jets event where the tauon decays into an electron or a muon passing the aforementioned cuts can enter this analysis. In principle, the presence of two neutrinos could also lead to kinematical effects on the hadronic side of such an event. To pass the requirement on the lepton p_{T}, slightly harder leptonically decaying top quarks are selected, which could also affect the mass reconstruction on the hadronic side of the event. However, the amount of τ+jets events in the selected sample is only around 7 %. Later on, it is seen that the inclusion of τ+jets events in the signal introduces no mass-shifting effect and thus they are not considered as a background to the analysis, but rather included in the signal.

To retain high reconstruction efficiency, all input objects are only considered in a pseudorapidity range of $|\eta| < 2.5$, which is the acceptance of the inner tracking detector.

The selection of l+jets top quark pair decay candidate events is based on these high-level input objects and requires:

- at least one lepton (electron or muon) (this effectively reduces background from the all-hadronic decay channel of $t\bar{t}$ pairs and non-resonant multijet processes),
- at least three high-p_{T} jets and an additional low-p_{T} jet (this reduces events with low jet multiplicity, e. g. W+jets events with only few additional jets).

Since the signal events contain a neutrino in the final state, missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is normally used as another requirement. In the given samples this requirement is not possible due to the way they are simulated: In 2008, one quadrant of one hadronic endcap calorimeter suffered from power supply failures and no signals could be read out. The simulation was altered to reflect this detector condition. Due to this failure, the overall balance of transverse energy reconstruction is not given and the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable is not easily usable. To cope with this limitation, a method was developed to identify events where the jet reconstruction could have been flawed. Events with less than 10 GeV activity in the electromagnetic calorimeter in front of this HEC quadrant are found to have a chance of 95 % that no jets are present that would have passed the jet cuts. Those events are retained. Events with more than 10 GeV electromagnetic activity in this quadrant are rejected. This procedure, termed HECQ veto, replaces the requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$.

In addition, a loose preselection is applied to all the samples before they are analysed with the top mass analysis software. This reduces processing time and makes local analysis of the samples possible. In addition, the preprocessing creates additional jet collections which are not standard content of the ATLAS AOD data format but should be available for the analysis. These are inclusive k_{T} jets and anti-k_{T} jets with $R = 0.4$ and $R = 0.6$, and cone jets with $R = 0.4$, all made using locally calibrated topological clusters. From these collections, the inclusive k_{T} jets with $R = 0.4$ are used further on. The preselection requires at least one lepton with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and at least 2 jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ plus a third jet with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. All objects have to lie within $|\eta| < 3$ to be considered.

4.2.2 Monte Carlo datasets

This work is focused on specific signal and background datasets from simulation. The used data samples were all simulated at a centre-of-mass energy of $\sqrt{s} = 10\,\mathrm{TeV}$, which was the value foreseen for the initial period of LHC running at the time of simulation.

Signal sample

The l+jets signal is contained in a sample together with ll decay processes from $t\bar{t}$ production simulated with the MC@NLO generator [55]; the QCD parton showering was simulated with Herwig [56, 57]. The simulation of the underlying event was carried out with the program Jimmy. The top quark mass was taken to be 172.5 GeV and a K-factor of 1.07 is applied to rescale the cross section from MC@NLO to the near-NNLO cross section from [58]. The PDFs used for this MC@NLO sample are CTEQ6.6 [59], the standard NLO CTEQ PDF in the $\overline{\mathrm{MS}}$ scheme from 2008. Due to its method to calculate NLO cross sections, MC@NLO produces a fraction of events with negative weights and all distributions of physical quantities are made by incorporating the event-wise weight. Finally, GEANT [60, 61] is used to simulate the detector response to the particles interacting with the detector components. For the analysis, only the l+jets part of the sample is used; ll events are dropped by inspecting the decay of the Monte Carlo contents of the events.

Background samples

Since the W+jets process shares the event topology with the l+jets top quark pair decay channel, it is an important background, and efficient methods for its rejection are needed. It was seen that W+jets events remain the dominant source of background after all described cuts [52]. Events from QCD multijet production can also enter the analysis with reasonable statistics. However, the large number of simulated events that is necessary due to the huge cross sections involved prohibit the use of Monte Carlo generated samples to fully study the multijet background. Thus, it is foreseen that the QCD multijet background will be extracted from real data. An intermediate solution that is followed currently, is to only estimate the expected total number of multijet events that enter the analysis from analysing the QCD Monte Carlo samples, while approximating distributions of their physical quantities by W+jets events. In the following, rejection methods against W+jets background are explored and therefore the analysis only uses W+jets background.

The used background samples consist of W+jets events with 2–5 additional partons simulated with AlpGen [62] and showered with Herwig. This AlpGen generation of events uses the CTEQ6L PDF set, which is the LO PDF compatible with CTEQ6.6. A K-factor of 1.22 scales the LO cross section determined by AlpGen to the near-NNLO prediction made with the program FEWZ [63]. The MLM matching algorithm [64] ensures that the samples are exclusive between matrix element partons and partons from the showering model, so that they can be used together in the analysis. Additionally, a filter at stable particle level is applied which requires at least 3 jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ to reduce the number of events that have to be simulated in the GEANT detector simulation.

The W mass is taken to be 80.4 GeV in all of the samples. Table 4.1 shows the ATLAS Monte Carlo datasets used in this study together with their cross sections and available statistics at the time of writing. In this table, $W(e\nu) + 2p$ denotes e. g. a final state after the AlpGen matrix element of a leptonically decaying W into e and ν_e plus 2 additional partons. The cross sections include branching ratios, filter efficiencies and MLM matching efficiencies for background samples. The final cross section that is used in the analysis is then given by $\sigma\times$K-factor. One can see a large difference in the generator cross sections for samples with the same multiplicity of additional partons, but a different lepton flavour from the leptonic W decay. This is a result of the aforementioned stable particle jet filter: To pass this filter, an event must contain three of those jets above a certain momentum. Since electrons create calorimeter clusters and are reconstructed in the jet making, it was decided that they also appear as input for jet making on stable particle level. The requirement of three jets is then easier to fulfil for $W \to e\nu$ events than for $W \to \mu\nu$ events, since the former get one jet (from the electron) "for free".

The actual analysis is performed using the ATLAS software Athena [65] in release 14.2.21 with the AthenaROOTAccess functionality. The code is split into steering files written in Python and algorithmic parts in C++ and can be accessed via the webfrontend of the ATLAS version control system [66].

For analysis, a fraction of events corresponding to $146\,\mathrm{pb}^{-1}$ is taken randomly from the available statistics of each sample. This value is inspired by a mixed sample that was used in the ATLAS top quark analysis group and is well within the luminosity that is expected for the first phase of LHC running. The numbers of events that pass the subsequent application of the given analysis requirements are shown in table 4.2. One

Table 4.1: *Signal and background Monte Carlo samples used for the analysis together with their available statistics. For all samples, the simulation of QCD showering is performed with the Herwig code. Sample 108251 was not available at the time of analysis.*

sample	physics process	σ [pb]	K-factor	events	$\mathscr{L}$ [pb^{-1}]	ME generator
105200	$t\bar{t}$ l+jets and ll	202.86	1.07	359174	1654.7	MC@NLO
108240	$W(e\nu) + 2p$	148.51	1.22	37714	205.5	AlpGen
108241	$W(e\nu) + 3p$	105.07	1.22	24480	208.8	AlpGen
108242	$W(e\nu) + 4p$	42.69	1.22	8500	207.0	AlpGen
108243	$W(e\nu) + 5p$	15.94	1.22	3000	213.6	AlpGen
108244	$W(\mu\nu) + 2p$	7.24	1.22	1981	239.1	AlpGen
108245	$W(\mu\nu) + 3p$	48.06	1.22	10500	205.6	AlpGen
108246	$W(\mu\nu) + 4p$	30.84	1.22	6313	206.1	AlpGen
108247	$W(\mu\nu) + 5p$	14.13	1.22	3000	213.3	AlpGen
108248	$W(\tau\nu) + 2p$	62.11	1.22	14974	205.3	AlpGen
108249	$W(\tau\nu) + 3p$	71.43	1.22	12750	155.0	AlpGen
108250	$W(\tau\nu) + 4p$	35.33	1.22	7927	210.8	AlpGen
108251	$W(\tau\nu) + 5p$	15.26	1.22	N/A		AlpGen

can already note that the initial event composition is dominated by background events and that these are unequally distributed between the different lepton flavours of the W decay. This is due to the described jet filter applied to those samples. In the course of the selection cuts, the overall background fraction within the selected events drops from 75 % to about 33 %. This analysis is the baseline for further studies on improved background rejection possibilities.

4.2.3 Mass reconstruction

The mass of the top quark m_t is reconstructed in the selected events in the following way: All possible three-jet combinations from the selected jets (i. e. the ones passing the low-p_{T} selection) are made. The particular combination maximising the p_{T} of the four-vector sum of the three jets is chosen to represent the top quark. This reconstruction is called $p_{\mathrm{T}}^{\mathrm{max}}$ method. A histogram for the reconstructed m_t is shown in figure 4.1. The black points represent the simulated data corresponding to 146 pb^{-1} of luminosity, as composed by the samples described above. The blue histogram represents the W+jets samples only, i. e. the physics background part of the distribution. The green histogram shows those signal events, where not all of the three jets chosen really stem from the hadronically decaying top quark, as determined by the generator evolution in the event. These are the “combinatorial background” events, where the $p_{\mathrm{T}}^{\mathrm{max}}$ method failed to determine the correct three jet combination. In contrast, the red histogram shows those signal events where the $p_{\mathrm{T}}^{\mathrm{max}}$ method correctly selected the jets coming from the hadronically decaying top quark. This is tested by assigning to each of the three simulated partons from the top decay the closest reconstructed jet in ΔR from the jet list. If no double assignments are present, all distances are below 0.2, and the same jets that $p_{\mathrm{T}}^{\mathrm{max}}$ chooses are selected, the top quark is considered being correctly reconstructed. In figure 4.1, 754 events are correctly reconstructed, in addition to 3614 combinatorial and 2164 W+jets background

Table 4.2: *Expected numbers of events for a luminosity of 146 pb^{-1} after the requirements described. The l+jets row shows the cut flow for the l+jets part of the signal sample.*

sample	initial	presel	1 lepton	4 low-p_T jets	3 high-p_T jets	HECQ veto
$t\bar{t}$ l+jets & ll	31691	22846	16058	8383.9	5659.8	4887.8
l+jets		17836	11885	7340.9	5044.9	4358.4
$W(e\nu)+2p$	26796	15917	13390	154.2	37.7	32.7
$W(e\nu)+3p$	17116	10662	8139	599.9	224.4	198.6
$W(e\nu)+4p$	5996	3905	2696	821.9	372.5	314.6
$W(e\nu)+5p$	2050	1402	856	516.9	296.1	235.9
$W(\mu\nu)+2p$	1209	896	644	87.9	24.4	22.6
$W(\mu\nu)+3p$	7456	5596	4444	632.7	281.2	254.9
$W(\mu\nu)+4p$	4473	3578	2820	1186.2	621.5	557.0
$W(\mu\nu)+5p$	2054	1718	1340	902.0	562.5	486.6
$W(\tau\nu)+2p$	10650	1545	590	14.2	4.3	2.8
$W(\tau\nu)+3p$	12011	2552	788	76.8	31.3	25.6
$W(\tau\nu)+4p$	5491	1616	412	153.8	81.8	74.1
sum backgd.	95302	49387	36119	5146.5	2537.7	2205.4

events, yielding a background fraction including combinatorics of about 88 % instead of 33 %.

The invariant mass distribution is then fitted with the sum of a Gaussian and a Chebychev polynomial of 8^{th} degree to extract the mass. The Chebychev polynomials of the first kind are used [67, 68], as defined within $x \in [-1; 1]$ via their recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \tag{4.1}$$

with

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \quad . \end{aligned} \tag{4.2}$$

The fit range [30; 500] (in GeV) is transformed to the interval [−1; 1] by calculating $T(x')$ with

$$x' = \frac{2(x - x_{max})}{x_{max} - x_{min}} + 1 \quad . \tag{4.3}$$

The Chebychev polynomials are highly suitable for polynomial fitting, since they form an orthogonal system of polynomials. In addition, they closely approximate the optimal fit polynomial, which is defined as the fit function that minimises the maximal deviation from the true function.

In figure 4.1, the Gaussian part of the fit is shown as black curve, the Chebychev part as green curve, and the upper black curve is their sum. The mean of the Gaussian is then identified with the reconstructed mass of the top quark. This assumption is only meaningful if the Gaussian is a proper estimate of the histogram of correct combinations, which is not guaranteed to be true. The combinatorial background has a peaking structure and as such can move the maxima of Chebychev and Gauss curves by some amount

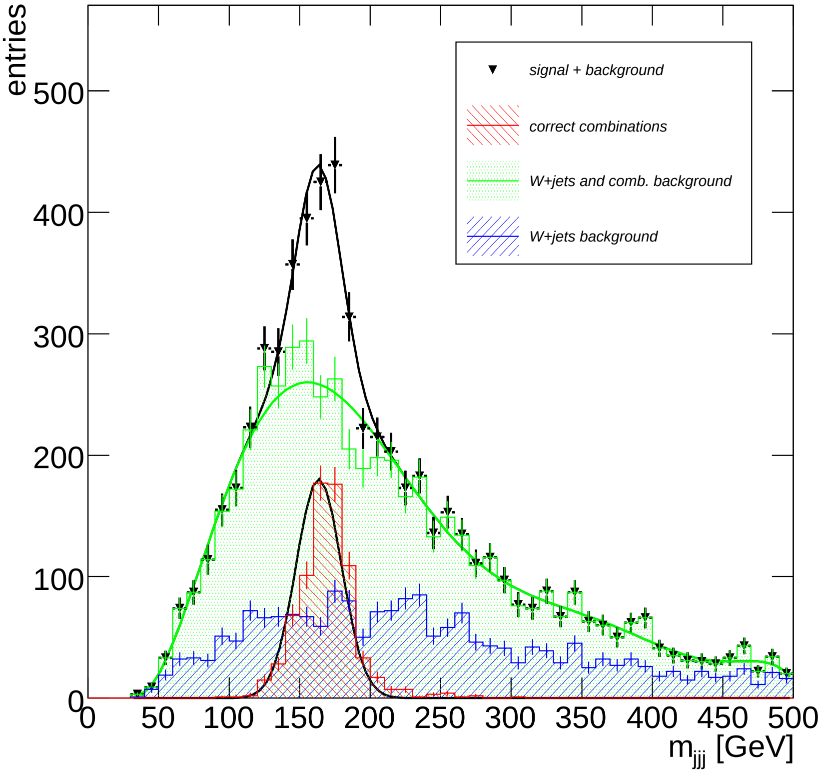


Figure 4.1: *The invariant mass of the three chosen jets as a measure of the top quark mass. The blue histogram represents the W+jets samples, the red histogram shows the signal events with the correct three jet combination. The green histogram is the sum of W+jets background and the not correctly assigned signal events. The black points represent the total distribution. The results from the fitting procedure to the total simulated data are shown as curves: the Gaussian part in black, the Chebychev part in green, the upper black curve is the sum of both.*

(usually to lower mean values for the Gaussian, since it absorbs some portion of the background). The histogram of correctly reconstructed jet triplets peaks at 168 GeV and not at the input top quark mass of 172.5 GeV due to the non-final local jet calibration and jet algorithm effects. The background fractions within the signal region, as defined by the interval of $\pm 1\sigma$ around the mean value of the Gaussian, are 17.5 % physics background and 67.7 % total background. It is noteworthy that with this amount of statistics, a higher order Chebychev polynomial is likely to closely follow the statistical fluctuations of the sample and therefore is only useful for higher statistics. However, it can then better follow the peaking combinatorial background and should yield smaller biases from the fitting procedure.

To test for potential mass-shifting effects by the inclusion of τ+jets events, the ratio of the reconstructed mass distributions for e/μ+jets and τ+jets $t\bar{t}$ decays is examined. A fit of the resulting distribution with a constant around the peak region ($50\,\text{GeV} < m_{jjj} < 250\,\text{GeV}$) shows an equal compatibility to the data compared to a straight-line fit. Both yield a χ^2/NDF of about 49/40 and the slope of the straight

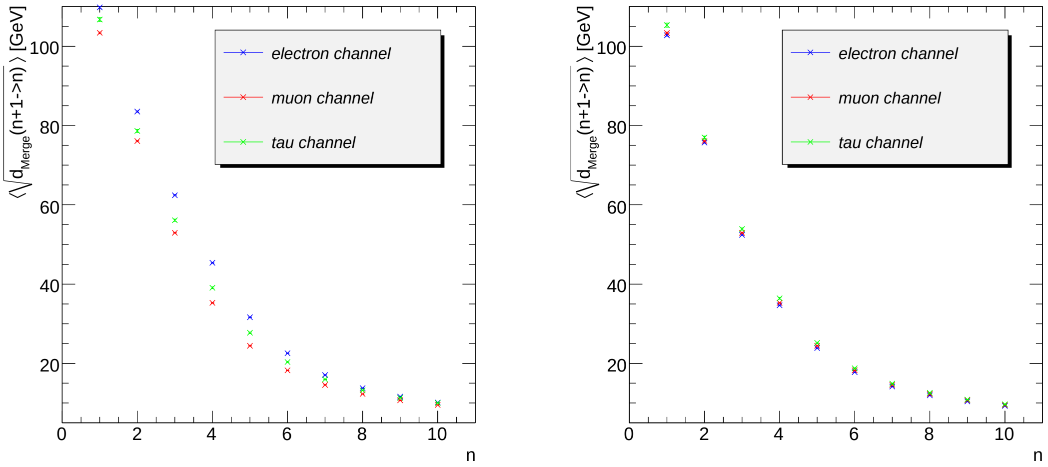


Figure 4.2: *Mean of various* d_{Merge} *values in the signal sample split by their W decay channel. The left figure shows the reconstructed* d_{Merge} *when all clusters are used, the right figure shows the same when applying the electron cluster removal procedure described in the text. The error bars (too small to be seen on any of the points) denote the RMS divided by the square root of entries of the distribution.*

line is only 15 % of its uncertainty. It is therefore concluded that τ+jets events introduce no systematic shift to the mass distribution.

4.3 Usage of d_{Merge} as event shape variables for background rejection

To characterise the evolution of the jet clustering, the observables known as $d_{\mathrm{Merge}}(n+1 \to n)$ are introduced. They are defined as the minimum d in an exclusive k_{T} algorithm at the jet clustering step with $n+1$ protojets left. That means that setting $d_{\mathrm{cut}} \geq d_{\mathrm{Merge}}(n+1 \to n)$ results in n or less jets, while setting $d_{\mathrm{cut}} < d_{\mathrm{Merge}}(n+1 \to n)$ yields at least $n+1$ jets.

In ATLAS, the energy that is deposited by electrons and muons is used in the clustering to calorimeter clusters. Electrons can therefore create jets on their own, as opposed to muons which do only rarely leave significant energy in the calorimeter. This creates an asymmetry in the jet evolution of muon and electron channels. To cope with this, the d_{Merge} values used in the following analysis are reconstructed by applying the k_{T} algorithm in the exclusive mode on a filtered list of calorimeter clusters. Every cluster that is closer in ΔR than 0.1 to a *selected* electron is taken off the input list for jet making. Figure 4.2 shows that the dependency on the lepton flavour in the W decay is essentially removed when applying this recipe to the signal sample. Before applying the electron cluster removal (left side of figure 4.2), the d_{Merge} distributions of different W decay lepton flavours have different mean values. After removing the electron clusters from the input of jet making, the mean values of d_{Merge} become compatible for all lepton flavours. In principle, one would need to resort to more elaborate techniques to match reconstructed electrons to clusters, since the electron reconstruction uses *sliding window* *clusters* on the calorimeter cells [38]. In contrast, the clusters used for jet making are topological clusters with non-regular shape and therefore cannot be matched one-to-one with the fixed-size sliding window. However, figure 4.2 shows that the simple method used here leads to satisfactory results.

Further on, the d_{Merge} distributions will be explored as event shape variables to further distinguish between the l+jets top decay signal and W+jets background. The idea is that the three jets from the hadronically decaying top quark create a different pattern in the relative transverse momentum than the uncorrelated jets from the background. This effect is expected to be most visible in the $d_{\mathrm{Merge}}(5 \to 4)$ and $d_{\mathrm{Merge}}(4 \to 3)$ variables, since at these clustering steps the four jet configuration of the signal events is reached and any further clustering leads to merging of the decay products of the top quarks.

4.3.1 Cut-based background rejection

For the cut-based analysis, the distributions of d_{Merge} variables are built for signal and background by an exclusive k_{T} algorithm and examined for their separation power. The k_{T} algorithm is set to perform the clustering procedure to the very end and return the d_{Merge} value for this event at each clustering step. The distributions of the first ten d_{Merge} variables are shown in figure 4.3. The dimension of d_{Merge} is that of momentum squared and for easier interpretation, the square root is displayed. Only events that pass the standard cut-based top analysis are shown in the histogram, again normalised to $146\,\mathrm{pb}^{-1}$ of luminosity. The signal events are denoted with solid line histograms, where the electron (blue), muon (red), and tau (green) channels are shown as summed up histograms. The W+jets background is denoted with dotted histograms, again split into the different subchannels. In this figure, a separation between top signal and W+jets background is visible, especially in the variables $d_{\mathrm{Merge}}(5 \to 4)$ and $d_{\mathrm{Merge}}(4 \to 3)$, as indicated by the largest difference in $|d^{\mathrm{mean}}_{\mathrm{Merge}_{top}} - d^{\mathrm{mean}}_{\mathrm{Merge}_{W+jets}}|$. However, the distributions are largely overlapping and any requirement placed on a d_{Merge} variable in order to increase signal purity will cut severely into the signal statistics.

To evaluate which of the variables is the most discriminating one, the background rejection versus signal efficiency curves are shown in figure 4.4. The first seven d_{Merge} variables are shown with the requirement on d_{Merge} scanned between $0 < d_{\mathrm{Merge,min}} < (300\,\mathrm{GeV})^2$. The background rejection is defined as the fraction of background events that does not fulfil the requirement, while the signal efficiency is defined as the fraction of signal events that fulfils the requirement. It can be seen that the area under the curve, which can be used as a measure of the effectiveness of the variable, is maximal for $d_{\mathrm{Merge}}(5 \to 4)$, while in some regions of efficiency $d_{\mathrm{Merge}}(6 \to 5)$ or $d_{\mathrm{Merge}}(4 \to 3)$ outperform it in rejection power. Higher and lower multiplicity variables are severely weaker in discrimination power or even not discriminating at all. The curve for $d_{\mathrm{Merge}}(2 \to 1)$ lies below the diagonal in this plot, which means that a selection of $d_{\mathrm{Merge}}(2 \to 1) < d_{\mathrm{Merge,max}}$ achieves separation between signal and background. The areas under the different curves are given in table 4.3. For $d_{\mathrm{Merge}}(2 \to 1)$, $(1 - area)$ is denoted in brackets as the real separation power for requiring an upper instead of a lower limit. The requirement on the most sensitive d_{Merge}, i. e. $d_{\mathrm{Merge}}(5 \to 4)$, is then optimised to maximise efficiency×rejection. For a value of $d_{\mathrm{Merge,min}} = (29.2\,\mathrm{GeV})^2$, efficiency×rejection is about 0.38. At this configuration, 60.3 % of signal events are retained, while 63.4 % of background events are rejected. The fraction

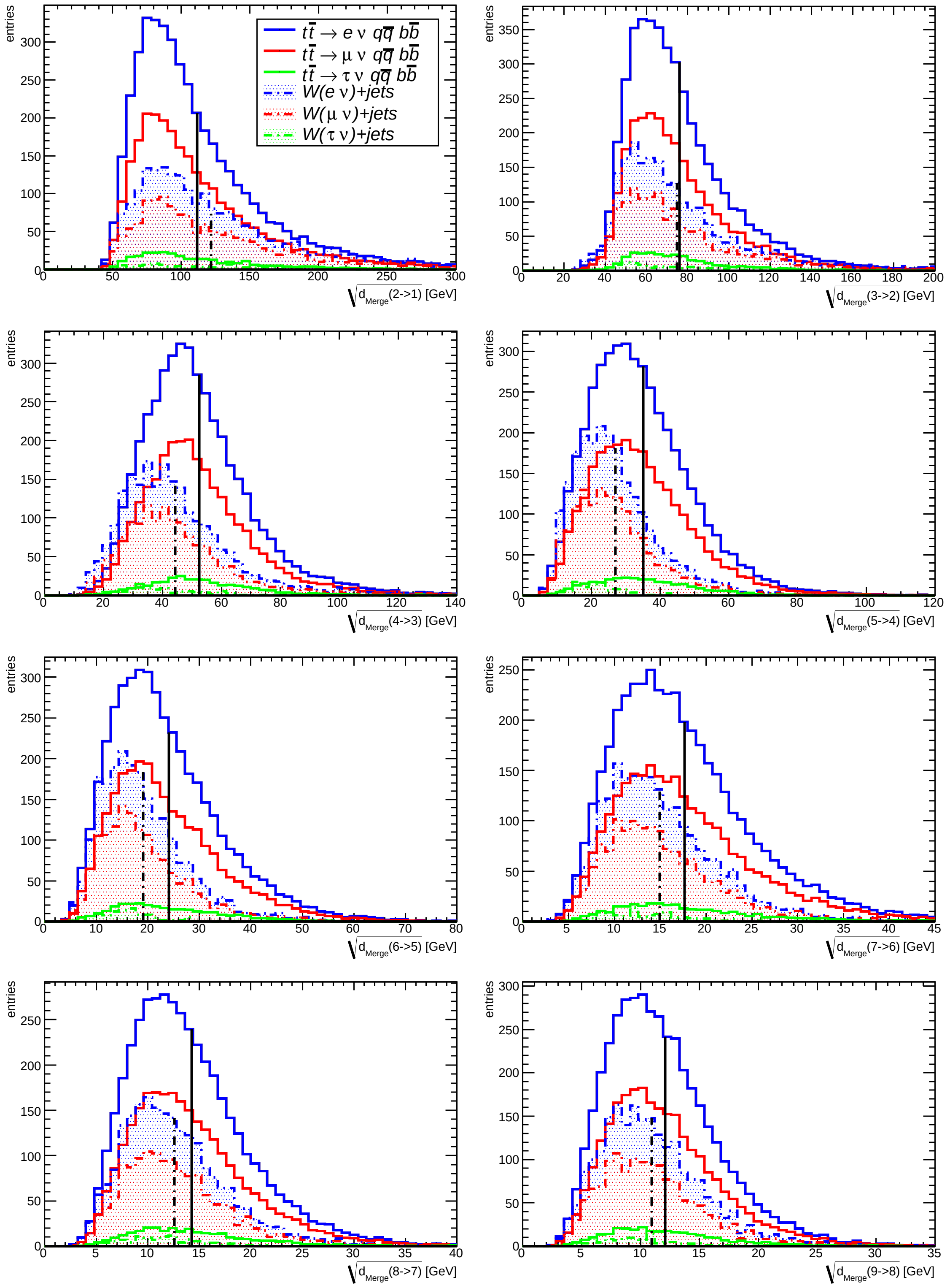


Figure 4.3: *The* $\sqrt{d_{\mathrm{Merge}}}$ *values at various steps of jet clustering. The electron (blue), muon (red) and tau (green) parts of the samples are shown as two sets of stacked histograms. The three signal channels are shown with solid lines, the three W+jets background samples with dotted lines. All histograms are scaled to a luminosity of* $146\,pb^{-1}$ *and only selected events are shown. The vertical lines represent the mean values of the signal (solid) and background (dashed) distribution.*

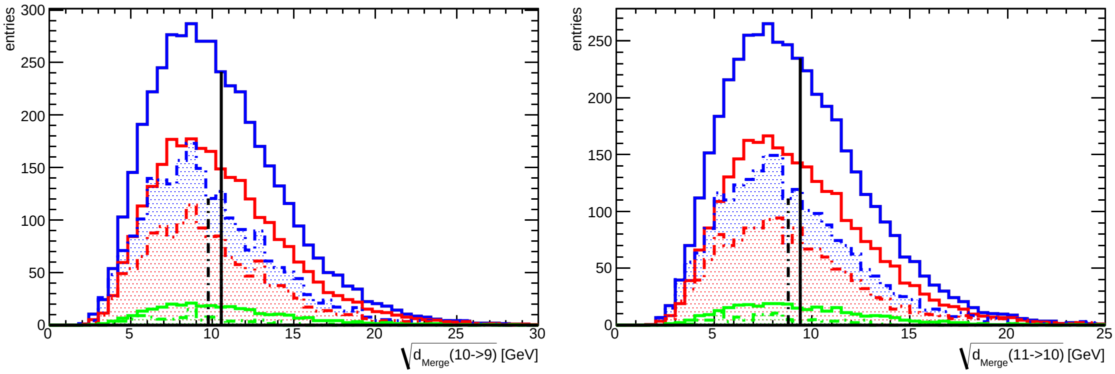


Figure 4.3: *. . . continued.*

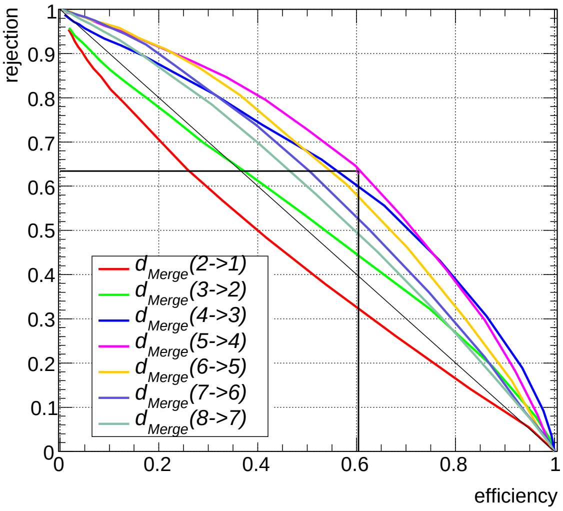


Figure 4.4: *Background rejection versus signal efficiency for different* d_{Merge} *variables. The black diagonal denotes the line where the separation power becomes zero. The chosen working point for the requirement on* $d_{\mathrm{Merge}}(5 \rightarrow 4)$ *is indicated.*

of physical background in the selected events is then reduced from 33 % to 23 %, and from 17.5 % to 10.4 % in the signal region around the Gauss peak.

The distribution of reconstructed three-jet masses using the $p_{\mathrm{T}}^{\mathrm{max}}$ method is given in figure 4.5. Only events satisfying the standard signal selection plus the requirement $d_{\mathrm{Merge}}(5 \rightarrow 4) > (29.2\,\mathrm{GeV})^2$ are displayed, normalised to $146\,\mathrm{pb}^{-1}$ of luminosity. The diagram shows that the relative strength of signal versus the W+jets background has increased compared to figure 4.1.

To investigate the possible shifts on the top quark mass due to the requirement on $d_{\mathrm{Merge}}(5 \rightarrow 4)$, the requirement is scanned from zero to $(55\,\mathrm{GeV})^2$ in steps of $(5\,\mathrm{GeV})^2$. For each point, the mean and width of the Gaussian from the combined Gauss+Chebychev fit are shown in figure 4.6a. The point without any requirement on d_{Merge} serves as the reference and its statistical error spans the turquoise band. Since every event passing tightened cuts is present in the samples with looser cuts, the points are strongly correlated.

Table 4.3: *Area under the rejection versus efficiency curve for different* d_{Merge} *variables. An area of 0.5 means no separation power for a variable.*

d_{Merge}	area
$d_{\mathrm{Merge}}(2 \rightarrow 1)$	0.434 (0.566)
$d_{\mathrm{Merge}}(3 \rightarrow 2)$	0.523
$d_{\mathrm{Merge}}(4 \rightarrow 3)$	0.633
$d_{\mathrm{Merge}}(5 \rightarrow 4)$	0.657
$d_{\mathrm{Merge}}(6 \rightarrow 5)$	0.626
$d_{\mathrm{Merge}}(7 \rightarrow 6)$	0.596
$d_{\mathrm{Merge}}(8 \rightarrow 7)$	0.570

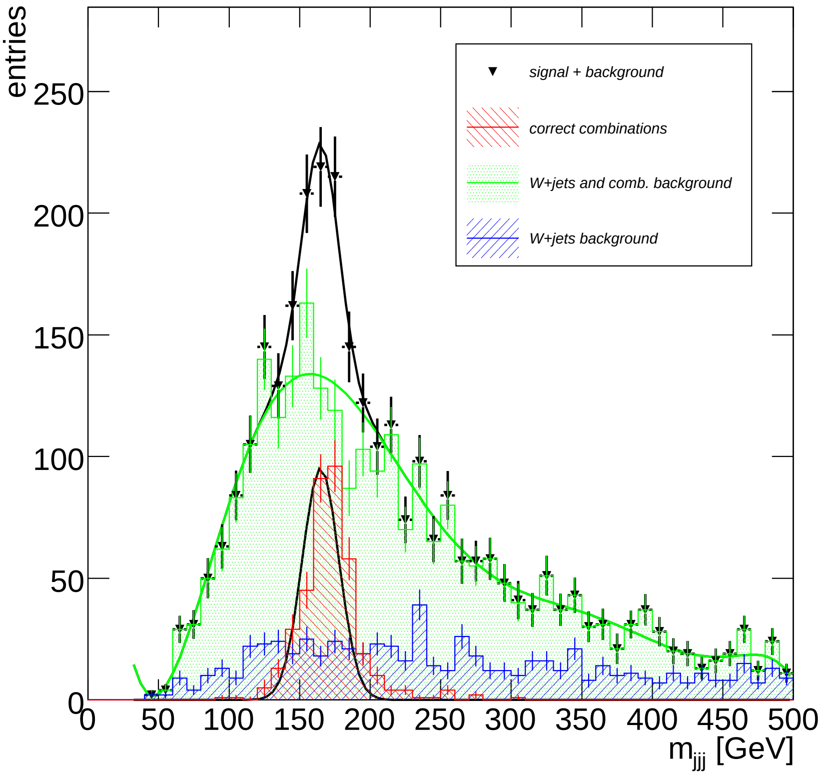


Figure 4.5: *Same as figure 4.1, but only events satisfying* $d_{\mathrm{Merge}}(5 \rightarrow 4) > (29.2\,\mathrm{GeV})^2$ *are shown.*

Subsequent points are assigned decorrelated uncertainties according to $\sigma_{decor}^2 = |\sigma^2 - \sigma_0^2|$, with σ_0 being the uncertainty of the leftmost point [69]. These represent the systematical uncertainty of the mean or width value when leaving out part of the events. Figure 4.6b shows the mean and width of a Gauss fit to the correctly assigned three jet combinations in the same manner. Thus one can deduce whether a possible systematic effect is acting on the real signal or only affecting the reconstruction of the total data including combinatorial and physics backgrounds.

Figure 4.6a shows that the peak of the "data" distribution stays stable against increasing the requirement. Within the decorrelated uncertainties, the mass values are still compatible with the initial mass value and the reconstructed masses stay within the statistical fluctuation of the first fit. The uncertainties become larger in the right direction since the samples lose a lot of statistics and the fit procedure can only find the peak of the data with larger uncertainty. The width of the Gaussian peak also essentially stays constant. In figure 4.6b, one can see a strong rise of the mass value for the correct jet triplets when increasing the cut. This is probably because higher d_{Merge} values at a particular clustering step mean higher separation between jets and thus also higher invariant masses. Placing requirements on a single d_{Merge} variable is thus dangerous, since it introduces a shift on the correctly reconstructed jet combinations that might go unnoticed due to the combinatorial and physics background not showing this behaviour.

4.3.2 Evaluation of multivariate discriminator techniques

Fisher discriminants

The Fisher discriminant method [70] is a multivariate analysis technique that classifies events into categories by analysing several input variables x_i. In high energy physics, usually only two categories—signal and background—are used. A Fisher discriminant is built by using a linear combination of the input variables, such that the output distribution is maximally separated for signal and background events. The coefficients are determined by projecting the n-dimensional input onto one dimension while maximising the separation between the classes. The ansatz for the Fisher discriminating function is

$$F(\boldsymbol{x}, \boldsymbol{w}) = w_0 + \sum_{k=1}^{n_{var}} w_k x_k \quad , \tag{4.4}$$

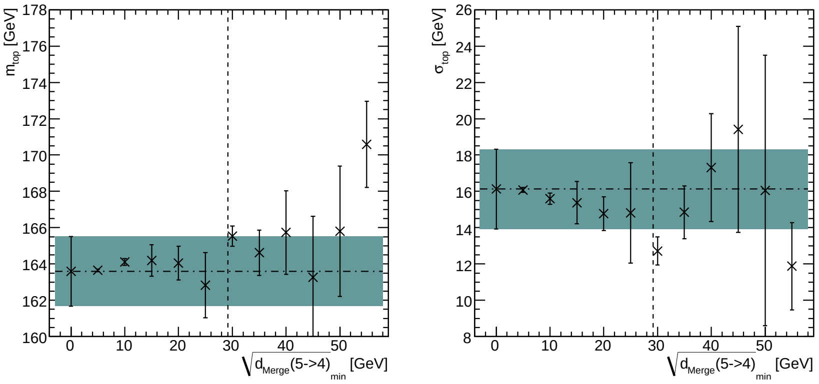
where $\boldsymbol{w}$ denotes the vector of weight coefficients to the input variables $\boldsymbol{x}$. The weights are then determined such that they maximise the following expression [71, 72]:

$$J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \mathscr{B} \boldsymbol{w}}{\boldsymbol{w}^T \mathscr{W} \boldsymbol{w}} \quad . \tag{4.5}$$

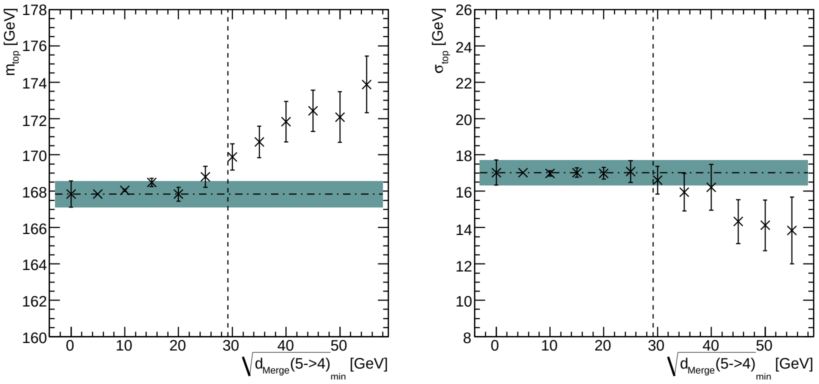
Here, $\mathscr{W}$ denotes the within-class covariance matrix

$$\mathscr{W}_{kl} = \sum_{U=S,B} \langle x_{U,k} - \langle x_{U,k} \rangle \rangle \langle x_{U,l} - \langle x_{U,l} \rangle \rangle \quad , \tag{4.6}$$

where the angle brackets denote the expectation value of the operand and the two event categories are labelled S for signal and B for background. The symbol $\mathscr{B}$ denotes the



(a) reconstructed m_t, i. e. Gauss part of the Chebychev+Gauss fit



(b) mass from Gauss fit to correct jet triplets

Figure 4.6: *Change of reconstructed top quark mass when increasing the requirement on* $\sqrt{d_{\text{Merge}}(5 \rightarrow 4)}$. *The left figures show the fitted means of the Gaussians with their fit uncertainties, the right figures show the widths. In each figure, the leftmost uncertainty bar is evaluated for the full sample and determines the statistical uncertainty* σ_0*, all other uncertainties are decorrelated according to* $\sigma^2_{decor} = |\sigma^2 - \sigma^2_0|$. *The dotted vertical line indicates the optimal requirement of* $d_{\text{Merge}}(5 \rightarrow 4)_{\text{min}} = (29.2\,\text{GeV})^2$.

between-class covariance matrix

$$\mathscr{B}_{kl} = \frac{1}{2} \sum_{U=S,B} \left(\langle x_{U,k} \rangle - \langle x_k \rangle \right) \left(\langle x_{U,l} \rangle - \langle x_l \rangle \right) \quad . \tag{4.7}$$

The total covariance matrix is then given by $\mathscr{C} = \mathscr{W} + \mathscr{B}$. In other words, the method tries to maximise the between-class separation $\boldsymbol{w}^T \mathscr{B} \boldsymbol{w}$, while trying to minimise the within-class dispersion $\boldsymbol{w}^T \mathscr{W} \boldsymbol{w}$. The result for the Fisher coefficients is then

$$w_k = \frac{\sqrt{N_S N_B}}{N_S + N_B} \sum_{l=1}^{n_{var}} \mathscr{W}_{kl}^{-1} \left(\langle x_{S,l} \rangle - \langle x_{B,l} \rangle \right) \quad . \tag{4.8}$$

A Fisher discriminant can deal with linear correlations between the input variables. However, it can only separate classes if the mean values of the input distributions differ; otherwise no projection separating them is possible. Unlike other multivariate classifiers, a Fisher discriminant is a parametric method and does not suffer from the "curse of dimensionality". This means that the drastically increased volume of phase space due to a high number of input variables does not diminish the classification performance of the method. It should therefore not suffer as much from a low amount of available training data as other classifiers do.

Application on Signal-Background classification

Since there are differences in the mean values of the various d_{Merge} distributions, a Fisher discriminant can be used to separate signal and background distributions. A Fisher discriminant using d_{Merge} variables was e. g. used in [73] to separate $t\bar{t}$ events at tree level from $t\bar{t}$ events with an additional radiated gluon. Hereafter, the potential of the method to better separate $t\bar{t}$ events from W+jets background than a single requirement on one d_{Merge} variable would do is evaluated. The Fisher discriminant method is supposed to be simple and robust with little or no tweaking of its parameters.

For this analysis, the "Toolkit for MultiVariate Analysis" (TMVA) [71] is used in version 3.9.5 in conjunction with ROOT 5.18 [74]. The TMVA package delivers a framework with many multivariate analysis methods applicable to the same data. The input that is taken here consists of the first 10 variables $\sqrt{d_{\text{Merge}}(2 \to 1)}$ to $\sqrt{d_{\text{Merge}}(11 \to 10)}$. A sample of 3000 $t\bar{t}$ and 1500 W+jets events is used to calculate the weights for the Fisher discriminant. Afterwards, the remaining statistics of the samples serves as test events to quantify the performance of the method.

The distributions of the input variables are shown in figure 4.7, where the d_{Merge} values are divided by the approximate peak value of their distribution and normalised to unit area for signal and background events, to compare the relative separation for the different variables. The ranking of input variables reported by the Fisher method corresponds directly to the separation that is visible between the mean values of signal and background distributions. The highest separating variables are $\sqrt{d_{\text{Merge}}(5 \to 4)}$, $\sqrt{d_{\text{Merge}}(6 \to 5)}$ and $\sqrt{d_{\text{Merge}}(2 \to 1)}$. Overall, higher d_{Merge} variables supply less information to separate the samples, since their separation between signal and background distributions becomes small. This is expected, since at those later stages of jet evolution, the d_{Merge} variables should not be different for different hard processes. The distribution

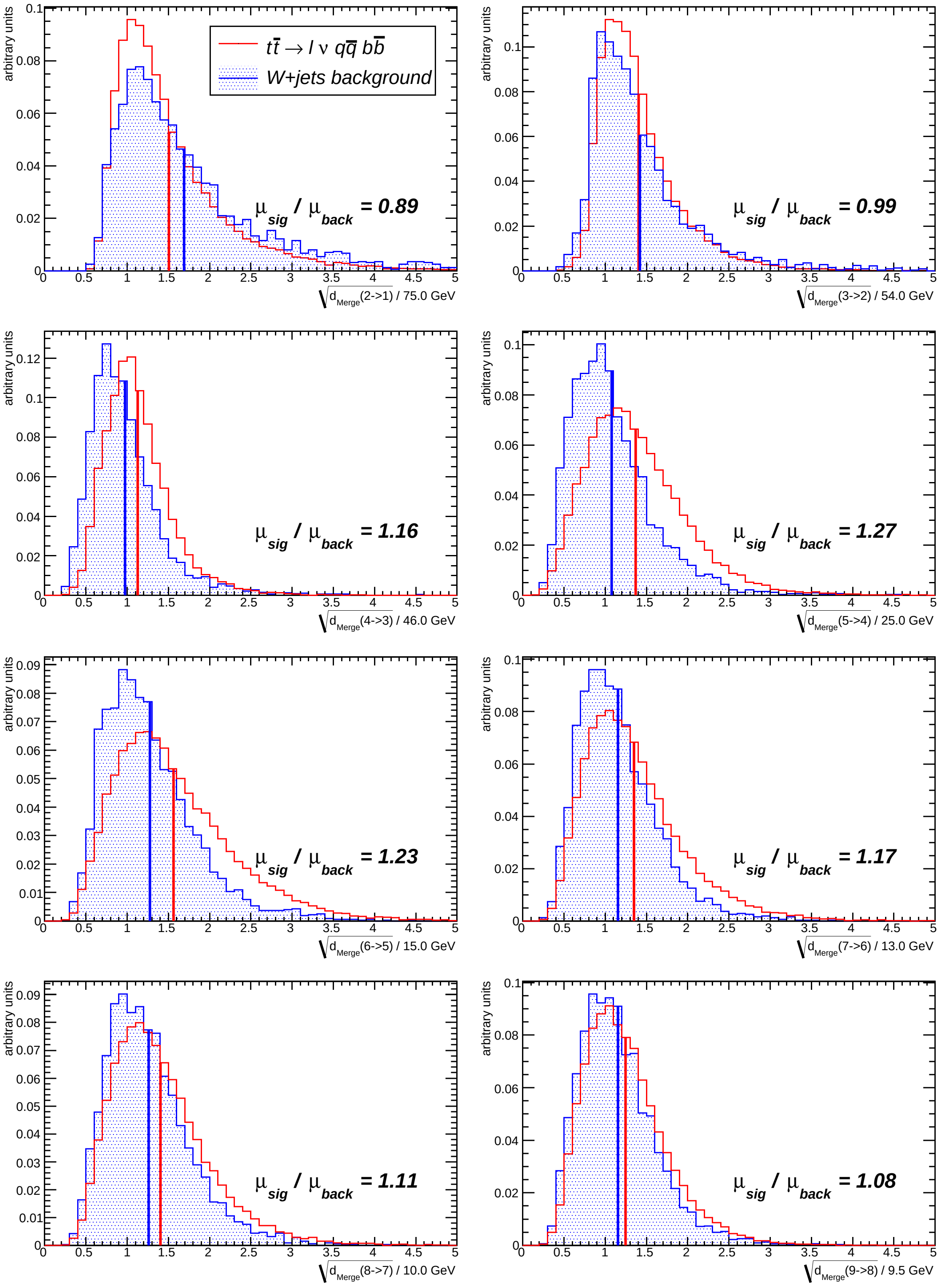


Figure 4.7: *Normalised distributions of the first 10 d_{Merge} input variables. Each variable is divided by its approximate peak value to compare the relative shapes. The red histograms depict the signal distributions, the blue histograms depict the W+jets background distributions. The ratios of the mean values of the distributions, denoted as vertical lines, are printed in each figure.*

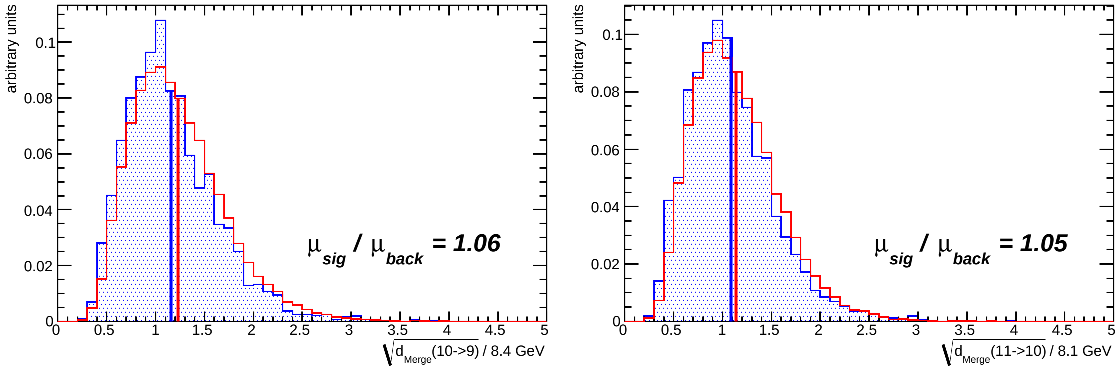


Figure 4.7: *. . . continued.*

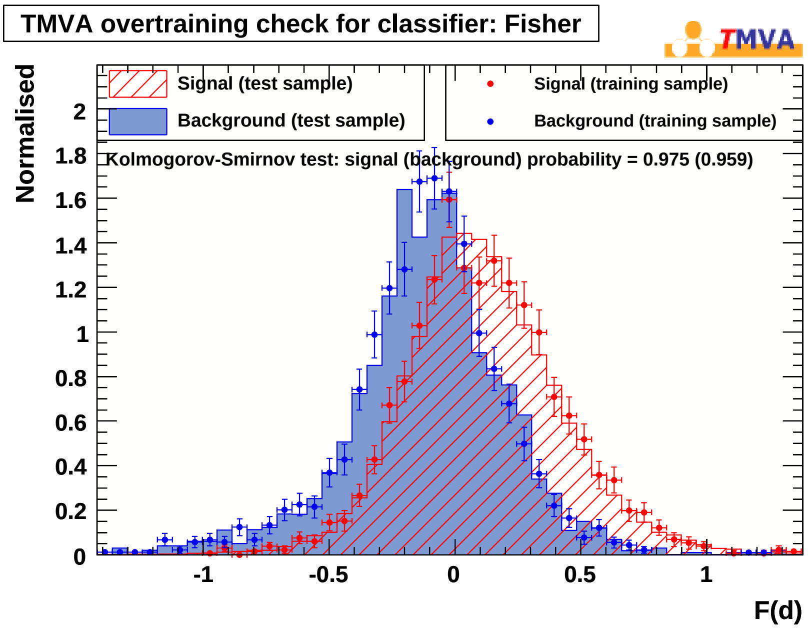


Figure 4.8: *The discriminator output of the Fisher method. The filled histograms denote the test sample output whereas the points with uncertainty markers denote the training sample output.*

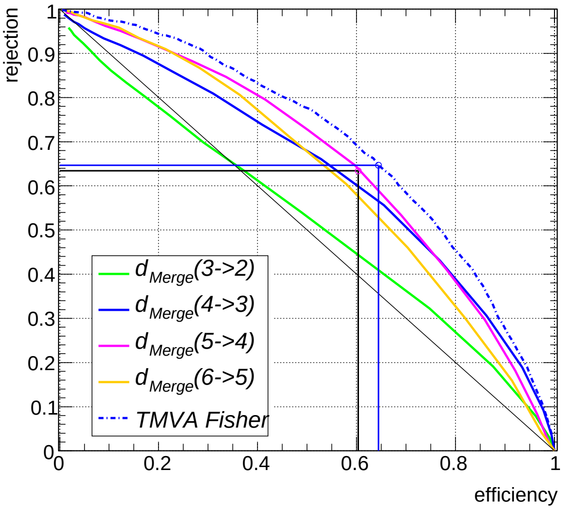


Figure 4.9: *The background rejection versus signal efficiency curve for the Fisher discriminant technique compared to the curves for cutting on single* d_{Merge} *variables.*

of the Fisher output variable $F(\boldsymbol{d})$ is shown in figure 4.8. Comparing to figure 4.7, one can see that the Fisher discriminant achieves more separation than individual requirements on a single d_{Merge} variable, although signal and background distributions still overlap substantially. The parameters and statistics that are used do not overtrain the discriminator. This is supported by a Kolmogorov-Smirnov probability of 0.975 (0.959) for the compatibility of the training and test output distributions on signal (background) events. This quantity measures the compatibility of two distributions based on both normalisation and shape by looking at the maximum difference between the two cumulative distribution functions of the test distributions [75]. The quoted numbers are the probabilities that the Kolmogorov-Smirnov test statistics exceeds the obtained value for two samples drawn from the same distribution. For distributions fulfilling this assumption, the result is uniformly distributed between 0 and 1 and can be directly used as confidence level, i. e. for a confidence level of 5 % the value must be above 0.05. Therefore, for both signal and background, training and test distributions are accepted as compatible.

Figure 4.9 presents the rejection versus efficiency curve—also known as Receiver-Output-Characteristics (ROC) curve—for this Fisher discriminant compared to several d_{Merge} variables from section 4.3. The Fisher discriminant achieves about 5 percent absolute higher background rejection at a given signal efficiency than using the most sensitive variable $\sqrt{d_{\mathrm{Merge}}(5 \to 4)}$. The absolute performance as given by the area under the ROC curve, although increased from 0.657 to 0.699, indicates that separating signal from background using d_{Merge} variables is difficult.

The Fisher discriminant is optimised in the same way as the single d_{Merge} requirement, by maximising efficiency×rejection. The resulting optimal value for the requirement on $F(\boldsymbol{d})$ is found to be practically zero (0.001). Setting $F(\boldsymbol{d})_{\mathrm{min}} = 0$ results in 63.5 % of signal being retained while 66.0 % of background is rejected. The fraction of physical background events over the whole mass range then drops to 21 %, and to 11.4 % within the signal region, compared to 23 % and 10.4 % for the requirement on the single d_{Merge} variable. The corresponding mass distribution for the $p_{\mathrm{T}}^{\mathrm{max}}$ three jet mass with this requirement

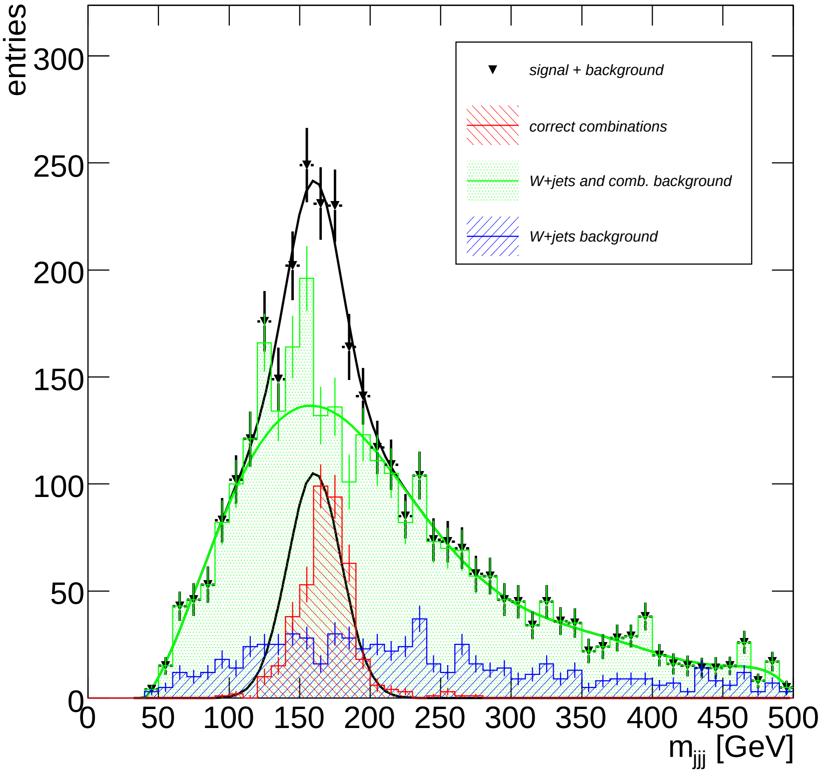
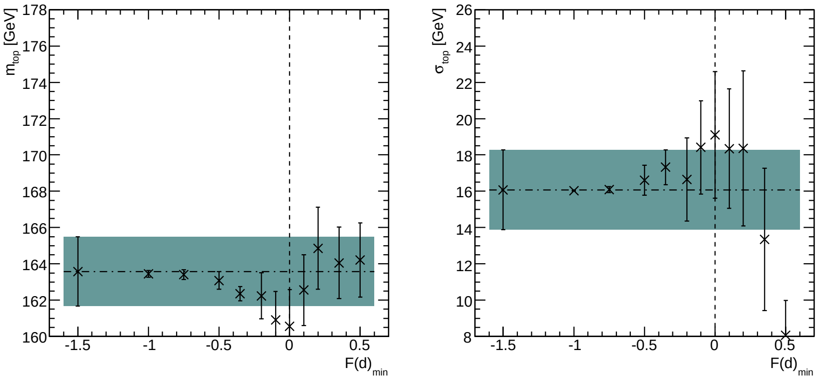


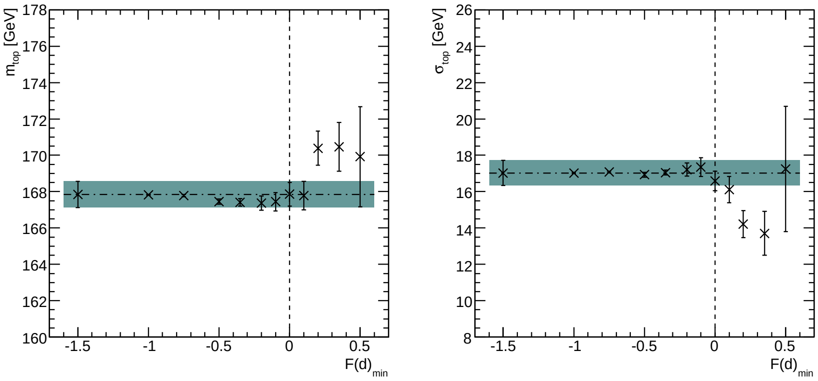
Figure 4.10: *Same as figure 4.1, but only events satisfying* $F(\boldsymbol{d}) > 0$ *are shown.*

applied is shown in figure 4.10. The slightly higher background fraction in the signal region compared to the requirement on a single d_{Merge} variable comes from the fact that the Gaussian fit does not approximate the correctly reconstructed mass distribution as well as in figure 4.5. To quantify this, the χ^2 between the Gaussian part of the fit to all data and the distribution of correct jet triplets is calculated by summing $(g(m_i) - n_i)^2/\sigma_{n_i}^2$ for all non-empty bins i, with n_i being the content of bin i, σ_{n_i} its uncertainty, and $g(m_i)$ the fitted Gaussian function at bin i. It is found that the value of $\chi^2/NDF = 144/18 = 8.0$ severely exceeds the χ^2/NDF for the full statistics $(77.6/20 = 3.9)$ or for putting a requirement on $d_{\mathrm{Merge}}(5 \to 4)$ $(64.9/19 = 3.4)$.

To test systematic changes in the reconstructed mass, the requirement on the discriminator output is scanned from -1.5 to 0.5. The result on the fitted Gauss mean and width to the summed “data” distribution can be seen in figure 4.11a. Figure 4.11b shows the mean and width of a Gauss fit to only those events where the $p_{\mathrm{T}}^{\mathrm{max}}$ method is able to find the correct three jets stemming from the top quark decay. The uncertainties are calculated in the same way as for figures 4.6a and 4.6b, and for ease of comparison, the vertical axes are also kept the same. These figures show that cutting on the Fisher discriminator output does not introduce as much systematic mass shifts for the correct triplets as cutting on a single d_{Merge} value does. Apparently the linear combination of these variables shows a lower correlation to the mass of the hadronically decaying top quark. While the mass of the correct combinations stays stable, the mass reconstructed from the fit to all “data” shows lower values for requirements around the optimal requirement of $F(\boldsymbol{d})_{\min} = 0$.

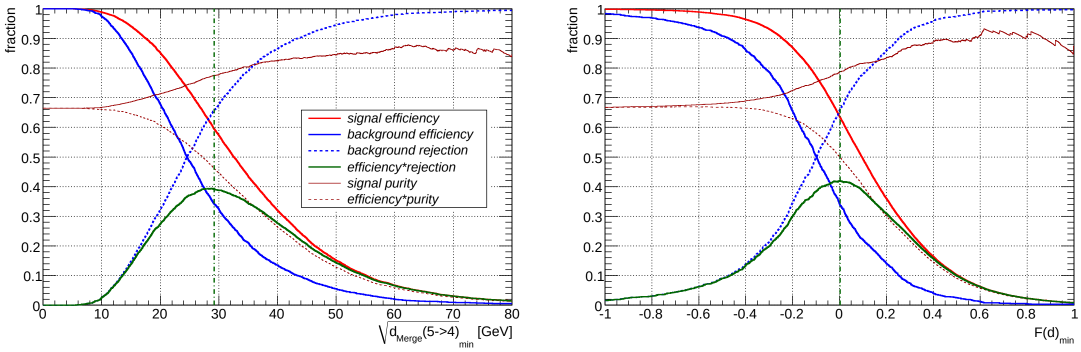


(a) reconstructed m_t, i. e. Gauss part of the Chebychev+Gauss fit



(b) mass from Gauss fit to correct jet triplets

Figure 4.11: *Reconstructed top quark masses and their means when increasing the requirement on $F(\boldsymbol{d})$, analogue to figure 4.6. The dotted vertical lines indicate $F(\boldsymbol{d})_{\text{min}} = 0$.*



(a) requirement on $d_{\mathrm{Merge}}(5 \rightarrow 4)$ (b) requirement on $F(\boldsymbol{d})$

Figure 4.12: *Efficiency, rejection and purity curves for suppression of* W*+jets background.*

This explains the worse compatibility of the Gaussian curve and the distribution of correct combinations found above. In figure 4.10, the combinatorial background exhibits a slightly stronger peak which the Chebychev polynomial is not able to follow, so that the Gaussian is representing more events on the left side of the correct jet triplets. While not being statistically significant, this effect needs to be scrutinised before the Gaussian mean value is used as a direct measure of the top quark mass, to see whether a bias and/or a systematic uncertainty need to be included in the mass reconstruction method. However, for analyses using other means to extract the top quark mass from the obtained distribution, notably template-based analyses, the change of combinatorial background should be tolerable within the method.

By placing a requirement on the Fisher discriminator output, the signal significance, as defined via $S/\sqrt{(S+B)}$, can hardly be increased. This is because of the high signal to background ratio of 1.98 already without the additional selection. However, the signal purity can be increased from initially 67 % up to about 95 %, if about 95 % of signal statistics is sacrificed. Figure 4.12 shows a summary of efficiency, purity and rejection numbers for the requirement on the Fisher discriminator output compared to the single d_{Merge} variable. The signal (background) efficiencies are given as red (blue) curves, and the background rejections as the dotted blue curves. The slim red lines denote the signal purities and efficiency×purity numbers (dotted), with the initial numbers of 4368 and 2205 signal and background events assumed. The green curves represent the efficiency×rejection numbers and the vertical lines denote the optimal requirements that are used for the two methods. The figure shows that the Fisher method can achieve slightly higher signal purities and signal×efficiency compared to the requirement on d_{Merge}.

4.4 Using d_{Merge} variables to suppress wrong combinations

The two presented methods will not drastically improve the top quark mass reconstruction. The dominant background within a $p_{\mathrm{T}}^{\mathrm{max}}$-based top mass analysis after the current event selection and a requirement on the Fisher discriminant output is composed of the incorrectly assigned jet triplets from the signal sample. These still contain some top quark mass information and peak in the region of the simulated top quark mass. Thus they are capable of shifting the Gaussian peak away from the true top quark mass. Therefore, it is

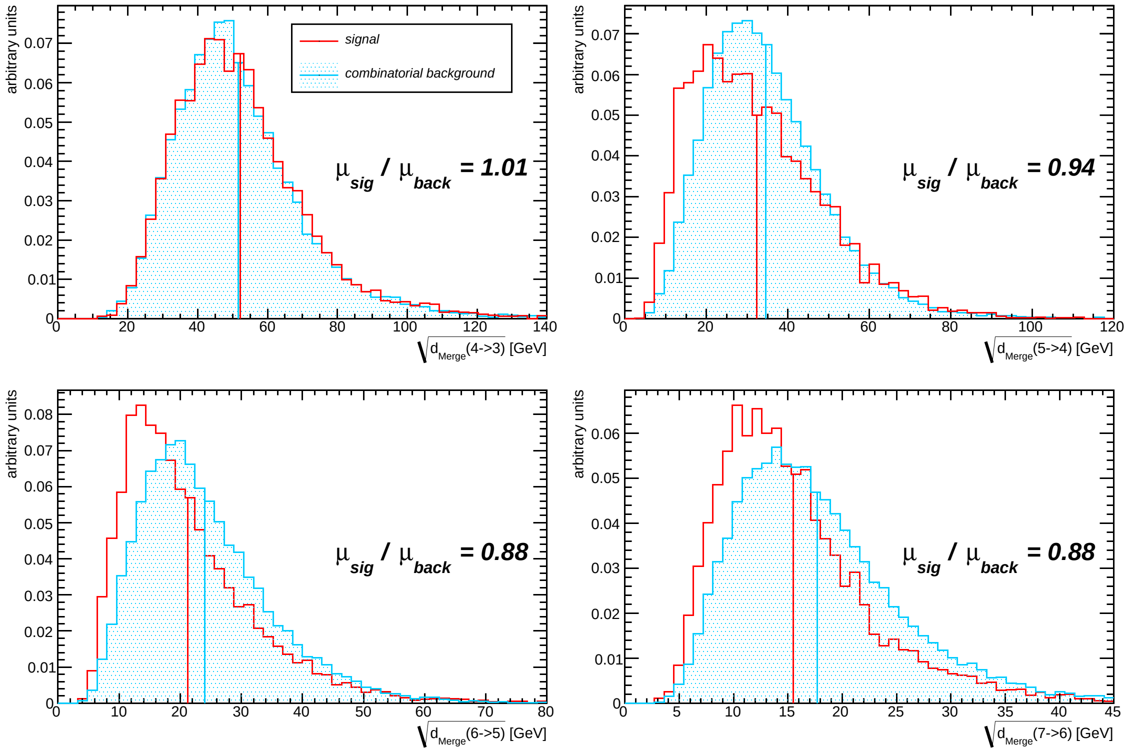


Figure 4.13: *Normalised distributions of four* d_{Merge} *variables within the* $t\bar{t}$ *sample. The red histograms depict the correctly found jet triplets, the light blue histograms depict the distributions from combinatorial background.*

also tested whether the d_{Merge} variables can improve the rejection of combinatorial background. Figure 4.13 shows the four d_{Merge} variables with the highest separation power, where the red histograms denote those events where the $p_{\mathrm{T}}^{\mathrm{max}}$ method was able to find the correct jet triplet, and the light blue histograms contain the incorrectly assigned events. The histograms are normalised to unit area to compare their shapes, showing that some discrimination between these two classes of events is possible. Apparently, the $p_{\mathrm{T}}^{\mathrm{max}}$ algorithm is able to find the correct combination more likely within certain types of event topologies. This knowledge could be used to suppress wrongly associated combinations. However, the mean values, denoted by the vertical lines in figure 4.13, are not very different and their order is inverted compared to the signal to physical background rejection case. For d_{Merge} variables above $d_{\mathrm{Merge}}(4 \rightarrow 3)$, the mean values of the combinatorial background distributions are higher than for correctly assigned top quark events, in contrast to the physical background, where they are lower than for the signal distributions (c. f. figure 4.3). This means that the simultaneous training of a Fisher discriminant against both types of background is unfavourable and a two-step procedure would be needed. This is visualised in figure 4.14, where the red histograms denote the correctly reconstructed jet triplets, while the green histograms display the sum of combinatorial and physical background, with the proper scaling according to the same luminosity taken into account. The two distributions are then normalised to unit area. It is seen that the separation for these four most discriminating variables from figures 4.3 and 4.13 is rather poor. The variable $d_{\mathrm{Merge}}(5 \rightarrow 4)$, which is used for W+jets background rejection, achieves nearly no

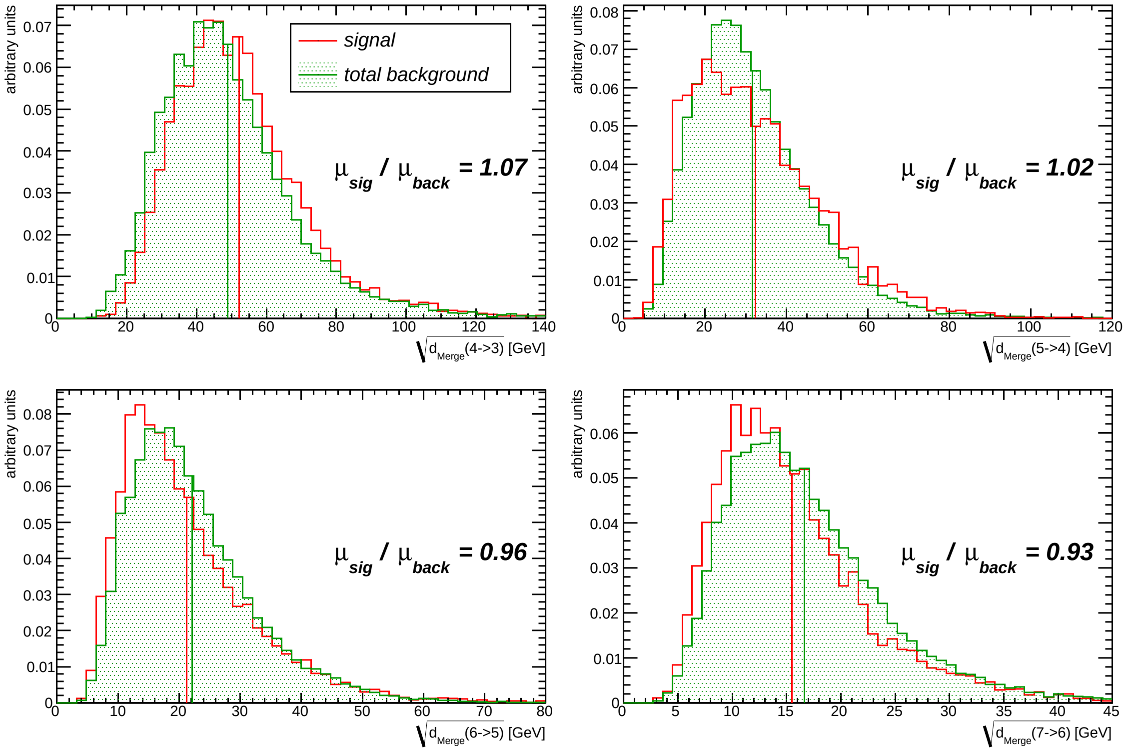


Figure 4.14: *Normalised distributions of four* d_{Merge} *variables. The red histograms depict the correctly found jet triplets within the signal samples, the green histograms depict the distributions from background, containing wrongly associated jet triplets and W+jets processes.*

separation at all, since the lower d_{Merge} values for physical background are cancelled by the larger d_{Merge} values for combinatorial background. On the other hand, a two-step rejection procedure which discriminates against physical background and afterwards against combinatorial background will severely decrease the signal statistics. Due to these findings, the d_{Merge} event shape variables are not used for discrimination against combinatorial background, and other means of suppressing combinatorial background are looked for in the following.

4.5 Background discrimination using kinematic variables

To evaluate further means of background suppression, several kinematic variables are examined for their potential to discriminate against W+jets background. A selection of them is shown in figure 4.15. For each event, the values of the variables are filled into one of three histograms: The red histograms contain $t\bar{t}$ l+jets signal events where the $p_{\mathrm{T}}^{\mathrm{max}}$ method found the correct jet triplet, the light blue histograms contain the rest of the l+jets events, i. e. combinatorial background, and the blue histograms contain W+jets events. All histograms are filled with events corresponding to $146\,\mathrm{pb}^{-1}$ of luminosity. Several invariant masses and ΔR distances and angles are tested and only the most prominent in separation are shown. These are:

- $m(b_l l)$ (figure 4.15a): the invariant mass of the candidate b-quark jet stemming from the leptonically decaying top quark and the chosen lepton. This b-quark jet candidate is taken as the highest-p_{T} jet not used for the hadronic top quark candidate. Due to kinematical reasons, this variable has an upper limit given by $m(b_l l) \leq \sqrt{m_t^2 - m_W^2} = 152.6\,\mathrm{GeV}$, which is seen as a sharp kink in the distribution of the correct triplets. Due to resolution effects and the sometimes incorrect assignment of the leptonic b-quark jet, the reconstructed mass can be larger than the theoretical limit.

- $m(tb_l)$ (figure 4.15b): the invariant mass of the hadronic top quark candidate and the leptonic b-quark candidate.

- ΔR_{max} (figure 4.15c): the largest of the three ΔR present in the three hadronic top quark candidate jets.

- ΔR_{min} (figure 4.15d): the smallest of the three ΔR present in the three hadronic top quark candidate jets.

- m_W (figure 4.15e): the W mass, reconstructed via the ΔR_{min}-method. This method chooses two candidate jets from the selected jet triplet to represent the hadronically decaying W by taking the pair that has the smallest ΔR in the rest frame of the top quark candidate. This is motivated by the fact that for a top quark decay the W and b must be oriented back-to-back in the top quark rest frame, i. e. have a ΔR of π. The correctly chosen jet triplets peak around the known W mass, whereas the backgrounds have a large proportion of lower masses. The combinatorial background has a second peak around the correct mass, since it is possible to select the two correct jets from the W decay, while missing the third jet from the top quark decay, therefore creating an entry in the combinatorial background. Despite the shape differences, the mean values of the distributions are similar, so that a Fisher discriminant needs a transformation of this variable to make it useful for discrimination.

- Δm_W (figure 4.15f): the absolute difference between the reconstructed W mass and its current world average value, i. e. $\Delta m_W = |m_W - 80.4\,\text{GeV}|$. This transformation makes the observable useful for discrimination using a Fisher discriminant. While m_W is correlated to the reconstructed m_t, the correlation coefficient for Δm_W and m_t is almost zero. Therefore Δm_W is used instead of m_W as discriminating variable.
- $\Delta R(Wb_h)$ (figure 4.15g): the ΔR between the hadronic W candidate reconstructed by the ΔR_{min}-method and the third jet in the hadronic top quark candidate.

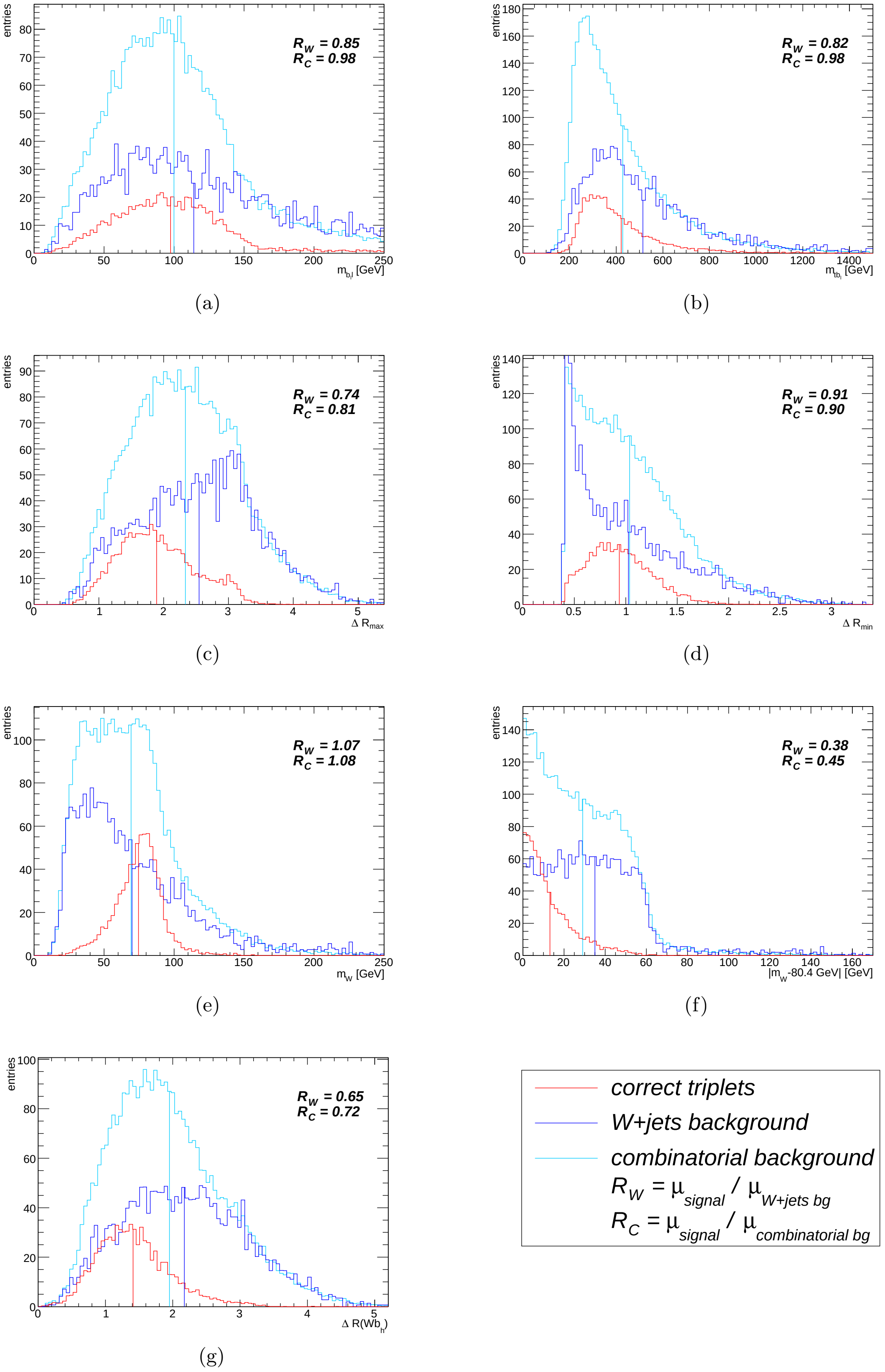
Some of these variables possess the convenient feature that they are also discriminating against combinatorial background, namely ΔR_{max}, Δm_W, and $\Delta R(Wb_h)$. The matrix of linear correlation coefficients between the seven investigated variables and the reconstructed top quark mass is shown in table 4.4 for the events where the correct jet triplets were selected within the signal sample. The two variables $m(tb_l)$ and m_W are correlated with m_t, which can be understood by the participation of at least two of the $p_{\text{T}}^{\text{max}}$-selected jets. The variable Δm_W removes the m_W-m_t correlation almost completely. The ΔR-variables are only marginally correlated with m_t. Apart from that, the highest correlations can be found amongst the ΔR-variables themselves. The largest absolute correlation coefficient is 0.76 between ΔR_{max} and $\Delta R(Wb_h)$, so that it is conceived that each variable will contribute to the separation power of a combined method.

Table 4.4: *Matrix of linear correlation coefficients of the selected variables and m_t for the events where the correct jet triplets were selected.*

	m_t	$m(b_l l)$	$m(tb_l)$	ΔR_{max}	ΔR_{min}	m_W	Δm_W	$\Delta R(Wb_h)$
m_t	1							
$m(b_l l)$	0.07	1						
$m(tb_l)$	0.26	0.26	1					
ΔR_{max}	-0.05	-0.11	-0.41	1				
ΔR_{min}	-0.05	-0.09	-0.39	0.55	1			
m_W	0.44	0.04	0.14	0.01	0.20	1		
Δm_W	-0.07	0.01	-0.02	-0.08	-0.22	-0.39	1	
$\Delta R(Wb_h)$	-0.02	-0.11	-0.39	0.76	0.40	-0.12	0.02	1

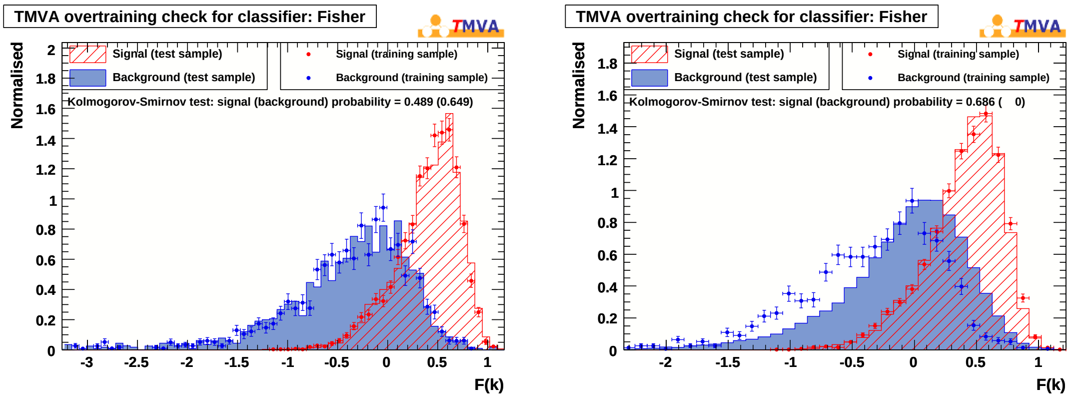
A Fisher discriminant is then trained upon the six chosen variables, hereafter denoted as $\boldsymbol{k}$, to separate correct jet triplets from W+jets background. Combinatorial background events are initially left out in the training process to not spoil the separation against physics background events. Half of the events available are used for training, i. e. about 1500 events for W+jets background and about 5300 events for the correct combinations. The Fisher discriminant method exhibits a good separation between signal and background, as seen in the distribution of discriminator output values shown in figure 4.16a. The Kolmogorov-Smirnoff-probabilities indicate compatibility between the test events and the training events regarding $F(\boldsymbol{k})$.

The most separating variables for the Fisher method are Δm_W, $\Delta R(Wb_h)$ and ΔR_{max}, followed by $m(tb_l)$, $m(b_l l)$ and ΔR_{min}. This is expected from the sizes of relative mean



(a) (b) (c) (d) (e) (f) (g)

Figure 4.15: *Several kinematic variables with possible separation power between signal and background. The ratios of the mean values of the distributions, denoted as vertical lines, are printed in each figure.*



(a) tested on physics background only

(b) tested on both physics and combinatorial background

Figure 4.16: *The discriminator output of the Fisher method when trained on the six kinematic variables. The filled histograms denote the test sample output, whereas the points with uncertainty markers denote the training sample output. The training process involves correct jet triplets as signal and W+jets events (but not combinatorial events) as background.*

value differences present in the input distributions. The area under the ROC curve is 0.898 when calculated without combinatorial background events. An advantage is that the highest ranked variables also show the largest differences between the correct triplet and combinatorial background distributions, which makes the method also efficiently reject this background. For an estimation of the rejection performance including combinatorial background, those background events are added to the testing events after the training process. By this, the ROC area drops to 0.805, which indicates a still good separation between signal and both of the backgrounds. Bearing in mind that the amount of combinatorial background events is more than 50 % larger than the amount of W+jets events and even almost 3 times as high in the signal peak region, this is an encouraging result.

The response of the Fisher discriminant is shown in figure 4.16b with combinatorial events included in the testing sample. For the background, the training and testing outputs are no longer compatible, since they now contain different event sources. The output distribution for background testing events moves to higher values compared to training events, i. e. the addition of combinatorial background makes the combined background more signal-like. On the other hand, there is still enough separation between the output for signal and background left. The optimal working point is again found by maximising efficiency×rejection, and the resulting requirement is $F(\boldsymbol{k}) > 0.14$ when considering W+jets background only. For the optimal separation against both backgrounds, the working point moves to a slightly higher value, to $F(\boldsymbol{k}) > 0.23$. The reconstructed top quark mass distribution with this requirement applied is shown in figure 4.17. There it is seen that the discriminator indeed suppresses not only W+jets background events, but is also sensitive to wrong choices of the $p_{\mathrm{T}}^{\mathrm{max}}$ method. The right tail of the mass distribution is suppressed stronger than the region around the peak and the correct triplets nearly compose half of the peak seen in the total "data". For the looser requirement, the fraction of physics background events is reduced to 16.2 % (7.9 %) over the whole mass range (in the peak region). Due to the suppression of combinatorial events, the fraction of total

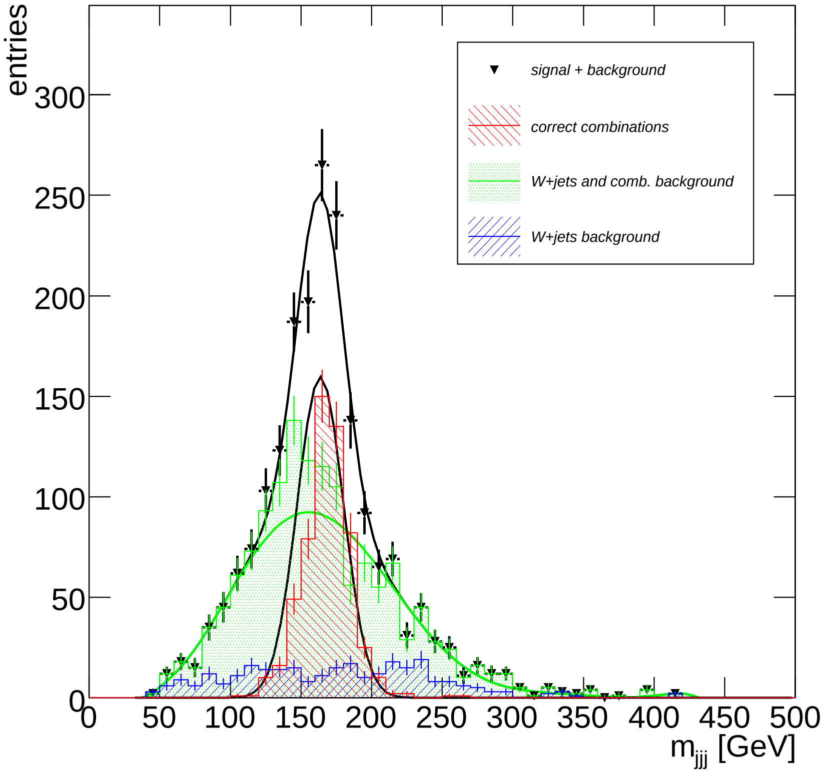
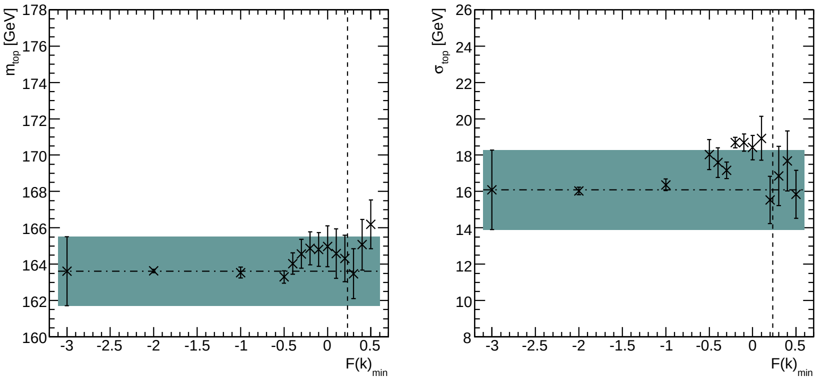
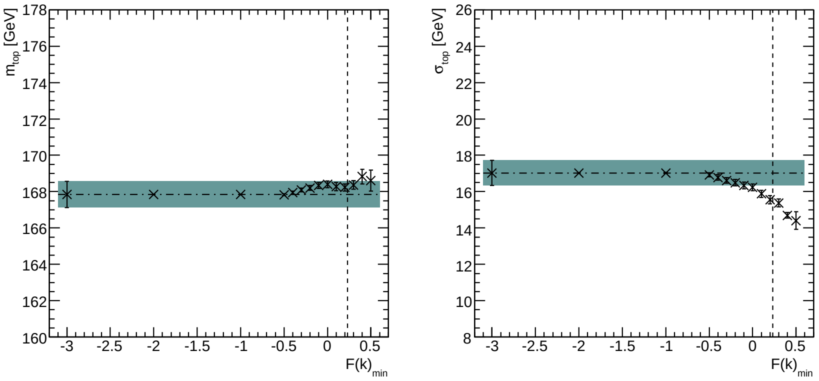


Figure 4.17: *Same as figure 4.1, but only events satisfying* $F(\boldsymbol{k}) > 0.23$ *are shown.*



(a) reconstructed m_t, i. e. Gauss part of the Chebychev+Gauss fit



(b) mass from Gauss fit to correct jet triplets

Figure 4.18: *Reconstructed top quark masses and their means when increasing the requirement on $F(\boldsymbol{k})$, analogue to figure 4.6. The dotted vertical lines indicate $F(\boldsymbol{k})_{\mathrm{min}} = 0.23$.*

background also drops, from 88 % (67.7 %) to 74 % (54.6 %) over the whole mass range (in the peak region). For the tighter requirement, the corresponding fractions of total background are 71 % (51.8 %).

The dependence of the top quark mass upon the discriminator output is tested by varying the requirement on $F(\boldsymbol{k})$ and reconstructing m_t. Figures 4.18a and 4.18b show the results for the fit to the reconstructed mass distribution and to the correctly assigned jet triplets, respectively. It is seen that the reconstructed mass stays very stable for $F_{min} \leq -0.5$. This is understandable since almost only background events are rejected up to this point. For tighter requirements, the reconstructed as well as the correct mass shift to higher values. The absolute values of the shifts of about 1 GeV and 0.5 GeV are still compatible with the initial mass within the statistical uncertainty. However, the uncertainties attached to the correct combinations indicate some effect with a systematical uncertainty that is smaller than the statistical uncertainty. With increasing requirement, the distribution of correct jet triplets becomes systematically slimmer, as seen from the right part of figure 4.18b. This selection method is a good candidate to be used as an extension to the template-based mass reconstruction [76] that is under development at the moment at the MPP top group. For the template method, the systematic uncertainty on the mass of the correct triplets is less relevant, since the mass is calibrated against the

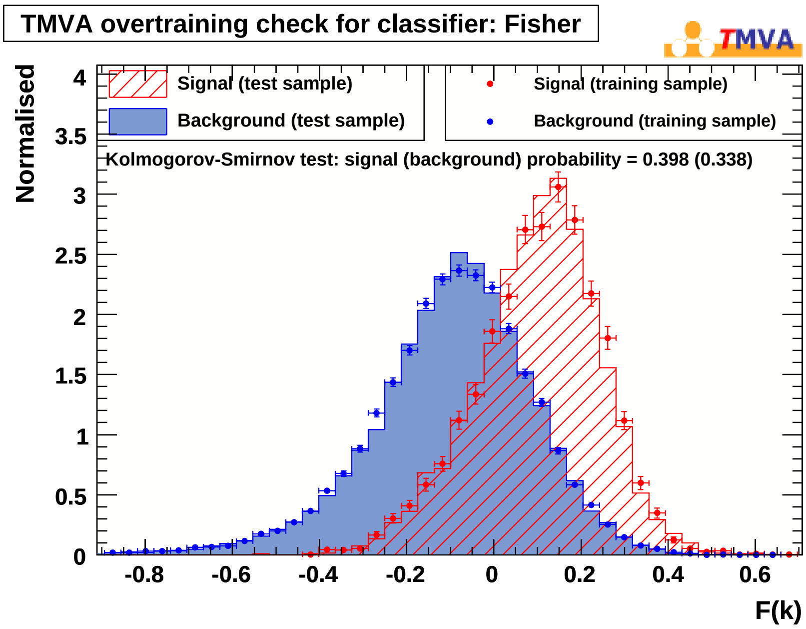


Figure 4.19: *Same as figure 4.16, but with combinatorial background events added to the test sample.*

Monte Carlo simulated top quark mass.

Since it is seen that combinatorial background events are distributed like signal events only for the three lower ranked variables, but are more background-like for the high ranked variables, another discriminant is trained, using the combinatorial events as additional background in the training process. Half of the available statistics is used for training, i. e. 5300 signal and about 29600 background events. The output from the Fisher discriminant is shown in figure 4.19 for training and testing events. The achieved separation power is about the same as in figure 4.16b, also indicated by a similar area under the ROC curve of 0.809. A summary of the three ROC curves is shown in figure 4.20, with the red curve belonging to the separation of correct combinations from W+jets background, the light blue curve using the same Fisher discriminant, but also testing the rejection against combinatorial background events, and the green curve showing the discrimination power for the discriminant trained against both backgrounds simultaneously. The latter two curves are almost overlapping, which means that training against only physics background also nearly optimally rejects combinatorial background. The optimal requirement on this Fisher discriminant is found to be $F(\boldsymbol{k})_{\mathrm{min}} = 0.02$, and the reconstructed top quark mass distribution with this requirement applied is shown in figure 4.21. It is seen that the high mass regime is not as fully suppressed as with the previous discriminant, while in the signal region the Gaussian part of the fit is an equally good description of the distribution of correct jet triplets ($\chi^2/NDF = 59.1/16 = 3.7$ versus $62.7/15 = 4.2$ in the previous case). The fraction of background that is rejected is 75.7 % at a signal efficiency of 74.1 %. This results in a physics background fraction of 19.1 % (7.3 % under the signal peak) and a total background fraction of 71.5 % (50.1 %), which is close to the values found for the previous case without combinatorial background events in the training sample.

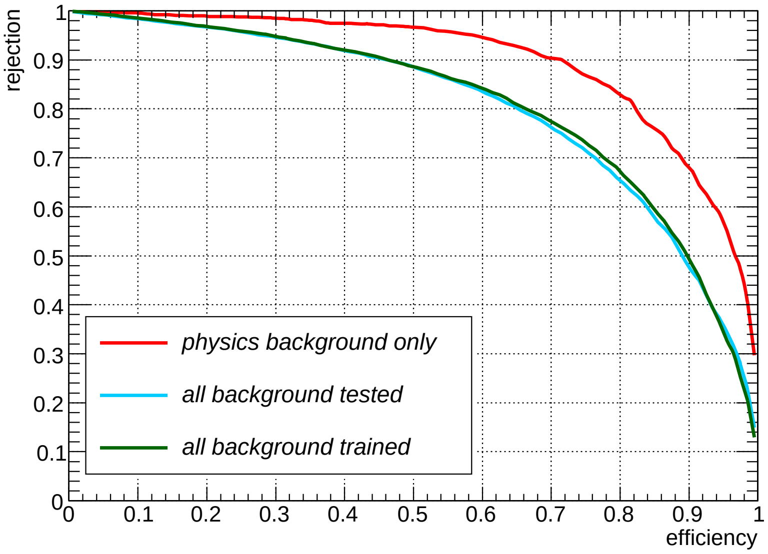


Figure 4.20: *The background rejection versus signal efficiency curves for the Fisher discriminant trained upon kinematic variables. The red curve reports the separation of the correct triplets against* W+*jets events only, while the light blue curve uses also combinatorial background for evaluating the background rejection. The green curve denotes the use of combinatorial background also for training the discriminant.*

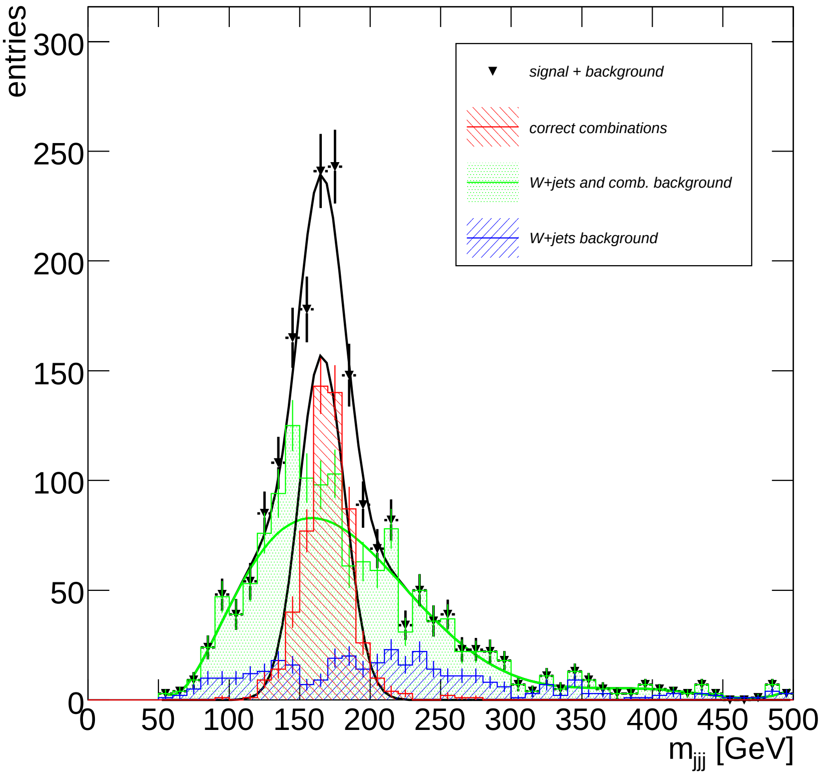
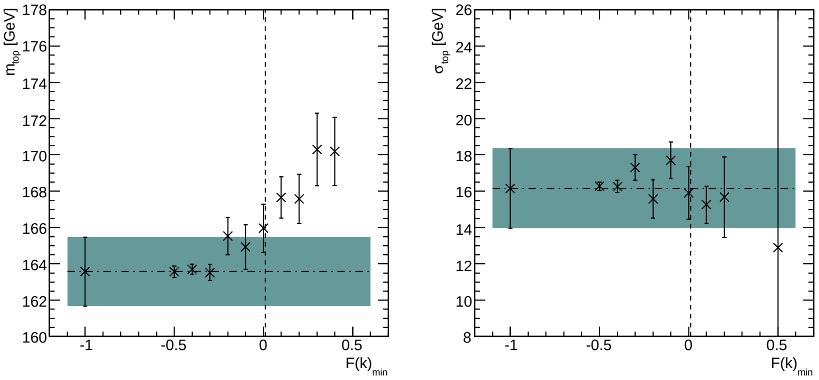
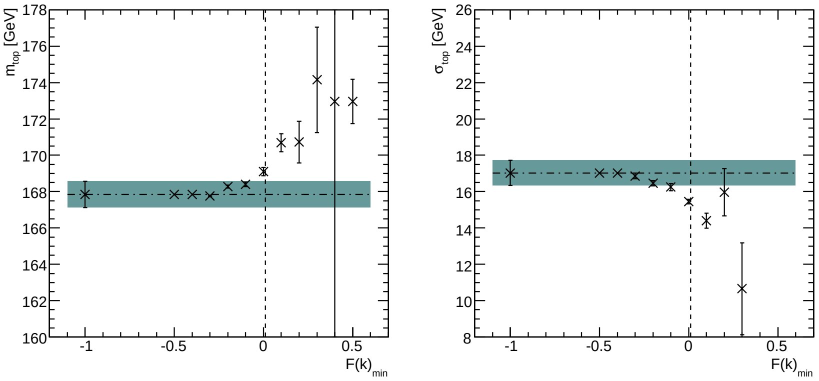


Figure 4.21: *Same as figure 4.1, but only events satisfying* $F(\boldsymbol{k}) > 0.02$ *for a Fisher discriminant trained on combinatorial and physics background are shown.*



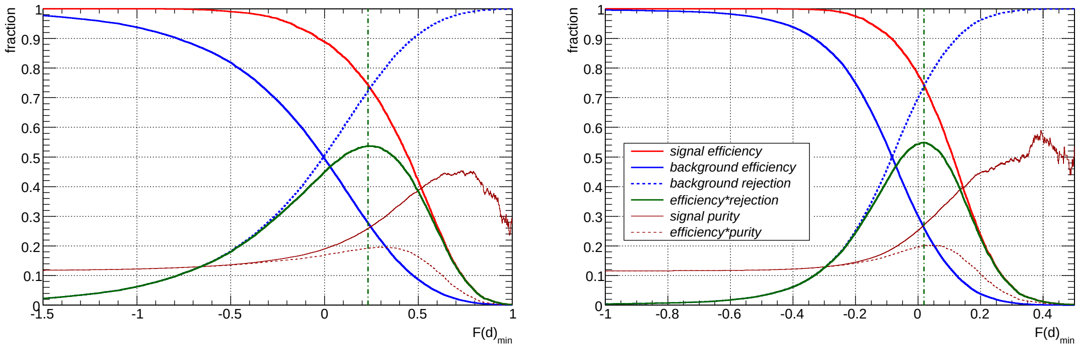
(a) reconstructed m_t, i. e. Gauss part of the Chebychev+Gauss fit



(b) mass from Gauss fit to correct jet triplets

Figure 4.22: *Reconstructed top quark masses and their means when increasing the requirement on $F(\boldsymbol{k})$ trained on both combinatorial and physics background, analogue to figure 4.6. The dotted vertical lines indicate $F(\boldsymbol{k})_{\text{min}} = 0.02$.*

The effects of the requirement on the reconstructed mass are shown in figure 4.22. Increasing the requirement on $F(\boldsymbol{k})$ has a high influence on the reconstructed mass (in figure 4.22a) as well as the mass of the correct triplets (in figure 4.22b). Adding the combinatorial background in the training process results in the training data being dominated by combinatorial events, which introduces a correlation between the discriminator output and the reconstructed mass in the correct jet triplets. Additional investigation could reveal whether and how the top mass dependence of the combinatorial background produces this effect. Using $t\bar{t}$ samples simulated with different top quark masses might reduce the dependence of $F(\boldsymbol{k})$ on the top quark mass. For now, only training on physics background and applying the resulting discriminant to all data is preferred. In figure 4.23, the achievable efficiency, rejection and purity numbers are summarised for both Fisher discriminants. It is seen that the discriminant trained on W+jets background only (figure 4.23a) achieves about the same maximum efficiency×rejection and efficiency×purity numbers as the discriminant trained on all background (figure 4.23b), so that its use does not diminish the power of the method.



(a) requirement on $F(\boldsymbol{k})$ trained on only W+jets background

(b) requirement on $F(\boldsymbol{k})$ trained on both backgrounds

Figure 4.23: *Efficiency, rejection and purity curves for suppression of W+jets and combinatorial background.*

4.6 Conclusions

The presented top quark mass reconstruction analysis uses solely information from the calorimeter to measure the top quark mass. The l+jets channel of top quark pair decays is used, with the events being tagged by the leptonically decaying top quark, and the top quark mass being reconstructed from the hadronically decaying top quark via the $p_{\mathrm{T}}^{\max}$ method. A simple and robust method is in place that uses no b-quark jet identification and is intended to be used for the first data that will be obtained from the ATLAS detector once the LHC starts. The major sources of systematic uncertainty in this analysis are the jet energy scale and the fitting procedure used to extract the mass value.

An analysis to further reduce the dominant source of physical background, the W+jets process, was presented. It tries to exploit the d_{Merge} observables, obtained by running a k_{T} jet algorithm in the exclusive mode. It was shown that a separation of $t\bar{t}$ l+jets signal from W+jets background is possible at the expense of signal statistics. The additional background rejection on top of the standard event selection is 63.4 % at a signal efficiency of 60.3 %, where the required minimal d_{Merge} value was optimised to maximise the product of the two fractions. Placing selection requirements on the $d_{\mathrm{Merge}}(5 \longrightarrow 4)$ variable, however, introduces a systematic bias on the top quark mass that one wants to reconstruct, since harder jet topologies are systematically selected.

A multivariate technique, the Fisher discriminant, was tested for its performance on the same problem. It uses several d_{Merge} variables together and achieves a still higher additional signal to background separation. At the same time, the dependence of the top quark mass on the output of the discriminator is much reduced. However, the combinatorial background has to be well represented by the Gauss+polynomial fit procedure if the mean of the Gauss curve is identified with the reconstructed mass, which is not the case with the current Chebychev polynomial fitting. Therefore, the obtained mass bias has to be accounted for. The optimal background rejection in terms of efficiency×rejection for this method is 66.0 % at a signal efficiency of 63.5 %.

Finally, several kinematic variables were identified to be discriminating against the W+jets backgrounds as well. To be able to use them for signal selection, another Fisher discriminator was trained and applied to the available data samples. Since part of the chosen variables also discriminate against wrong choices of the jet association to the hadronically decaying top quark, the combinatorial background events are not treated as belonging to the signal in the training process. This method is able to suppress 87.1 % of W+jets background at an efficiency of 74.8 % for the correctly assigned jet triplets. Moreover, also 69.4 % of combinatorial background events are rejected at the same time. The overall signal to background ratio is much enhanced and this selection presents a promising improvement also as an input to more advanced methods of mass reconstruction, like the template method that is prepared within the MPP top group. By using the good separation against wrongly associated jet triplets and therefore including combinatorial background in the training process, the achievable separation power stays at the same level at the cost of a top quark mass dependency of the discriminator output.

Further developments could lead to additional rejection of combinatorial background by exploiting the discrimination of kinematic variables against wrong jet triplets. Instead of rejecting wrong choices of the $p_{\mathrm{T}}^{\mathrm{max}}$ method, a combination of these variables can be used to judge on the most likely correct jet triplet, therefore increasing the signal to background ratio without diminishing the available event statistics. It remains to be seen whether such a method can achieve the correct jet association fraction that e. g. kinematical fit-based methods provide [77].

Eventually, the use of b-quark jet identification methods will improve the selection process of $t\bar{t}$ l+jets pair decays both against physical background as well as wrong jet combinations. This will need a firm understanding and an excellent calibration of the ATLAS detectors, especially the Inner Detector tracking devices.

Chapter 5

Alignment studies for the silicon part of the ATLAS Inner Detector

The accurate calibration of the tracking devices of a particle detector also involves the exact determination of the positions of all detector modules in space, termed *alignment*, to provide the best possible track parameter resolution for charged particles, as well as the reliable reconstruction of vertex multiplicities and positions. The latter is especially beneficial for identification of b-quark induced jets, which is needed for high suppression of physical and combinatorial background in many physics analyses, including those involving top quarks. At MPP, an alignment algorithm for the pixel and SCT detectors was developed and applied to testbeam data from the 2004 Combined TestBeam (CTB) and cosmic ray data from the 2008 cosmic data taking [53]. The MPP algorithm is called *local χ^2 alignment algorithm* and is described in detail in [78, 79]. Here only a brief review of the procedure is given. Afterwards, the creation of semi-realistic global systematic deformations of the Inner Detector is described. These deformation sets are used in the alignment and physics groups of ATLAS to measure the performance of the alignment procedures and test the impact of a remaining systematic misalignment to the physics performance of the detector. The application of the local χ^2 alignment algorithm using the cosmic ray data taken in autumn 2008 is presented. Additions and improvements to the method are highlighted.

5.1 The local χ^2 alignment algorithm

5.1.1 Track-based alignment

Track-based alignment procedures are used to accurately determine the positions of all detector elements of a tracking detector. As opposed to external alignment methods, like survey measurements or length-monitoring methods for the support structures, they work by using the hits from the recorded detector data themselves. Just as the laser-based system that is in place to monitor the geometry of the SCT system [80], track-based alignment will be performed continuously during running of the accelerator. Track-based methods are used to obtain the final precision that can be achieved on the determination of the actual position of each detector element, after a sufficient external base alignment allowing a reliable track reconstruction has been achieved.

All track-based alignment methods work by minimising *track residuals* on the given data sample with respect to the *alignment parameters* under consideration. A track residual r is defined as the distance between the measurement point on a detector module and the corresponding track extrapolation onto the surface of the module. On a perfectly aligned detector, all residuals yield Gaussian distributions centred around zero with a width determined by the sensor resolution and tracking uncertainties. Shifted modules usually possess a shifted residual distribution, which gives a handle on the real position of the detector element. Once a correction on the position is obtained, the geometry of the detector in the reconstruction software is updated and future reconstruction of the data takes the calculated positions and irregularities of the detector elements into account. It is important to note, that centred residuals are required, but not sufficient, for a perfect alignment knowledge. As will be discussed later, certain detector deformations exist that still result in centred residuals.

In general, residuals depend on the alignment parameters $\boldsymbol{a}$ of all modules (i. e. the corrections on the position of each module) as well as on the track parameters $\boldsymbol{\pi}$ of the track that created the residual: $r = r(\boldsymbol{a}, \boldsymbol{\pi})$.

5.1.2 The local χ^2 formalism

The local χ^2 algorithm works by minimising track residuals on all modules using all tracks in the input dataset. It does so by using a χ^2-based method, where the total χ^2 to be minimised is given by the sum of all track residual χ^2:

$$\chi^2(\boldsymbol{a}, \boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_m) = \sum_{i \,\in\, tracks}^{m} \boldsymbol{r}_i{}^T(\boldsymbol{a}, \boldsymbol{\pi}_i) \cdot V_i^{-1} \cdot \boldsymbol{r}_i(\boldsymbol{a}, \boldsymbol{\pi}_i) \quad . \tag{5.1}$$

Here, $\boldsymbol{a}$ denotes the vector of all alignment parameters of all modules, $\boldsymbol{\pi}_i$ is the vector of track parameters of track i, $\boldsymbol{r}_i$ denotes the vector of all residual measurements on the modules hit by track i and V_i is the covariance matrix of the measurements of track i.

The minimising condition is then formulated as

$$\frac{d\chi^2(\boldsymbol{a})}{d\boldsymbol{a}} = 0 \quad . \tag{5.2}$$

By using a linear assumption on the dependency of the χ^2 on the residuals, this yields the solution for the alignment parameter corrections:

$$\begin{aligned} \Delta\boldsymbol{a} &= -\left(\sum_{i \in tracks} \left(\frac{d\boldsymbol{r}_i}{d\boldsymbol{a}}\right)^T \cdot V_i^{-1} \cdot \left(\frac{d\boldsymbol{r}_i}{d\boldsymbol{a}}\right) \right)^{-1} \cdot \left(\sum_{i \in tracks} \left(\frac{d\boldsymbol{r}_i}{d\boldsymbol{a}}\right)^T \cdot V_i^{-1} \cdot \boldsymbol{r}_i \right) \qquad (5.3) \\ &\equiv -A^{-1}\boldsymbol{b} \quad . \end{aligned}$$

The alignment solution $\Delta\boldsymbol{a}$ is a vector of $6N$ dimensions, where N is the number of detector structures to be aligned. The matrix A is a $6N \times 6N$ matrix determined by the derivatives of the residuals with respect to the alignment parameters of the detector structures and the covariance matrix of the residuals. The vector $\boldsymbol{b}$ is a vector of $6N$ dimensions. Complete derivations of the solution can be found in [81, 82].

For the silicon part of the ATLAS Inner Detector, computing the solution means therefore solving this system of linear equations with 34992 degrees of freedom (6 degrees of freedom of a rigid body for each of the 5832 silicon modules). The global χ^2 algorithm, which is also in use for the ATLAS Inner Detector, does so by using a dedicated computer cluster for a full inversion of the matrix or fast solving techniques to calculate the alignment solution.

The local χ^2 method circumvents the complicated solving of the multi-dimensional system of equations by ignoring the interdependence of residuals that lie on the same track but on different modules. This is a consequence of the assumptions

- that alignment parameters of other modules on the same track do not influence the residual on the module under consideration (this gives the local approach its name);
- that measurements on different detector surfaces are uncorrelated (this neglects multiple Coulomb scattering effects and makes the covariance matrix of the residuals diagonal);
- and that the contribution of the residual derivatives with respect to the track parameters is negligible compared to the size of the derivatives with respect to the module alignment parameters.

To be able to make the last assumption, the local χ^2 approach uses *unbiased residuals*, i. e. residuals where the hit under consideration has been removed from the track fit. The uncertainty assigned to such a residual is given by the quadratic sum of measurement and track extrapolation uncertainties.

These three assumptions lead to a replacement of the total derivatives with partial derivatives in the solution, and the matrix becomes 6×6-block-diagonal. Therefore the local χ^2 algorithm can solve the minimisation for each module separately and the solution reads:

$$\begin{aligned}\Delta\boldsymbol{a}_k &= -\left(\sum_{i\in hits}\frac{1}{\sigma_{ik}^2}\left(\frac{\partial r_{ik}}{\partial\boldsymbol{a}_k}\right)^T\left(\frac{\partial r_{ik}}{\partial\boldsymbol{a}_k}\right)\right)^{-1}\cdot\left(\sum_{i\in hits}\frac{1}{\sigma_{ik}^2}\left(\frac{\partial r_{ik}}{\partial\boldsymbol{a}_k}\right)^T r_{ik}\right) \\ &\equiv -A_k^{-1}\boldsymbol{b_k} \quad ,\end{aligned} \tag{5.4}$$

with k being the module index. The neglected correlations between the modules then have to be accounted for by iterating the alignment procedure several times.

5.1.3 Tracking and the choice of residuals

The tracks that are currently used for the local χ^2 alignment are reconstructed using the CTB (combined test beam) tracking algorithm, which is optimised to reconstruct tracks not stemming from the nominal interaction points. The CTB tracking uses a χ^2 fit of the full track, together with additional χ^2 terms for scattering angles, to incorporate Coulomb scattering and energy loss of particles [83]. The pattern recognition is tailored to single track events, which makes it also usable for cosmic ray data taking if the orientation of cosmic ray tracks is taken into account. The track fitting part of the CTB algorithm can also be used for collision events as a global χ^2 track fitting algorithm.

There are several definitions of residuals that are used within the ATLAS alignment approaches. *In-plane* residuals are defined as the distances in Cartesian coordinates between the piercing point of the track through the detector element and the readout cluster position on the detector surface ("hit position"). For SCT endcap modules, the readout strips do not run parallel. Therefore the residual definition is either ambiguous or *angular residuals* have to be used. The local χ^2 algorithm uses *distance-of-closest-approach* (DOCA) residuals, where the residual can lie outside of the module plane. The DOCA residuals are defined as the shortest distance in space between the readout strip and the track. For pixel residuals, two virtual readout strips running parallel to the long and short module sides are introduced to be able to calculate DOCA residuals. Pixel hits deliver two residuals (x- and y-residuals from the short and long readout coordinate), while a hit on one SCT side delivers one x-residual. Usually, two sides of an SCT module are hit, and consequently the module acquires two x-residuals per track.

5.1.4 Levels of alignment granularity

The alignment of 5832 individual silicon detector modules is a time-consuming and delicate task. To collect enough hits per module in order to obtain statistically precise residual distributions, a huge number of tracks is needed, which are problematic to acquire during cosmic ray data taking. Detector modules are mounted together on different physical structures, like barrel layers or endcap disks, and thus are expected to move in a correlated way. Therefore, several *levels* of alignment granularity were introduced and encoded in the alignment algorithms. They agglomerate different numbers of individual modules to larger superstructures, in order to initially align only the 6 degrees of freedom (DOF) of this superstructure to determine movements common to all its modules. Therefore, the number of alignment parameters to solve for drops considerably (which is especially useful for the global χ^2 alignment) and the number of hits on each superstructure grows, since it is the sum of all hits in the constituent modules (which is beneficial for every alignment algorithm). The levels that are used within the local χ^2 alignment are:

- Level 1 (L1): This level subsumes modules at the physical subdetector granularity, i. e. it contains the following 4 superstructures: the whole Pixel detector, the SCT barrel and the two SCT endcaps.
- Level 2 (L2): This level contains individual barrel layers and endcap disks as separate structures, i. e. it has 31 alignable structures.
- Level 3 (L3): This level is the module level, where each silicon module is individually aligned as one rigid body. This is the level with the highest granularity and possesses 5832 alignable structures.
- Level 32 (L32): This level is a mixture between L2 and L3 which is particularly useful for the alignment using cosmic ray data. This level uses individual modules in the barrel layers (L3 granularity), while combining modules on each endcap disk (L2 granularity). For a detailed description, c. f. section 5.3.2.

It is foreseen to produce alignment constants for L1 and L2 structures more frequently than for L3.

5.2 Creation of realistic systematic deformations of the ATLAS Inner Detector

A key feature of track based alignment is that, based on the track topology that is available to populate the matrix A, eigenvalues of it can be degenerate. This means that there are detector movements that leave the track residuals and therefore the χ^2 unchanged. The 6 degrees of freedom (3 translational, 3 rotational) of the whole Inner Detector have to be fixed by an external reference, e. g. the centre-of-gravity algorithm that the alignment algorithms use to place the centre of the Inner Detector at the origin of the coordinate system. Additionally, there are detector deformations that leave the helical particle trajectories in a magnetic field helical. For such a deformation, the track fitting routine will find the same residual distributions and the alignment will determine the same alignment correction as for the undeformed case. However, the performance of track reconstruction will be degraded with respect to the optimal alignment, since the track parameters may well change due to the movement. Such systematic detector deformations with only insignificant changes in the residuals are called *weak modes* to emphasize the difficulty to constrain them. Alignment with e. g. tracks from cosmic rays suffers from different weak modes than alignment using tracks from collisions at the interaction point, since the particles traverse the detector differently. Therefore, non-collision tracks are valuable input for determining the detector geometry and it is desirable to have as many different sources of tracks as possible to enhance the eigenvalue spectrum. Consequently, also tracks coming from beam halo and beam-gas interactions will be included into the alignment procedure once the LHC is running.

In this work, a number of systematic deformations are created to study the ability of the alignment algorithms to discover them. They are included as separate database tags in the ATLAS conditions database, and can be used to reconstruct Monte Carlo data samples with a different geometry than the one used for simulation. The deformations are chosen to represent possible simple modes of deformation of the real detector which leave helical tracks from the interaction point unchanged. They are not necessarily expected to appear in real operation (e. g. due to physical effects like gravitational sagging of the detector or temperature changes). The main goal is to study how well the alignment algorithms are able to correct for them, which residual deformation is left after the alignment procedure, and how big an impact these modes will have to the physics performance of the detector.

Technically, the deformations are created at module level via modifying the conditions database alignment constants for all ATLAS Inner Detector modules. For SCT and Pixel, the representation of individual modules in the conditions database allows for basically every parameterised deformation, as long as the module shape stays unaffected. Therefore, almost all deformations can be approximated reasonably well by the silicon modules. For the TRT detector, the coarser granularity of alignment constants somewhat limits the representable movements. The TRT barrel is segmented in two halves for positive and negative Z-coordinate, each possessing 32 phi sectors with three modules. In the database, the corresponding modules for positive and negative Z are grouped and only

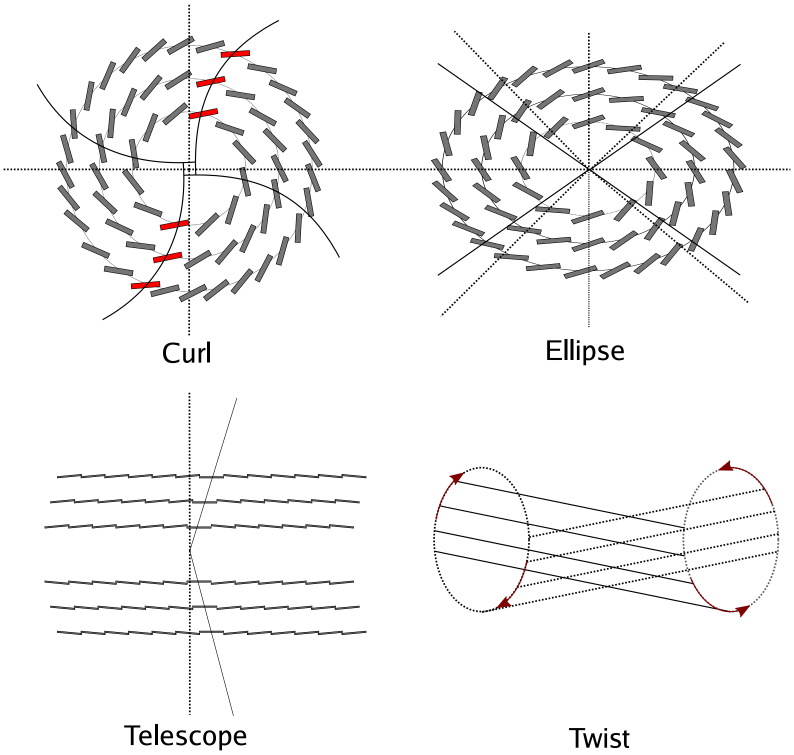


Figure 5.1: *Schematic view of the four chosen deformations.*

one transformation for the two of them is stored. The TRT endcaps are both segmented in 14 wheels with 32 phi sectors each. Every disk has a common alignment transformation in the database without representation of the phi sectors. Therefore, some of the systematic deformations are not fully applicable to the TRT detector description.

A parameterisation of systematic deformations is chosen that varies global R, Φ and Z coordinates as a function of R, Φ and Z, as inspired by studies made by the BaBar experiment [84]. Each of the nine possible combinations is labeled in a manner analogue to R$\Delta\Phi$. This example denotes a systematic shift in the module azimuth angle Φ as a function of the module radial coordinate R. All nine different modes are implemented and four of them are thoroughly tested and used in further studies, since they are thought to be the most difficult ones for the alignment procedures and have the largest impact on physical observables. Specifically, these are R$\Delta\Phi$, $\Phi\Delta$R, RΔZ and Z$\Delta\Phi$. The chosen modes of deformation are explained in detail in the following; a pictorial representation of them is shown in figure 5.1.

- Curl (R$\Delta\Phi$): The curl mode rotates barrel layers proportional to their radial distance from the beamline. This straightens or additionally curls tracks while leaving them helical and thus imbalances the reconstruction of transverse momentum. In addition, a reciprocal term rotates the barrel layers inversely proportional to their distance from the beampipe. This moves the track impact parameters away from the interaction point and spoils the vertex resolution. The parameterisation used is

$$\Delta\Phi = c_1 R + c_2/R \quad , \tag{5.5}$$

with the parameters $c_1 \approx 7.6 \cdot 10^{-4}\,\frac{\mathrm{mrad}}{\mathrm{mm}}$ and $c_2 \approx 50\,\mathrm{mrad\,mm}$ chosen such that the outer SCT barrel layer moves by 200 µm and the inner Pixel layer by 50 µm. Since the TRT barrel modules have a sizable extension in the radial direction, they are additionally rotated by the same angle $\Delta\Phi$ around their centrepoints to better approximate a curl mode. The TRT endcap disks are all rotated by a fixed angle, determined by the radial coordinate of their strawcentres.

- Ellipse (ΦΔR): The elliptical mode moves modules away from or towards the beamline depending on their azimuthal position, which produces a Φ-dependent effect on the distance scales. This should affect reconstructed vertex masses, since the angles between pairs of tracks are affected by this deformation. The deformation is parameterised as

  $$\Delta R = c\cos(2\Phi)R \quad , \tag{5.6}$$

  such that the upper and lower parts of barrel layers move inwards, while the horizontal parts move outwards. The constant $c \approx 4.9 \cdot 10^{-4}$ is chosen such that the SCT modules in the outer barrel layer move by at most 250 µm. The TRT endcap is not moved at all, since only complete discs can be aligned. For them, creating an elliptical deformation is not possible.

- Telescope (RΔZ): The telescope mode shifts modules along the beamline proportional to their distance from the beamline. This creates an asymmetry in the rapidity distribution and therefore mimics a boost along Z. This mode is parameterised as

  $$\Delta Z = cR \quad , \tag{5.7}$$

  with $c \approx 9.7 \cdot 10^{-4}$ chosen such that the SCT modules on the outer barrel layer move by 500 µm. The TRT endcaps cannot bend into a true telescope mode, they are moved as a whole along Z by an amount determined by the radial coordinate of their strawcentres.

- Twist (ZΔΦ): The twist mode rotates rings in the barrel layers proportional to their Z position. This curls the tracks in the positive direction for $Z > 0$ and negatively for $Z < 0$. The result is a transverse momentum bias as a function of pseudorapidity or polar angle. The used parameterisation is

  $$\Delta\Phi = cZ \quad , \tag{5.8}$$

  with $c \approx 2.5 \cdot 10^{-4}\,\frac{\mathrm{mrad}}{\mathrm{mm}}$ chosen such that the SCT modules on the outer barrel rings move by at most 100 µm. Since the TRT barrel modules span the whole length of the barrel without segmentation, they cannot follow the twist by a $\Delta\Phi$ rotation. Instead, each module is rotated around an axis through its centrepoint which points radially outwards from the beamline. This follows a twist movement well enough for small angles. The necessary rotation angle δ is calculated as follows:

  $$\delta \approx -\frac{\Delta\Phi \times R}{Z} = -cR \quad . \tag{5.9}$$

  The minus sign corrects for the direction of rotation between Pixel/SCT and TRT.

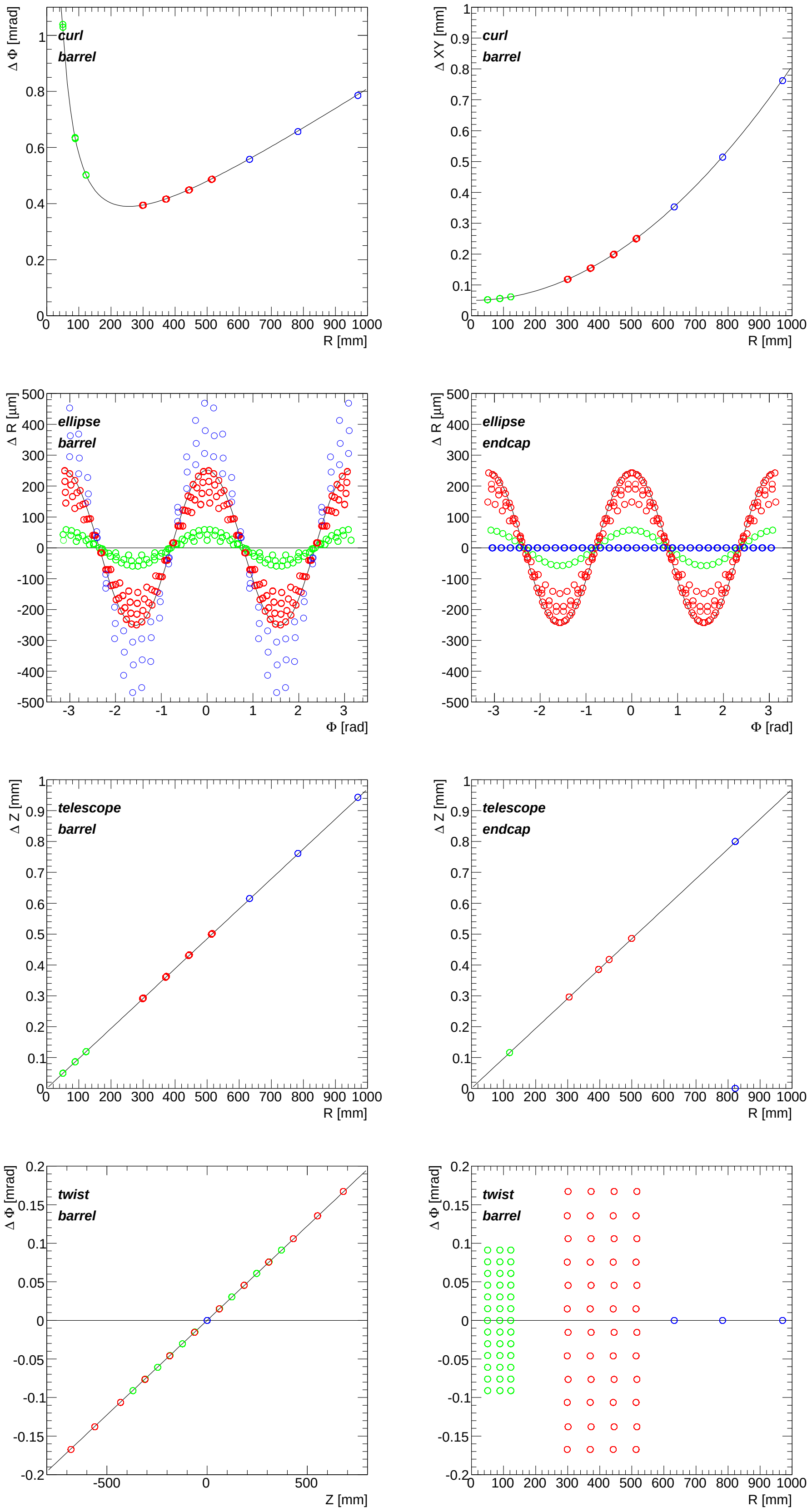


Figure 5.2: *Visualisation of the created systematic deformations. The figures show the movements of Inner Detector modules as function of their original positions. Pixel modules are depicted in green, SCT modules in red and TRT modules in blue. The individual figures are described in detail in the text.*

To test whether the four deformations have been created correctly, a set of visualisation tools are applied and control distributions are produced with physics monitoring tools.

In figure 5.2, some examples of control graphs of the module movements are shown. Each circle in the graph represents a module: Pixel modules are coloured in green, SCT modules in red and TRT modules in blue. The black curves represent the parameterised movements that the modules should follow. The top row of figures shows the $\Delta\Phi$ values (left) and $\Delta XY = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$ values (right) as functions of the module radius for the Curl misalignment in the barrel. Each barrel layer therefore produces several entries of modules sitting at the same radial coordinate. One can see the reciprocal decline plus the linear ascent for $\Delta\Phi$. This achieves that the parabola in the ΔXY graph does not start at zero displacement, but at the desired 50 μm for the innermost Pixel modules. Each of the three Pixel layers and four SCT layers consists indeed of two points, which is a result of the stave structure in the barrels.

The second row shows the radial displacements of barrel (left) and endcap (right) modules versus their Φ-coordinate in the detector for the Ellipse deformation. The parameterisation as a cosine function is visible. The black curve denotes the desired shifts for the SCT modules in the outer barrel layer and endcap rings. In the barrel, the three Pixel, four SCT and three TRT layers can be distinguished. In the endcap, the TRT modules are not shifted at all, since their database representation cannot radially shrink or expand. The Pixel endcaps only have one ring of modules per disk, whereas the SCT disks have up to three rings of modules at different radial locations. The middle ring for some disks consists of short middle modules instead of long middle modules, therefore this ring is split again into two different radii.

The third row of figure 5.2 shows the displacements in global Z versus the radial coordinate of barrel (left) and endcap modules (right) in the Telescope mode. Once again, the layer structure is visible and the displacements behave as expected.

Finally, the fourth row depicts the Φ-displacements as functions of the module Z (left) and R-coordinates (right) in the barrel for the Twist geometry. This shows how the different rings of the barrel extend in Z: The Pixel detector has 13 rings per layer (partially hidden by SCT and TRT points), the SCT has 12 rings per layer. The TRT modules span the whole length of the barrel and therefore all of their centres lie at $Z = 0$. They are not rotated in $\Delta\Phi$, but the compensation rotation described above is applied, which is invisible in this figure.

These figures, together with figures of other coordinate combinations, show that the module movements follow the desired misalignments and that no accidental movements have been introduced into the geometry of the Inner Detector. For this work, the four deformations are then created on top of the misaligned geometry that was implemented for the Computing Systems Commissioning Challenge (CSC), to be analysed and used by the alignment and physics groups of ATLAS.

The resulting geometry sets are then used to infer the effects on reconstructed quantities when the Inner Detector is systematically deformed [85]. For this purpose, the Curl and Twist deformations are compared to the perfect knowledge of the CSC misaligned detector, both before and after an alignment procedure. This procedure uses 100k simulated multimuon events for a global χ^2 alignment of the silicon part of the Inner Detector and an alignment of the TRT barrel using the TRT alignment algorithm. The impact on physics observables is additionally tested with a sample of 20k simulated $Z \rightarrow \mu\mu$ decays. It is found that the Curl deformation indeed forms a weak mode, since it leaves the residual and track χ^2 distributions unchanged. However, as expected, it biases curvature and therefore momentum measurement. After alignment, this bias is greatly reduced, but not fully removed. The deterioration of the d_0 measurement, which is introduced by the $1/R$ term, is removed by imposing a beamline constraint. The mass resolution of the reconstructed dimuon invariant mass in the $Z \rightarrow \mu\mu$ sample is improved. Figure 5.3 exemplarily shows the event-by-event difference between the reconstructed dimuon mass and the invariant mass of the two corresponding simulated muons from [86]. As can be seen from the red distribution, the Curl deformation before any alignment procedure broadens this resolution and additionally shifts it to higher masses. This is from the fact that the momentum increase of positively charged muons exceeds the momentum decrease of negatively charged muons in the Curl deformation. Performing the alignment procedure results in the blue mass distribution, which is centred and strongly improves the mass resolution for muons reconstructed in the barrel part of the Inner Detector. When using muons from the entire acceptance region of the Inner Detector, the improvement is not as good, since the TRT endcaps are not aligned.

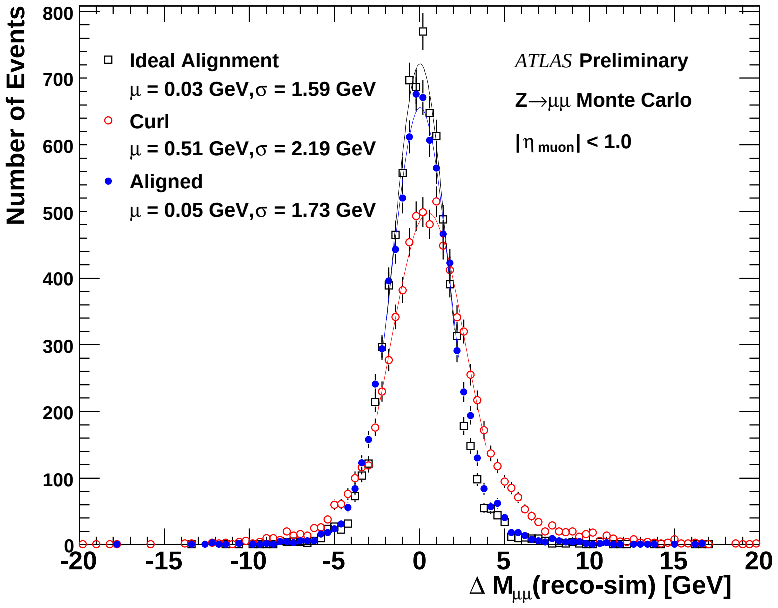


Figure 5.3: *Event-by-event difference of the reconstructed and simulated dimuon invariant mass in the* $Z \rightarrow \mu\mu$ *sample from [86, 87]. Black points denote the full knowledge of the simulated geometry, red points the curl deformation. The blue points emerge after the alignment procedure. Only events where both muon candidates are reconstructed within the barrel region (*$|\eta| < 1.0$*) are shown. The mean and width of a Gaussian fit in the range [*$\mu-RMS,\mu+RMS$*] are listed. This figure is an updated version of figure 14e) in [85].*

For the Twist geometry, it is seen that a weak mode is created, since the residuals and track χ^2 distributions stay unaffected, but the dimuon invariant mass resolution is degraded by the η-dependent curvature bias. Omitting endcap tracks improves the resolution, since the Twist affects tracks stronger if they are more forward-oriented. Performing the alignment procedure improves the resolution for silicon-only tracks. Since no TRT end-cap alignment is performed, the dimuon mass resolution stays degraded for TRT-extended tracks. For the Telescope deformation with the default value of c, no impact on the physical observables is seen. Therefore an exaggerated version with the movements increased by a factor of six is created. Even for this larger deformation, the impact on the track pseudorapidities is small. It is concluded that the Telescope deformation represents a true weak mode, but which is unlikely to be dangerous, since it needs an unphysically large magnitude to affect the detector resolution. For the Ellipse geometry, another version with misalignments increased by a factor of four is produced. This larger version exhibits a degradation of track χ^2 distribution and can therefore not be termed a true weak mode. This should enable alignment algorithms to correct for this kind of deformation, which is found to predominantly worsen the η and z_0 resolution.

Further on, studies from the physics groups showed that the impact of the Curl misalignment on the reconstruction of $B_d^0 \rightarrow J/\Psi K^{0*}$-decays is minor compared to the well aligned case, where material interactions dominate the mass resolution [86]. Other studies have been performed on the impact of systematic misalignments on the b quark identification and found that the Curl deformation severely lowers the b-tagging efficiency due to the misalignment of the inner pixel layers coming from the $1/R$ term [88].

5.3 Local χ^2 alignment using cosmic ray data

The local χ^2 alignment algorithm has been developed within the ATLAS software framework Athena since 2004. It is designed to align the silicon part of the Inner Detector by using track residuals from different sources - collision tracks, cosmic ray tracks and beam-gas interaction or beam halo tracks. It is a part of the Athena releases and can be accessed from the ATLAS code repository [89].

The local χ^2 algorithm has been applied to simulated as well as to real data from ATLAS. Most notably, it was used in the analyses of the Combined Testbeam from 2004, the CSC effort from 2007, and the cosmic runs in the ATLAS pit before the initial LHC startup in 2008. This chapter deals with the improvement of results obtained using cosmic ray data from autumn 2008. A description of the alignment method and the input to it is given, followed by an assessment of the improvements that are made to the alignment method.

5.3.1 Cosmic ray reconstruction and datasets

The muon tracks originating from cosmic rays that are used were taken during 64 ATLAS runs in September and October 2008. These runs are selected according to the data quality flag in the ATLAS run database. Only runs with the SCT detector marked as green and—if the Pixel detector is included in the run—also Pixel detector marked as green are used. The green data quality flag denotes runs that possess good data quality based on specific criteria defined per subdetector. These runs are meant to be directly usable for analysis without any further data quality scrutiny [90]. There are 13 runs taken without magnetic field, and 51 runs taken with the solenoidal and toroidal magnets operating at their nominal field strength that fulfill this requirement.

The version of the Athena framework used for reconstruction of the hits was Athena 14.5.2.4. The reconstruction produces Event Summary Data (ESD) files, taken as input for the local χ^2 alignment efforts. The tracks that are used for the local χ^2 algorithm are re-reconstructed within each alignment iteration using the Tier0 version of Athena 15.0.0.5 to incorporate the change of detector geometry from the previous iteration. This is an Athena version that is also used for prompt reconstruction of ATLAS data at the CERN Tier0 computer farm. As mentioned before, the tracks are reconstructed using the CTB tracking algorithm.

The following quality criteria are applied to the data: only tracks with at least 8 hits and—if the solenoidal field was on during data taking—a momentum of at least 1 GeV are used as input to the local χ^2 alignment procedure. These requirements suppress ill-reconstructed tracks and large residuals e. g. coming from material scattering. In total about 683k tracks fulfill these requirements, 144k from runs taken without magnetic field and 539k from runs with magnetic field. Additionally, the `HitQualSelTool` [91] is used to discard hits which potentially have deteriorating effects to the alignment. The standard settings are used, which only accept hits with a transverse track incidence angle of less than 0.8 rad ($\approx 45°$) and less than 6 strips or pixels contributing to a cluster. These requirements were found to sufficiently suppress pathological hit configurations, e. g. large clusters from bremsstrahlung or detector noise and tracks with shallow incidence that are more prone to non-optimal cluster reconstruction. In addition, hits on edge strips and pixels as well as outliers are rejected (outliers are hits that contribute a too high χ^2 value to the track fit and therefore do not participate in the fit, while still being associated to the track). In this track sample, nearly 80 % of the hits on the selected tracks are accepted.

The distributions of track parameters for the tracks used in the alignment are shown in figure 5.4. These are the perigee parameters [92, 93], i. e. the track parameters at the point of closest approach towards the origin of the ATLAS global coordinate system. The blue histograms denote track parameters of tracks measured with magnetic field, the black dashed histograms of tracks without magnetic field. The parameters are:

- z_0, the longitudinal impact parameter, and d_0, the transverse impact parameter. These measure the distance of the track from the origin of the coordinate system. Since cosmic ray tracks do not point to the interaction point, these distributions are very broad and flat instead of Gaussian and are determined by the size of the Inner Detector and the trigger acceptances.
- the azimuthal angle of the track ϕ and the polar angle θ. Since most of the tracks come from above and are directed downwards, the most prominent peak is situated at $\phi = -1.5$. The different sizes of the two access shafts results in two peaks with different heights in the θ distribution.
- the inverse charged momentum q/p. The q/p distributions show only tracks measured with the magnetic field switched on; tracks measured without magnetic field have no momentum measurement attributed to them. The $(q/p)^{-1}$ distribution shows momenta up to 10 GeV, while the q/p distribution shows momenta from 10 GeV upwards, so that the whole momentum range is covered. The 1 GeV requirement on the track momentum is visible as a step in the lower right histogram. The momentum distributions indicate an excess of positively charged muons from

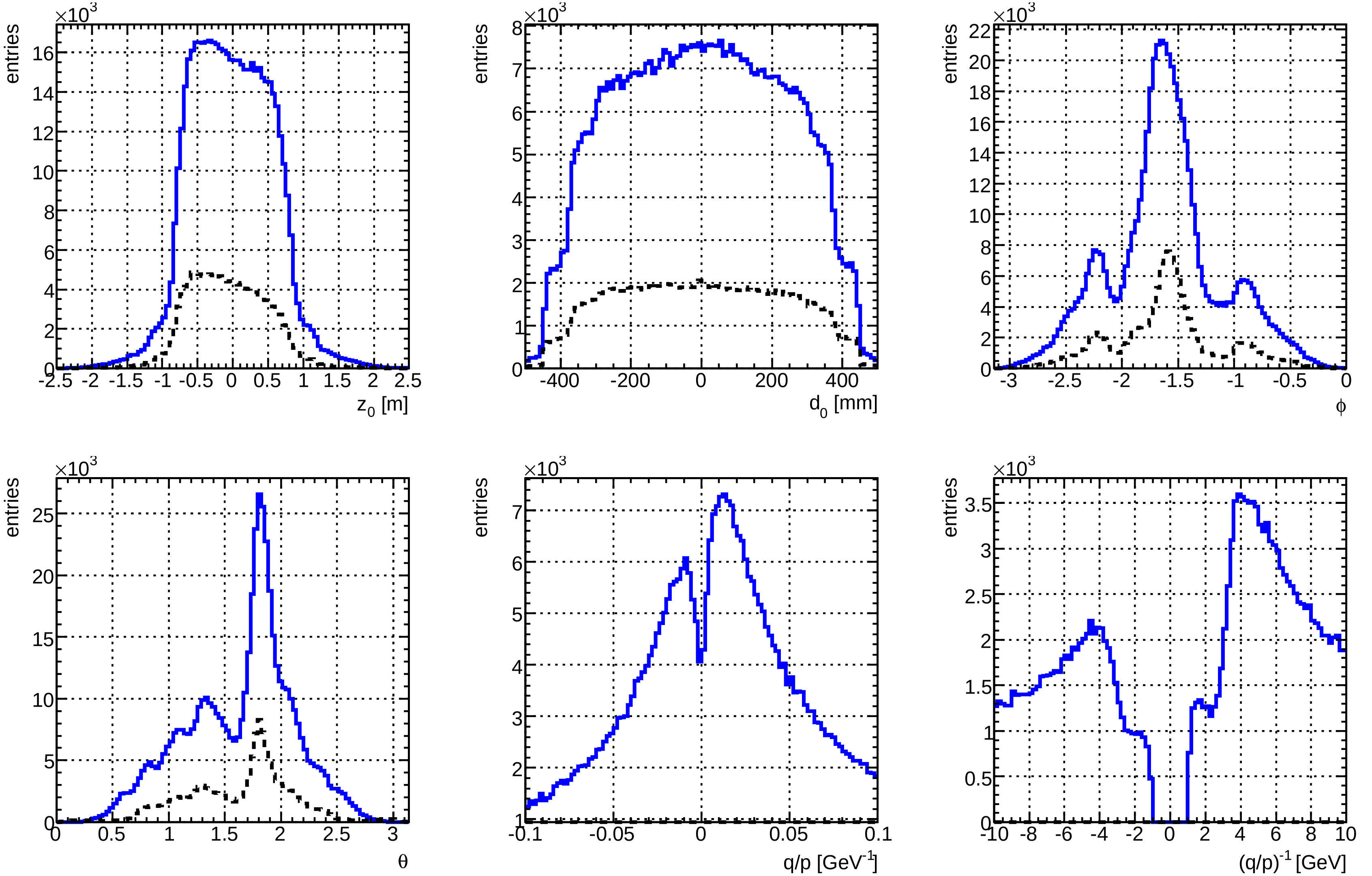


Figure 5.4: *The five track parameters of the cosmic ray tracks used for alignment. Tracks from runs without (with) magnetic field are displayed in black dashed (blue). For the momentum distribution, both q/p and its inverse are shown to cover all momenta.*

cosmic rays. This is expected from the known charge distribution of cosmic ray particles hitting the atmosphere [18].

5.3.2 Local χ^2 alignment procedure

The local χ^2 alignment is performed on this data with the following strategy: First, 4 iterations at L1 are performed to determine large movements of the silicon subdetectors with respect to each other. This aligns the Pixel detector and the three SCT parts with respect to each other, taking out movements that would need a higher number of iterations to converge on more granular levels of alignment. Afterwards, 6 iterations at L2 will better determine layer and disk positions. The low number of hits on the endcap modules would result in large uncertainties on a per-module alignment and so the endcaps are only aligned per disk. Additionally, a soft constraint is added to smoothen the convergence behaviour of the endcaps. Then 15 iterations on L32 align the barrel modules, which have enough hits for a per-module alignment, to a high precision. A pseudo survey constraint is added to enforce consistent alignment corrections also for modules without any hits, as well as to constrain movements with large statistical uncertainties. The endcap disks are included in the alignment, which makes them converge further if necessary, as well as adapt to changes in the barrel geometry. This procedures ensures a consistent alignment without the need for additional iterations to reconnect the endcap and barrel geometry, as previously used in [53]. Finally, 10 more iterations at L2 are performed to reach the final alignment constants and to judge whether the solution has converged.

The movements of the L2 structures for the first 10 iterations without a soft constraint are shown in figures 5.5–5.8. In each of these figures, the top rows depict the alignment parameters as a function of iteration for the three translational degrees of freedom, while the bottom rows show the corresponding rotational alignment parameters. The coordinate system that is used for alignment parameters at L1 and L2 is the global coordinate system. During the first four L1 iterations the endcap disks and barrel layers of the same L1 structure share common alignment parameters.

Figure 5.5 shows that within the ten iterations the Pixel barrel layers show good convergence, while the endcap disks show ongoing movements in some DOFs (e. g. β). The SCT barrel layers (figure 5.6) show a very fast convergence. Apparently the high number of hits and the larger spread in incident angles compared to endcap disks lead to a stable alignment solution very fast. The behaviour in Z, where the L1 shift of the whole barrel is followed by a shift of all layers in the opposite direction is a known effect, which at present is not fully understood. The SCT disks of endcap A (figure 5.7) show a converging behaviour for the more inner disks, i. e. the ones at smaller $|Z|$, whereas it is harder to align more outer disks at larger $|Z|$. This can e. g. be seen by the increasing divergence of disks 6, 7, and 8 for the β-rotations. The disk 8 seems to be diverging in the α-rotation. The translation in Y acquires large alignment parameters for this disk as well, since it is connected with the α-rotation via the large distance between the centre of the coordinate system and the centre of the disk. Thus the Y-shift is an artifact that merely compensates the large shift caused by the rotation about the origin of the coordinate system, to yield only a net rotation of the disk, while the actual shift in space of the disk is much smaller. This is shown in figure 5.9, which depicts the X- and Y-positions in global coordinates for the disk centres during the iterations, so that one can see the actual translations that

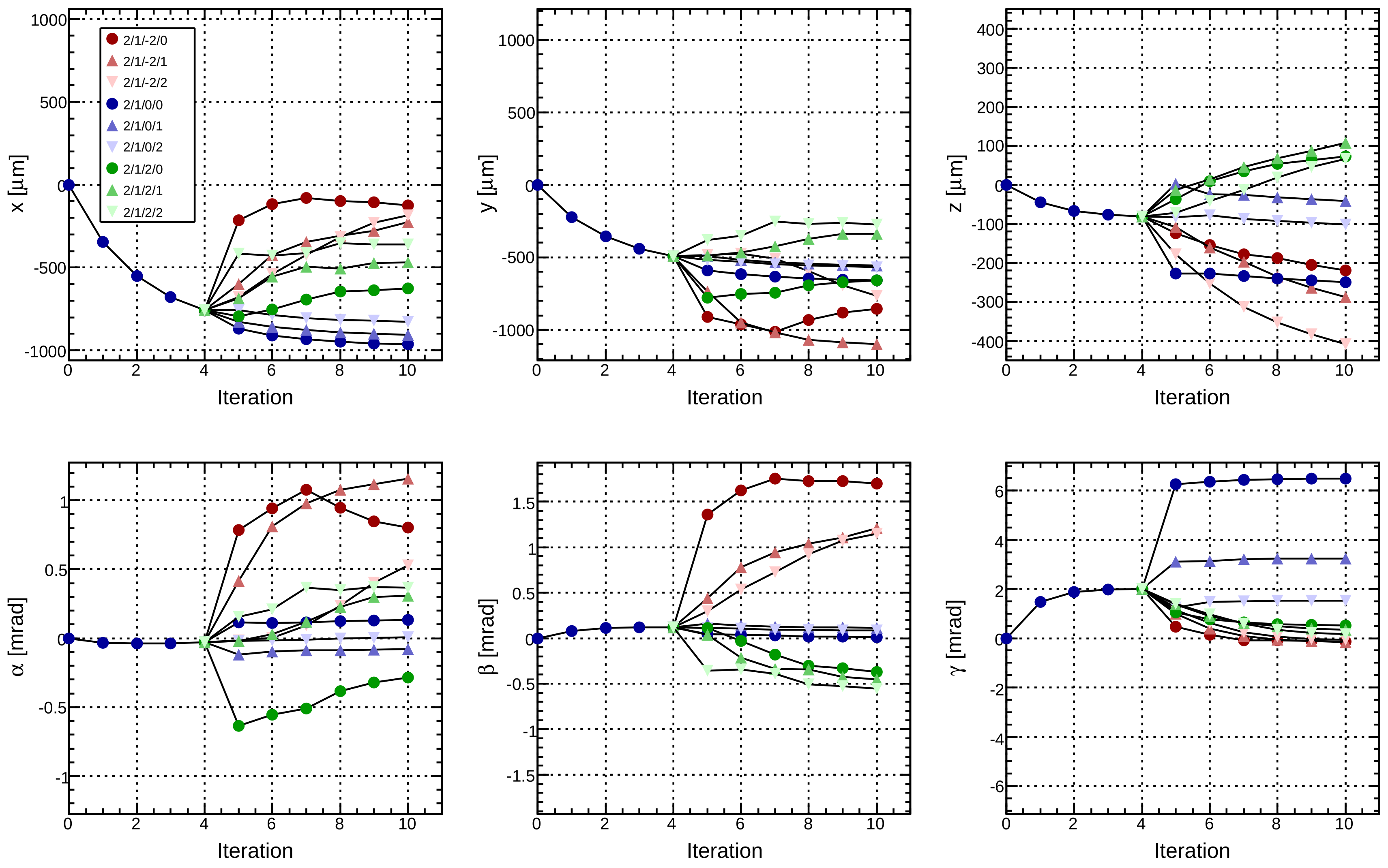


Figure 5.5: *Evolution of the six alignment parameters of the Pixel layers and disks during the first 10 iterations. No soft mode constraint is applied. The Inner Detector readout identifiers, as described in section 3.2.2, are used for numbering individual L2 structures.*

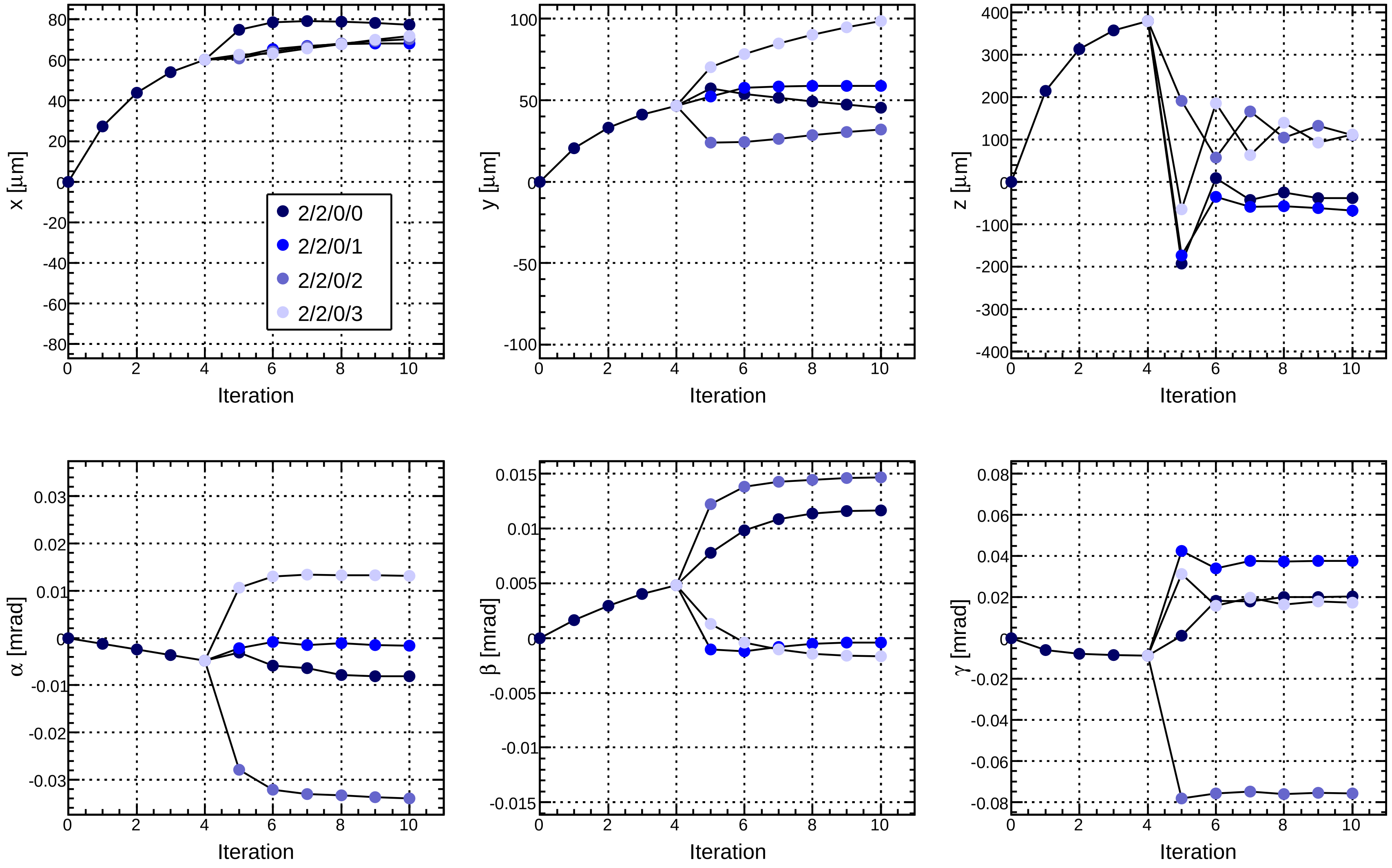


Figure 5.6: *Same as figure 5.5, but for the SCT barrel layers.*

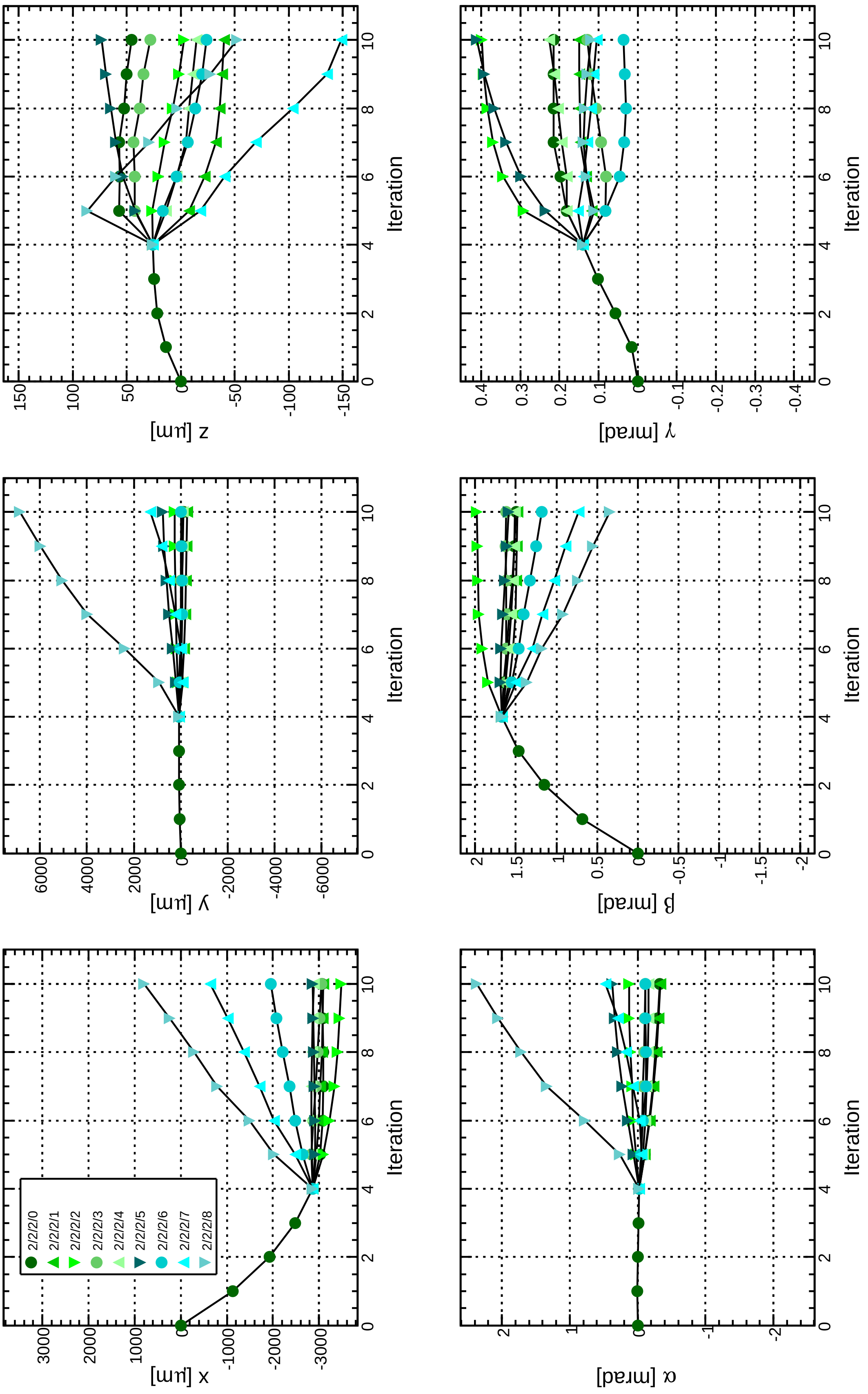


Figure 5.7: *Same as figure 5.5, but for the SCT endcap A disks.*

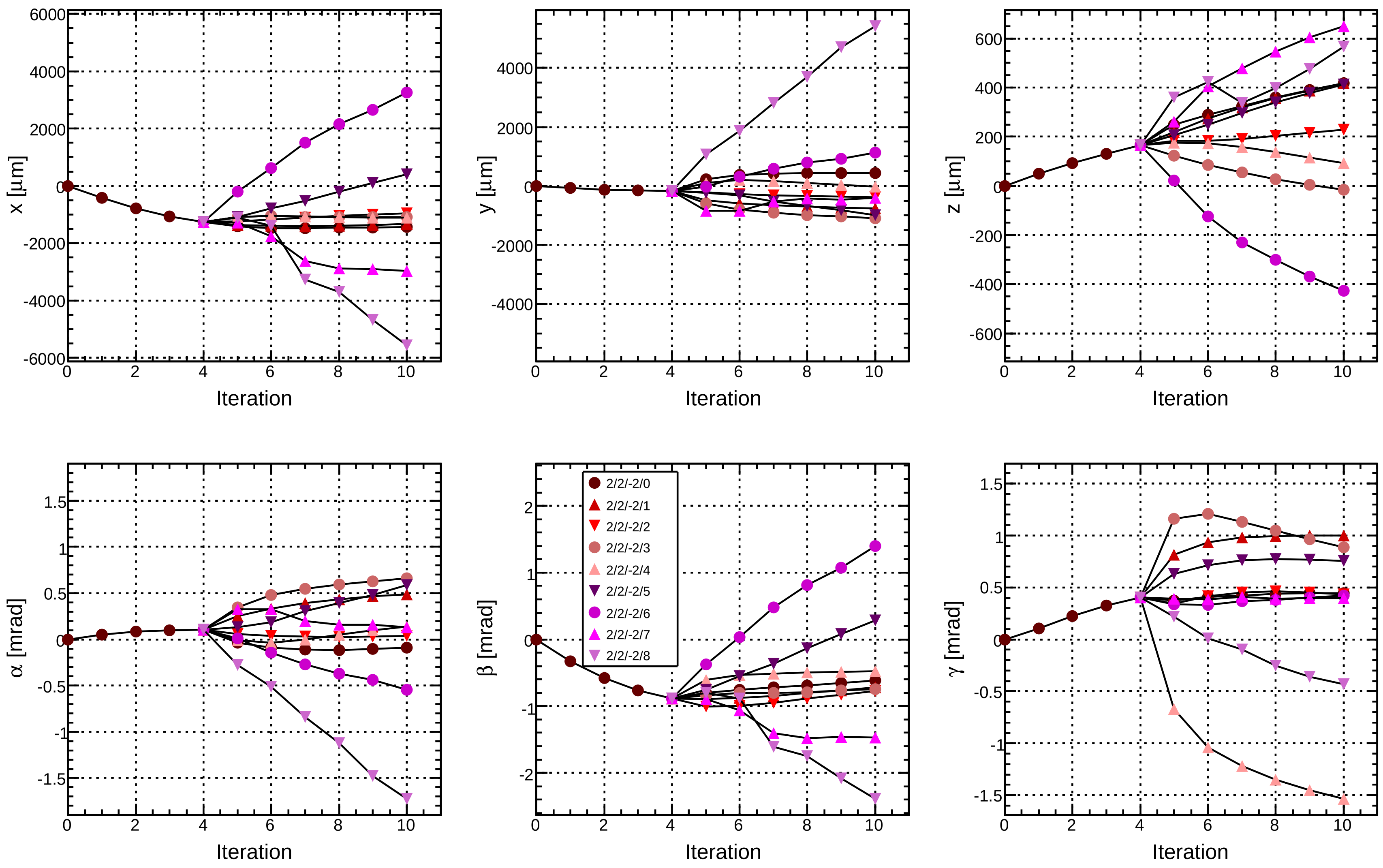


Figure 5.8: *Same as figure 5.5, but for the SCT endcap C disks.*

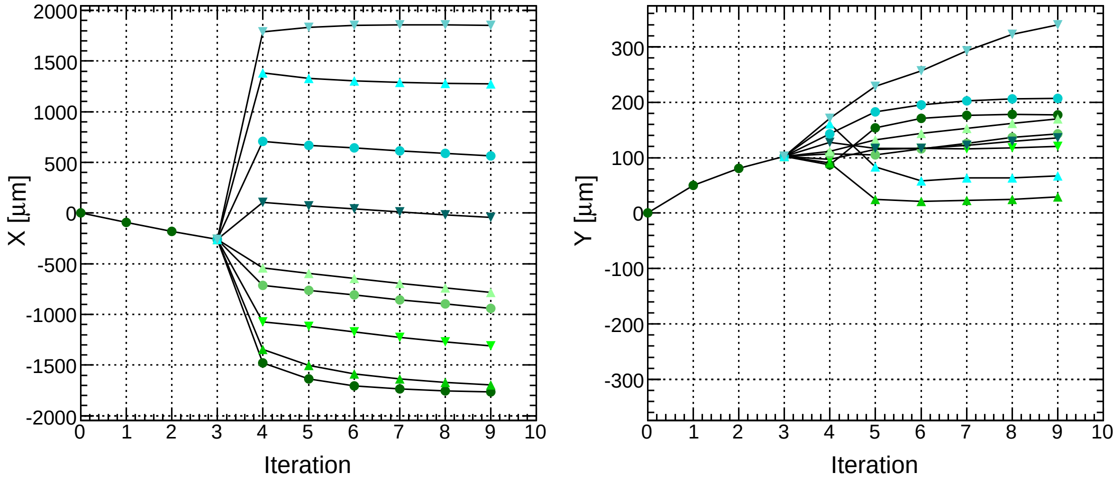


Figure 5.9: *Global X- and Y-positions of the SCT endcap A disks during the first 9 iterations. No soft mode constraint is applied.*

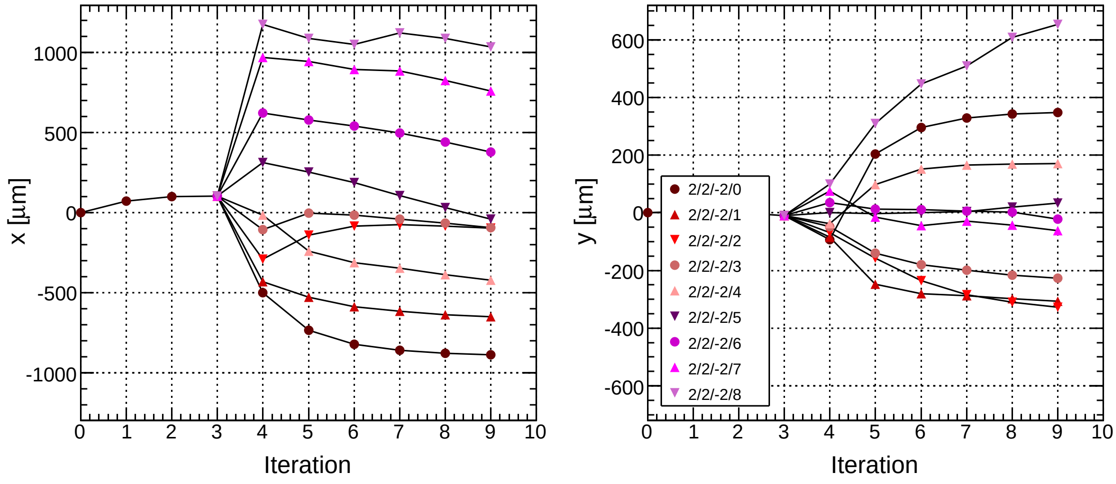


Figure 5.10: *Same as figure 5.9, but for the SCT endcap C disks.*

are calculated during the alignment. Here, disk 8 shows a signs of convergence for the Y-translation and the scale on which X and Y alignment corrections occur is much reduced. In this figure, the disks split up in their X-position in iteration 4 because the β-rotation of the entire endcap of about 1.5 mrad tilts them away from the beamline. After that, the X-movements are minor. The ongoing alignment change in the X alignment parameter in figure 5.7 merely compensates the β-rotation so that only the net rotations without further translations are applied. For the SCT endcap C disks (figure 5.8), it is seen that also preferentially the more outer disks 5–8 acquire notable alignment corrections, e. g. rotation angles for β of more than 1 mrad for the disks 6, 7 and 8. Disk 6 of SCT endcap C is a problematic disk that reads out many noise hits, leading to a bad distribution of residuals. This also affects the alignment of the neighbouring disks 5 and 7, which also show no convergence in some DOFs. From figure 5.10, which depicts the X- and Y-positions of the endcap C disk centres, it is seen that the Y-translations converge, except for disk 8, where the large Y-translations in figure 5.8 compensate the α-rotations not completely. These compensating effects do not appear for the barrel layers, since their centre of gravity (initially) coincides with the global coordinate origin.

It is more difficult to align the more outer disks, since they acquire a lot fewer hits than the more inner disks and the distribution of track directions is more narrow. Since every track needs to pass at least 4 modules to acquire the required 8 hits, the fraction of cosmic ray tracks that reach the outermost disks is very small. Overall, it is seen, as expected, that the X- and Y-translations, as well as the γ-rotations are the best-constrained DOFs for alignment, since these are directly measured by the detectors. For the endcap disks, the Z-translations and especially α- and β-rotations show more non-converging behaviour, the more limited the incident angle distribution and hit statistics becomes.

Addition of a soft constraint

To still be able to align the endcap disks in all six degrees of freedom, the local χ^2 algorithm is augmented by a soft constraint for all endcap disks. This constraint is aimed to reduce the movements of badly determined degrees of freedom and thus stabilises the alignment of the endcaps. It is implemented by adding one pseudo-hit to each endcap disk and barrel layer at the end of each iteration. This pseudo-hit mimics a hit with zero residual and a residual uncertainty that can be adjusted to determine the tightness of the constraint. Therefore the vector $\boldsymbol{b}$ stays unchanged in the formalism, and the matrix A acquires a summand of $\mathrm{diag}(\sigma_\mathrm{x}^{-2}, \sigma_\mathrm{y}^{-2}, \sigma_\mathrm{z}^{-2}, \sigma_\alpha^{-2}, \sigma_\beta^{-2}, \sigma_\gamma^{-2})$. Unlike for the global χ^2 alignment, where the correlations between the alignable structures are not neglected, for the local χ^2 a constraint on a full alignment mode of all structures cannot be imposed, but only each disk can be constrained individually.

To determine adequate values for the uncertainties on the pseudo-hits, the alignment parameter uncertainties without the soft constraint are analysed. Table 5.1 shows the number of hits and the calculated uncertainties on the six alignment corrections of each L2 alignable structure after iteration 5. These uncertainties are of statistical nature and stem from the accumulated error matrices of the residuals. A clear dependence of their magnitudes on the number of accumulated hits can be seen. Since 50–100 hits lead to satisfying statistical alignment correction uncertainties (c. f. section 5.3.2), it is decided to adjust the pseudo-hit uncertainties to the values one can expect for 50 hits times the number of modules on a disk. This is justified by the wish to constrain correlated movements of all modules on a disk more strongly than individual movements of a single module. By this constraint, the movements with larger statistical uncertainties than this value are dampened, while alignment corrections with a higher precision are affected only very little by the constraint. The values of the soft mode uncertainties are determined from table 5.1 by taking about 2500 hits as reference for a pixel disk and about 5000 hits for SCT disks (pixel disks have 48 modules, SCT disks have 52, 96 or 132). The values chosen are listed in table 5.2. The L2 iterations 5–10 are performed with these constraints being applied to all L2 alignable structures.

By applying this constraint, the alignment parameter evolution for the Pixel disks is smoothened out, as can be seen by comparing figure 5.11 (with soft constraint) with figure 5.5 (without constraint). In the end all structures seem to converge to the same values as without the constraint. For the barrel layers the soft constraints have practically no effect, since there the alignment solution was a lot more accurate than the constraint in the beginning. This is also visible for the SCT barrel layers from figure 5.12. For the SCT endcaps, the addition of the constraint has a beneficial effect on the alignment parameter evolution. For SCT endcap A (figure 5.13), the ongoing rotation of disk 8 in α is greatly reduced, but not brought to convergence. The endcap C disks show the biggest effect, as seen by comparing figure 5.14 with 5.8. Here the alignment corrections are smaller and smoother, especially for the non-converging rotations in disks 5–8. Now disk 8 even seems to achieve convergence in its α-rotation. However, the soft constraint cannot force convergence on each alignment parameter of a disk, but should give a more stable alignment parameter evolution towards the χ^2 minimum. Overall, the α-rotations of the more outer SCT endcap disks benefit the most, with some beneficial effect also on the β-rotations and Z-translations.

Residuals after L2 alignment

Figure 5.15 shows the DOCA residuals that result after 0, 4, and 10 iterations of the local χ^2 algorithm, split by subdetectors. The black histogram shows the initial residuals before the first iteration. The SCT detector is already prealigned in the reconstruction geometry, as can be seen by the centred and Gaussian-shaped residual distributions. However, their width is larger than that of the residual distributions of a perfectly aligned detector. The green histogram denotes the residuals after the 4 L1 iterations. Finally, the blue histogram shows the residuals after 10 iterations. The distributions become centred and Gaussian during the iterations, with their widths shrinking. For the Pixel residuals, the lower amount of statistics, especially for the endcaps, is visible in the larger bin-by-bin fluctuations. Figure 5.16 depicts the y-residuals for the Pixel modules. These are wider than x-residuals, but the alignment procedure also minimises them adequately.

The perfect alignment knowledge would result in perfectly Gaussian residual distributions with their widths determined by the sensor resolution plus the track fit uncertainty. Assuming that the residual uncertainties are the quadratic sum of the sensor resolution and the misalignment of sensors, an estimate of the remaining random L3 misalignments can be deduced for those residual distributions that are centred after the L2 alignment. For this purpose, the distributions are fitted with the sum of two Gaussian distributions

Table 5.1: *Statistical uncertainties on the alignment corrections in iteration 5 for barrel layers and endcap disks and the corresponding number of hits.*

Pixel structures

		$\sigma(a_x)$	$\sigma(a_y)$	$\sigma(a_z)$	$\sigma(a_\alpha)$	$\sigma(a_\beta)$	$\sigma(a_\gamma)$
structure	# hits	[µm]	[µm]	[µm]	[µrad]	[µrad]	[µrad]
endcap C disk 0	2412	1.45	2.41	3.52	34.15	48.12	11.37
endcap C disk 1	2524	1.45	2.36	3.27	33.67	45.44	11.14
endcap C disk 2	2853	1.85	2.48	3.55	42.66	43.35	13.34
barrel layer 0	49369	0.15	0.28	0.67	1.23	0.63	2.81
barrel layer 1	86665	0.11	0.22	0.53	0.94	0.47	1.21
barrel layer 2	114221	0.10	0.20	0.50	0.84	0.44	0.79
endcap A disk 0	3008	1.31	1.91	2.74	30.70	35.54	9.50
endcap A disk 1	2827	1.54	1.92	2.90	35.28	33.72	10.67
endcap A disk 2	1917	2.51	2.45	4.00	57.95	40.59	16.52

SCT structures

		$\sigma(a_x)$	$\sigma(a_y)$	$\sigma(a_z)$	$\sigma(a_\alpha)$	$\sigma(a_\beta)$	$\sigma(a_\gamma)$
structure	# hits	[µm]	[µm]	[µm]	[µrad]	[µrad]	[µrad]
endcap C disk 0	59531	0.77	1.10	1.50	4.72	4.90	1.41
endcap C disk 1	82784	0.64	0.85	1.07	3.61	3.64	1.20
endcap C disk 2	65839	0.84	1.03	1.33	4.73	4.46	1.52
endcap C disk 3	46125	1.03	1.21	1.56	5.75	5.15	1.80
endcap C disk 4	40427	1.25	1.43	1.93	6.90	6.25	2.11
endcap C disk 5	15239	2.70	2.91	4.46	16.92	14.48	4.36
endcap C disk 6	5271	5.78	5.37	12.26	42.42	38.80	8.12
endcap C disk 7	2563	6.63	8.23	18.39	58.37	60.03	10.00
endcap C disk 8	1093	11.85	15.11	47.21	114.90	131.42	20.32
barrel layer 0	1291792	0.04	0.07	1.73	0.19	0.09	0.16
barrel layer 1	1657170	0.03	0.06	1.49	0.16	0.08	0.11
barrel layer 2	1876009	0.03	0.06	1.48	0.16	0.07	0.08
barrel layer 3	2062932	0.03	0.07	1.53	0.17	0.08	0.08
endcap A disk 0	63233	0.67	0.95	1.33	4.21	4.39	1.21
endcap A disk 1	83158	0.64	0.82	1.05	3.55	3.52	1.16
endcap A disk 2	63905	0.86	1.02	1.36	4.90	4.40	1.53
endcap A disk 3	47854	1.02	1.14	1.54	5.81	4.91	1.76
endcap A disk 4	40732	1.24	1.36	1.89	7.00	6.00	2.07
endcap A disk 5	18150	2.23	2.46	3.87	15.22	12.64	3.75
endcap A disk 6	8587	3.44	3.90	7.19	25.96	22.22	5.33
endcap A disk 7	3634	4.72	5.76	12.54	42.81	38.17	7.23
endcap A disk 8	1963	6.95	9.86	23.40	71.17	70.33	10.35

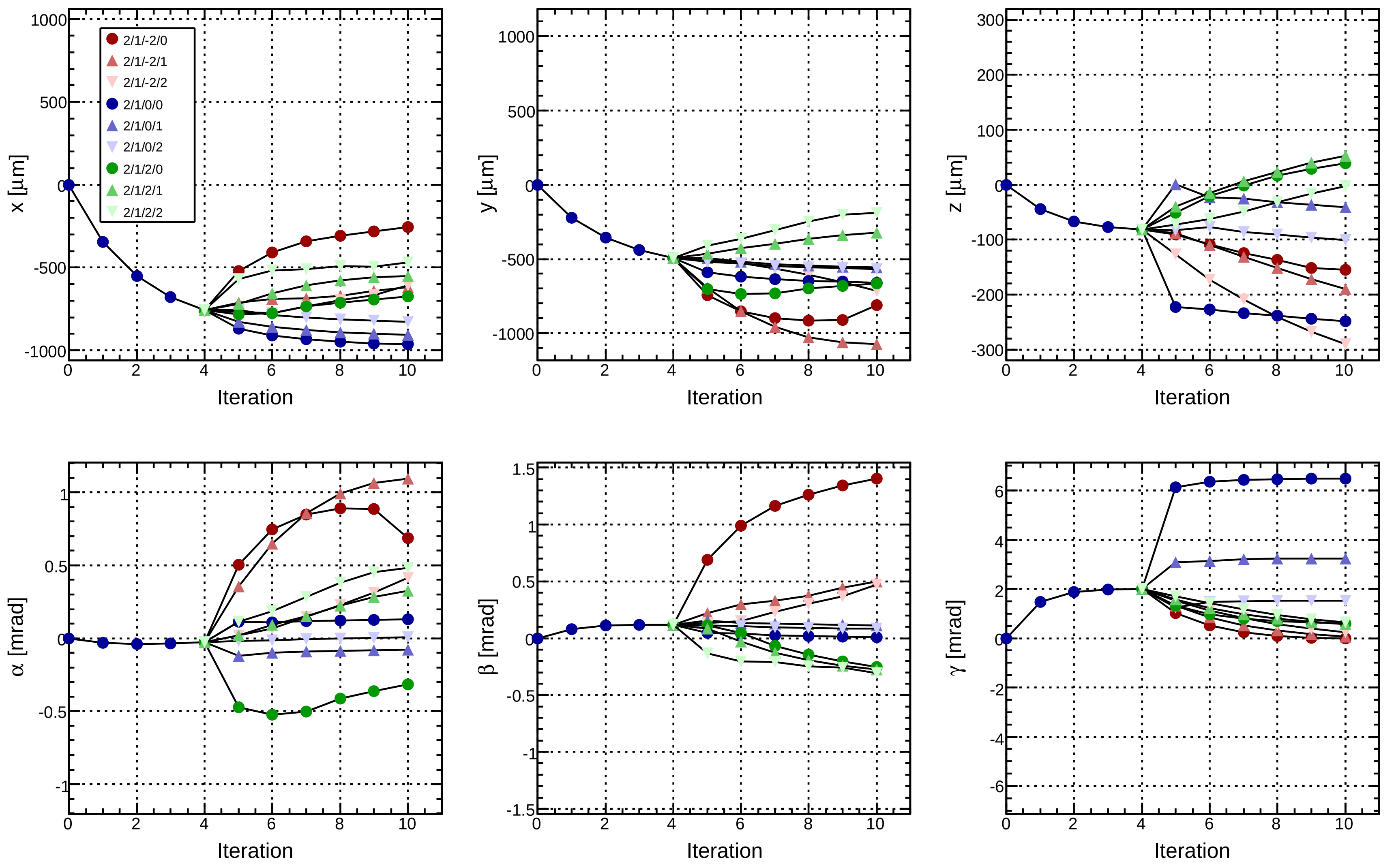


Figure 5.11: *Same as figure 5.5, but with a soft mode constraint applied to dampen the movements of weakly constrained degrees of freedom.*

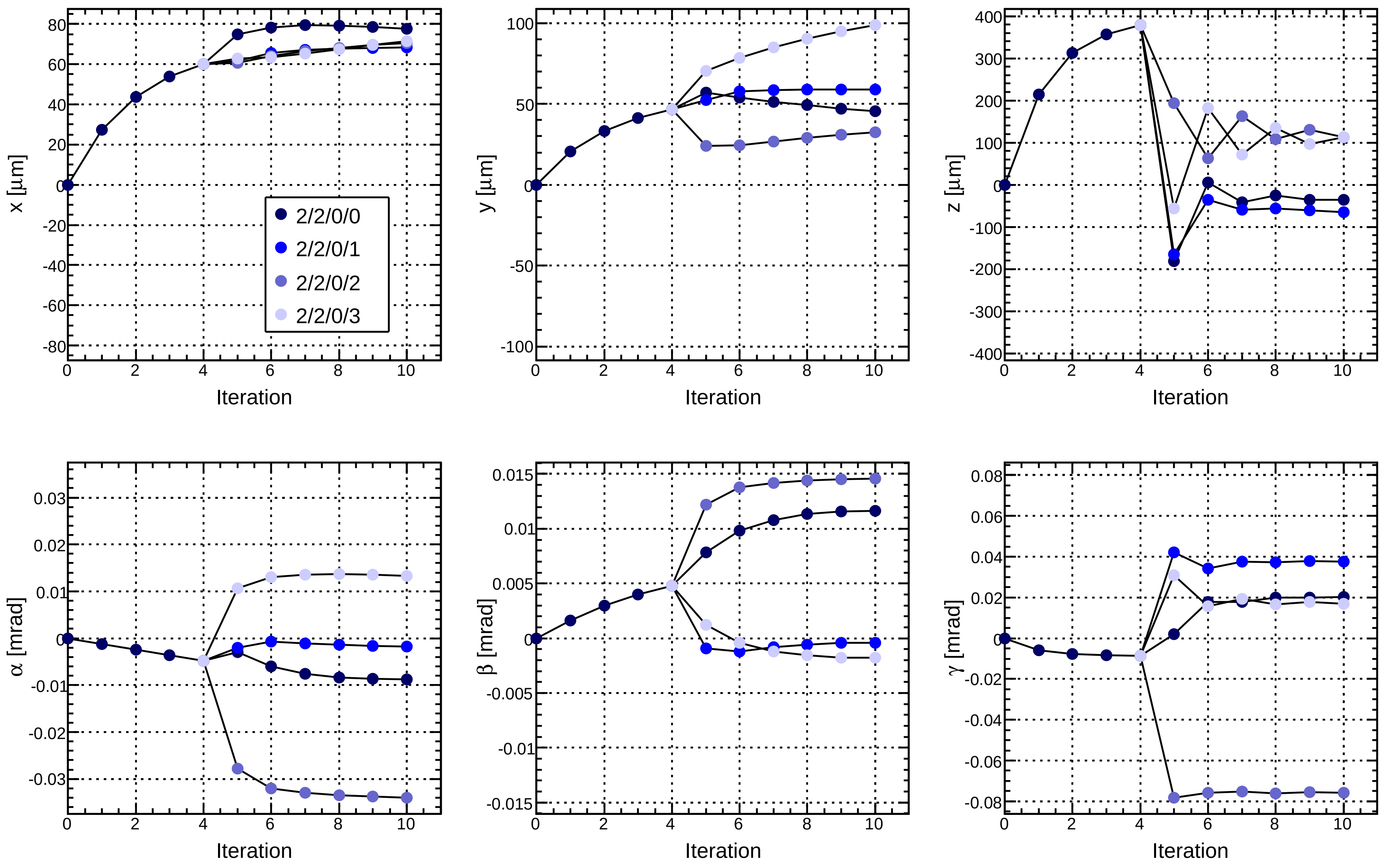


Figure 5.12: *Same as figure 5.6, but with a soft mode constraint applied to dampen the movements of weakly constrained degrees of freedom.*

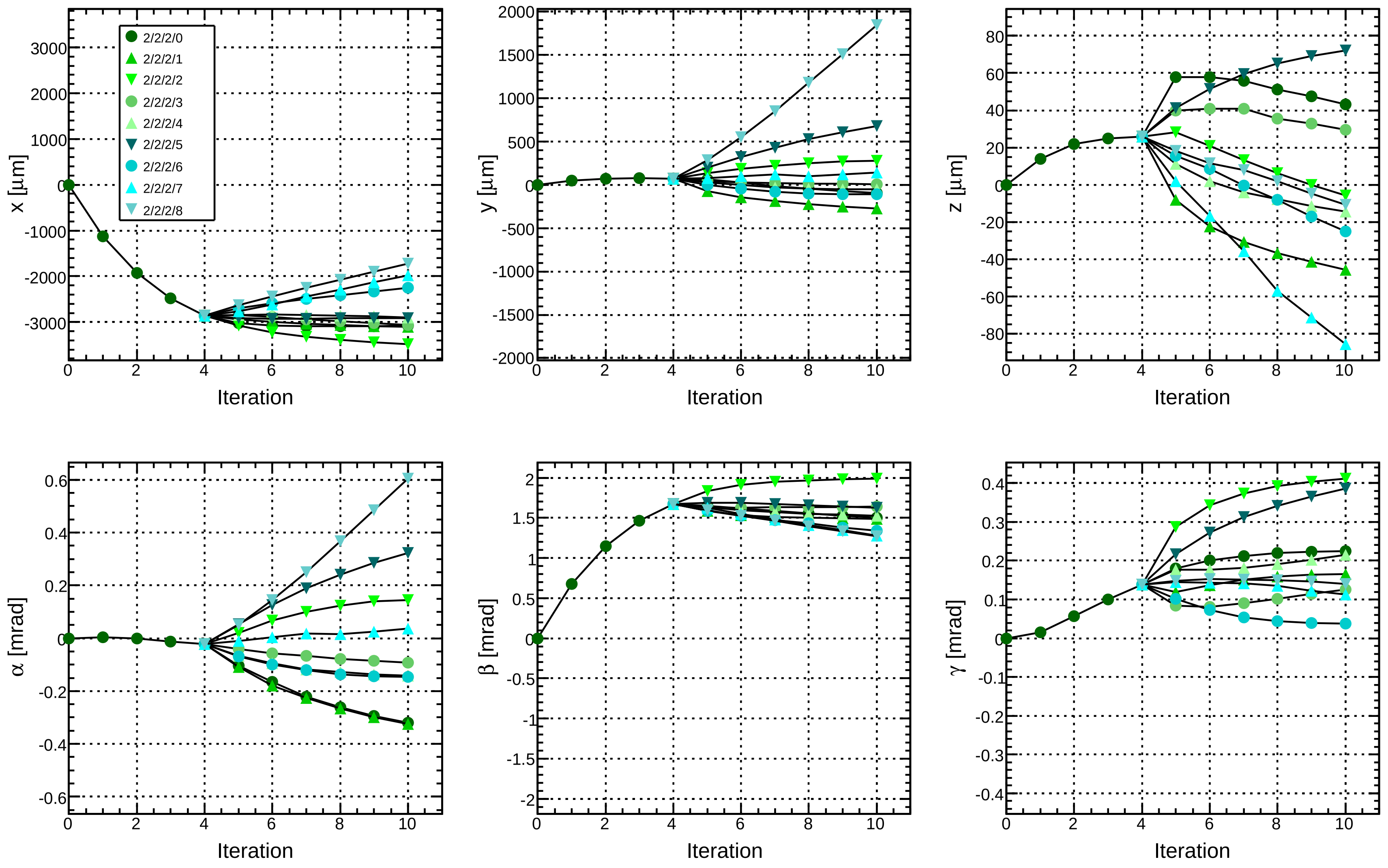


Figure 5.13: *Same as figure 5.7, but with a soft mode constraint applied to dampen the movements of weakly constrained degrees of freedom. Mind the different scales on the vertical axes compared to figure 5.7.*

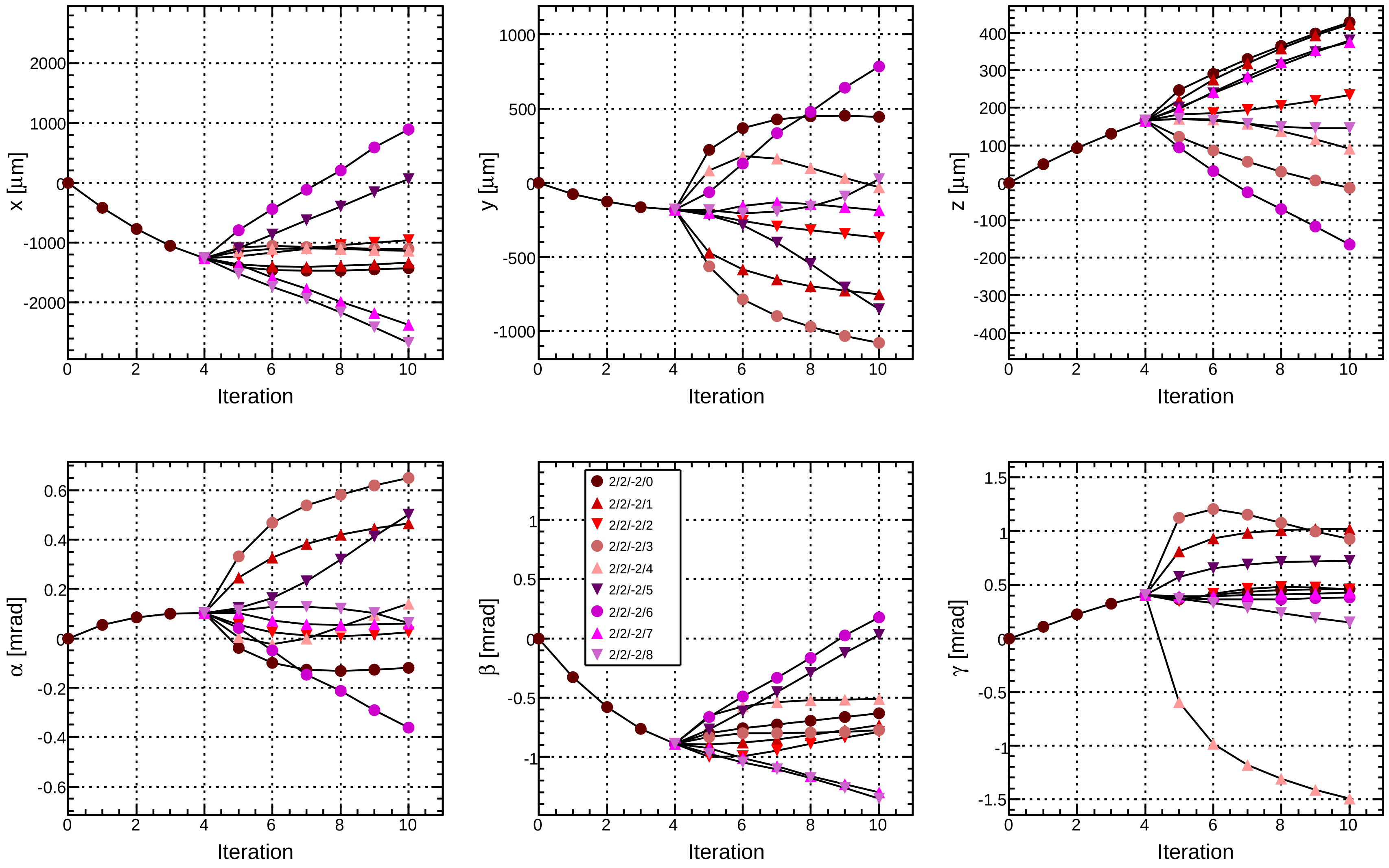


Figure 5.14: *Evolution of the six alignment parameters of the SCT endcap C disks during the first 10 iterations. A soft mode constraint is applied to dampen the movements of weakly constrained degrees of freedom compared to figure 5.8. Mind the different scales on the vertical axes compared to figure 5.8.*

Table 5.2: *Values chosen for the soft mode constraint uncertainties.*

	Pixel	SCT
σ_x	2 µm	5 µm
σ_y	2.5 µm	5 µm
σ_z	3.5 µm	12 µm
σ_α	50 μrad	40 μrad
σ_β	50 μrad	40 μrad
σ_γ	15 μrad	8 μrad

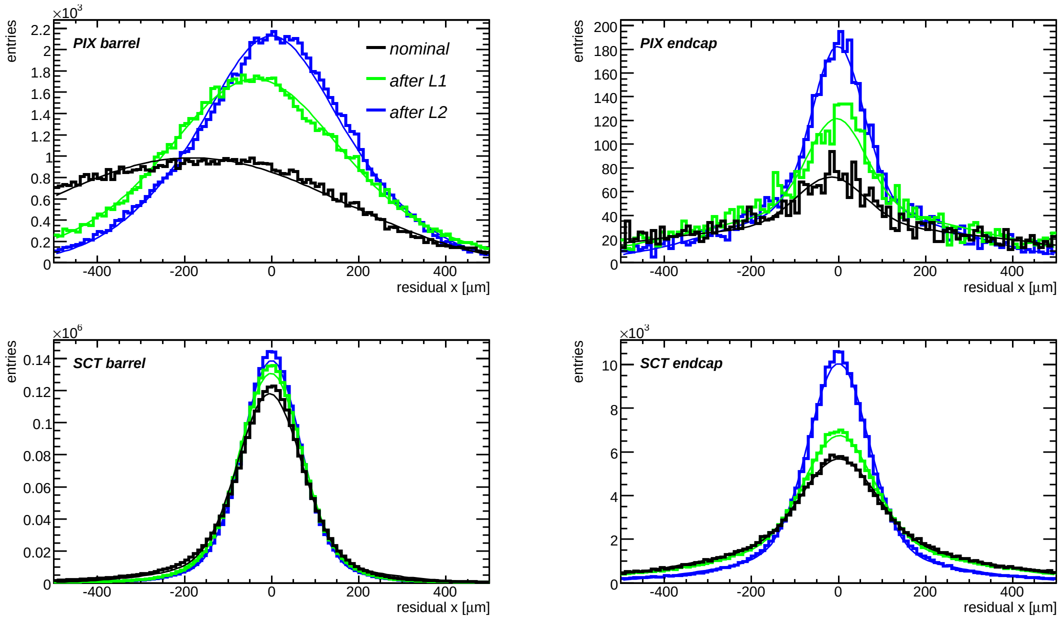


Figure 5.15: *Residual distributions on the four silicon subdetectors for x-residuals before alignment (nominal), after L1 alignment, and after L2 alignment.*

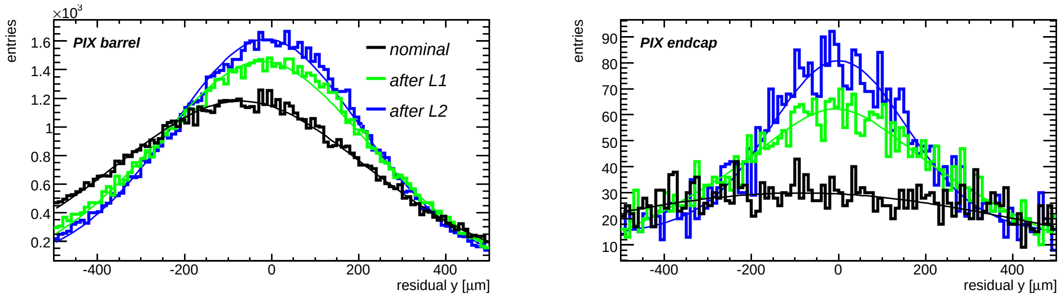


Figure 5.16: *Same as figure 5.15, but for the Pixel y-residuals.*

that share a common mean value μ, as given by

$$f(x) = c_1 e^{-0.5\left(\frac{x-\mu}{\sigma_1}\right)^2} + c_2 e^{-0.5\left(\frac{x-\mu}{\sigma_2}\right)^2} \quad . \tag{5.10}$$

This considers the fact that a narrow core distribution of well-aligned residuals is present, together with broader tails e. g. from noise, non-aligned modules and outliers. This procedure is also performed for the residual distribution of each individual layer and disk, and the resulting optimal parameters together with their one standard deviation statistical uncertainties are listed in tables 5.3 and 5.4.

Table 5.3: *Mean and widths of double-Gaussian fits to the DOCA residual distributions of Pixel structures after L2 alignment. The uncertainties quoted are one standard deviation from the fit procedure. For the x-residuals in barrel layer 0, the wider Gaussian does not converge.*

Pixel x-residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	-7.93 ± 2.65	62.52 ± 3.83	230.31 ± 11.36
endcap C disk 1	-3.87 ± 2.48	59.97 ± 3.64	250.77 ± 14.19
endcap C disk 2	-1.32 ± 2.47	46.56 ± 3.27	289.86 ± 15.87
barrel layer 0	10.74 ± 1.03	205.88 ± 2.10	N/A
barrel layer 1	23.32 ± 0.82	141.74 ± 2.76	368.90 ± 50.18
barrel layer 2	-16.16 ± 0.57	103.38 ± 5.09	195.60 ± 1.81
endcap A disk 0	3.41 ± 2.69	71.87 ± 5.17	241.16 ± 14.68
endcap A disk 1	-0.67 ± 2.39	59.04 ± 3.60	250.54 ± 13.98
endcap A disk 2	4.09 ± 2.66	51.73 ± 3.41	259.12 ± 17.09

Pixel y-residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	-14.26 ± 6.03	150.73 ± 12.17	459.29 ± 29.39
endcap C disk 1	-21.03 ± 6.02	152.67 ± 8.68	590.14 ± 48.34
endcap C disk 2	-5.12 ± 7.54	175.30 ± 9.96	720.69 ± 79.83
barrel layer 0	14.87 ± 0.93	192.38 ± 1.30	469.62 ± 34.25
barrel layer 1	-20.30 ± 0.80	209.27 ± 1.33	490.90 ± 26.80
barrel layer 2	-29.03 ± 0.85	258.30 ± 1.48	822.72 ± 118.94
endcap A disk 0	14.77 ± 5.50	186.98 ± 10.34	534.65 ± 56.51
endcap A disk 1	15.38 ± 5.60	150.78 ± 8.47	548.74 ± 37.78
endcap A disk 2	0.16 ± 7.13	169.28 ± 12.85	590.90 ± 73.70

The Pixel endcap x-residuals are considerably narrower and better centred than the barrel residuals, which means that the disk modules are placed more accurately. The tendency remains also for the y-residuals, however not as strong. The residual distribution of barrel layer 1 exhibits a shoulder on one side of the Gaussian which was not removed by L2 alignment. The later iterations will align the barrel modules individually, so that the widths of their residual distributions should decrease drastically. Since the endcap disks will not be aligned at module level, only minor improvements due to an overall improved alignment are expected there. The SCT residual distributions are well centred

Table 5.4: *Same as table 5.3, but for the SCT structures.*

SCT residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	6.59 ± 0.47	75.34 ± 0.75	235.42 ± 3.91
endcap C disk 1	-1.56 ± 0.40	74.22 ± 0.65	232.12 ± 2.97
endcap C disk 2	1.17 ± 0.49	79.57 ± 0.79	249.04 ± 4.13
endcap C disk 3	-4.40 ± 0.56	86.72 ± 0.95	244.70 ± 6.36
endcap C disk 4	-2.22 ± 0.54	76.19 ± 0.85	228.82 ± 5.08
endcap C disk 5	-2.47 ± 0.92	72.93 ± 1.23	256.54 ± 8.28
endcap C disk 6	5.93 ± 5.28	78.09 ± 9.80	380.03 ± 41.47
endcap C disk 7	-11.34 ± 2.48	72.42 ± 4.25	216.15 ± 17.74
endcap C disk 8	19.18 ± 4.12	84.13 ± 6.21	310.15 ± 78.27
barrel layer 0	0.57 ± 0.07	68.09 ± 0.11	175.79 ± 1.03
barrel layer 1	-1.19 ± 0.06	63.85 ± 0.09	168.17 ± 0.86
barrel layer 2	0.23 ± 0.06	65.25 ± 0.08	174.04 ± 0.75
barrel layer 3	-0.82 ± 0.06	72.00 ± 0.10	182.15 ± 0.62
endcap A disk 0	-6.90 ± 0.39	64.85 ± 0.52	232.55 ± 3.12
endcap A disk 1	1.97 ± 0.35	65.18 ± 0.49	230.81 ± 2.71
endcap A disk 2	-3.03 ± 0.38	61.64 ± 0.53	233.90 ± 3.09
endcap A disk 3	6.49 ± 0.38	59.72 ± 0.44	230.17 ± 4.16
endcap A disk 4	1.96 ± 0.42	62.49 ± 0.53	219.62 ± 4.24
endcap A disk 5	-0.08 ± 0.58	53.14 ± 0.74	200.54 ± 4.63
endcap A disk 6	0.37 ± 0.88	59.98 ± 1.20	203.19 ± 10.27
endcap A disk 7	-5.06 ± 1.62	65.57 ± 2.85	185.82 ± 13.85
endcap A disk 8	3.45 ± 2.19	63.75 ± 3.01	231.44 ± 22.60

and their widths are about the same size for barrel and endcap, which is caused by the L2 prealignment for the SCT.

To infer the amount of remaining random L3 module misalignment, residual distributions are also obtained for a simulated cosmic ray data sample. In this dataset, 86k tracks going through the Inner Detector volume were simulated with a perfectly known detector geometry. The distribution of DOCA residuals is histogrammed and fitted as for the data and the results are given in tables 5.5 and 5.6. The resulting widths of the Gaussian distributions give an estimate of the final width expected from a perfect alignment of the detector. For the y-residuals of half of the Pixel structures, a single Gaussian curve is obviously a good approximation of the residual distributions with perfect alignment knowledge; in these cases the width of the larger Gaussian acquires a large and, within its uncertainty, insignificant value.

Table 5.5: *Same as table 5.3, but for the simulated 86k cosmic tracks.*

Pixel x-residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	1.87 ± 2.36	28.02 ± 3.30	171.97 ± 37.32
endcap C disk 1	1.85 ± 2.29	29.78 ± 3.06	230.17 ± 66.45
endcap C disk 2	-0.14 ± 2.61	33.11 ± 3.27	274.65 ± 57.03
barrel layer 0	-0.33 ± 0.23	14.87 ± 0.31	45.16 ± 2.38
barrel layer 1	-0.06 ± 0.18	15.11 ± 0.28	38.85 ± 1.24
barrel layer 2	0.32 ± 0.18	17.01 ± 0.29	50.24 ± 1.55
endcap A disk 0	-0.53 ± 2.12	30.73 ± 3.05	249.99 ± 54.89
endcap A disk 1	3.63 ± 1.98	24.19 ± 2.23	294.61 ± 75.90
endcap A disk 2	-2.44 ± 2.83	27.19 ± 3.66	320.54 ± 88.20

Pixel y-residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	16.59 ± 7.73	111.25 ± 9.62	943.21 ± 1046.90
endcap C disk 1	8.39 ± 6.76	102.18 ± 4.82	N/A
endcap C disk 2	-0.56 ± 7.62	105.37 ± 8.23	923.88 ± 543.09
barrel layer 0	-0.15 ± 1.44	113.59 ± 1.17	1081.78 ± 805.45
barrel layer 1	1.96 ± 1.13	118.67 ± 0.89	785.00 ± 247.44
barrel layer 2	-1.77 ± 1.06	125.90 ± 1.02	468.55 ± 45.86
endcap A disk 0	-8.11 ± 5.84	71.13 ± 9.11	253.14 ± 44.60
endcap A disk 1	-18.61 ± 7.68	111.85 ± 7.81	1250.60 ± 1479.77
endcap A disk 2	-17.61 ± 11.42	120.07 ± 11.77	1950.59 ± 6669.27

The remaining L3 misalignments at this stage are then estimated by comparing tables 5.3 and 5.4 with 5.5 and 5.6, assuming a quadratic sum of the intrinsic resolution and the remaining misalignment. The remaining misalignments are found to be well below 100 µm for the x-direction of the Pixel endcaps and in the order of 50–60 µm. The barrel layers are not fully centred yet, although the alignment parameters have converged on L2. For the y-direction, the misalignments are in the order of 100–200 µm. For the SCT, the misalignments are in the order of 50–80 µm.

Table 5.6: *Same as table 5.5, but for the SCT structures.*

SCT residuals

structure	μ [µm]	σ_{small} [µm]	σ_{large} [µm]
endcap C disk 0	0.34 ± 0.37	30.09 ± 0.54	99.53 ± 6.28
endcap C disk 1	-0.04 ± 0.35	33.29 ± 0.40	160.47 ± 9.58
endcap C disk 2	-0.24 ± 0.43	33.01 ± 0.57	130.29 ± 5.99
endcap C disk 3	0.04 ± 0.41	31.09 ± 0.86	107.00 ± 15.39
endcap C disk 4	0.14 ± 0.45	29.46 ± 0.86	72.40 ± 4.14
endcap C disk 5	0.24 ± 0.81	37.08 ± 1.25	121.21 ± 13.29
endcap C disk 6	0.39 ± 1.17	24.20 ± 3.94	58.64 ± 5.24
endcap C disk 7	-0.54 ± 1.98	35.04 ± 1.85	357.26 ± 159.36
endcap C disk 8	1.93 ± 3.44	22.98 ± 12.00	63.59 ± 19.56
barrel layer 0	0.06 ± 0.08	25.37 ± 0.11	64.15 ± 1.00
barrel layer 1	0.02 ± 0.07	24.62 ± 0.11	61.79 ± 0.71
barrel layer 2	-0.01 ± 0.07	25.32 ± 0.10	68.44 ± 0.67
barrel layer 3	0.06 ± 0.08	30.04 ± 0.11	87.94 ± 0.82
endcap A disk 0	-0.38 ± 0.38	31.43 ± 0.55	124.37 ± 10.20
endcap A disk 1	-0.09 ± 0.36	30.09 ± 0.72	89.33 ± 6.31
endcap A disk 2	-0.00 ± 0.43	34.36 ± 0.49	190.64 ± 10.42
endcap A disk 3	0.54 ± 0.42	30.64 ± 0.84	91.25 ± 9.50
endcap A disk 4	-0.14 ± 0.46	30.63 ± 0.85	74.35 ± 4.98
endcap A disk 5	0.12 ± 0.77	33.91 ± 1.11	124.71 ± 11.09
endcap A disk 6	-0.50 ± 1.16	38.58 ± 1.20	284.59 ± 74.36
endcap A disk 7	1.00 ± 1.55	30.34 ± 1.49	389.58 ± 153.90
endcap A disk 8	-0.93 ± 3.00	32.34 ± 3.49	259.23 ± 107.07

Introduction of Level32 alignment

Since the Pixel and SCT barrel layers seem to have reached full convergence after 10 iterations, there is no need for further L2 iterations. Therefore the next 15 iterations are performed on L32, to make optimal use of the statistics available in the cosmic ray data. The L32 alignment replaces L3 alignment with fixed endcap disks [53]. In the latter case, only barrel modules were aligned, since the number of hits is not sufficient for module alignment on the endcaps. Afterwards, further L2 iterations were introduced to reconnect the barrel and endcaps. The L32 alignment overcomes this restriction, which should manifest itself in a smooth alignment parameter evolution without any kinks in the endcap disks. Technically, the local χ^2 algorithm combines endcap modules to disks and treats them as in L2 alignment, i. e. calculates residual derivatives and the alignment correction in the global coordinate frame. Barrel modules are treated as in L3, i. e. with the derivatives and alignment corrections determined in the local frames of the modules.

Addition of pseudo survey constraints

For the L32 iterations, another constraint is introduced to avoid a strict requirement on the number of hits a module acquires. Figure 5.17 shows the number of hits of all pixel barrel modules versus the errors of the six alignment parameter corrections. The $1/\sqrt{N}$-behaviour of the statistical uncertainty on the number of hits N is visible. Usually, a minimum number of hits ensures that the alignment solution fulfills certain criteria on the alignment parameter uncertainty. However, modules with too few or no hits at all—e. g. because they are switched off during data taking—are then not moved at all, leading to a discrepancy between aligned and non-aligned modules. To overcome the strict limit on the number of hits, the `InDetSurveyConstraintTool` [33] is used to provide pseudo survey constraints for the barrel modules. These constraints tie neighbouring modules on a barrel stave together by introducing penalty χ^2 terms on the deviation of the module distances on the stave from the surveyed distances. That means that alignment corrections that introduce a relative shift of neighbouring modules with respect to each other will be penalised. The procedure compares the “current” geometry, which is to be aligned, to a “reference” geometry, taken e. g. from a survey. For each stave, the relative movements between the two geometries are taken out by a χ^2-based minimisation procedure that overlays both stave representations. The distance between the “current” and the “reference” position of each module on the stave then serves as an additional residual, with its error determined by the survey precision.

Since the survey measurements were done in the assembly hall SR1 prior to lowering the barrel into the detector cavern, and prior to the complete Inner Detector assembly, deviations from the surveyed distances could have occured. Hereafter, it is not investigated how accurate the survey measurements still are, but the `InDetSurveyConstraintTool` is used with a constraint to nominal positions, hence the term pseudo survey constraint. The uncertainties that are assigned to the χ^2 penalty terms are not taken from the precision of the survey measurements (since those measurements are not even used), but derived in a similar fashion as in section 5.3.2. A minimum number of 100 hits is required for the pixel modules and the average uncertainty of the alignment parameter correction is read off figure 5.17. This value is the number of hits that was used for earlier alignment efforts, and lies in the region where the number of hits versus alignment correction uncertainty distribution becomes almost flat. The resulting uncertainties are also about the size of the alignment precision that is required by the ATLAS TDR (c. f. table 3.1). For the SCT, all active modules acquire more than 100 hits, and the uncertainties are set to constrain the few outlier points that can be seen in the alignment correction uncertainties in figure 5.18. An exception is made for the local y alignment correction in the SCT barrel: this is a weakly constrained degree of freedom, as can be seen from the large y alignment correction uncertainties in the figure, and often huge alignment corrections in y are assigned to SCT modules (in the order of mm, which is unphysical if the SCT barrel stays intact). Therefore, the constraint is placed at the most likely value of 75 µm (red point in figure 5.18. The values chosen for Pixel and SCT modules are summarised in table 5.7.

Table 5.7: *Values chosen for the pseudo survey constraint uncertainties.*

	Pixel	SCT
σ_x	4.5 µm	4 µm
σ_y	16 µm	75 µm
σ_z	12 µm	7 µm
σ_α	0.65 mrad	0.2 mrad
σ_β	2.00 mrad	0.3 mrad
σ_γ	0.25 mrad	0.1 mrad

The requirement on the minimum number of hits is then dropped and the ratio of the χ^2 from residuals and from the penalty term should determine how much the alignment correction is dominated by the track measurements or by the constraint. A module without any hit would be moved in correlation with its stave neighbours, such that those modules acquire a sensible pre-alignment, should they later be used in subsequent runs. The more accurate the alignment solution based on tracks gets, the less the module would be tied to its neighbours. Unfortunately it was found that a software defect prevents the pseudo survey constraint from properly working, which could not be resolved by the time of writing. Instead of a correction determined by relative movement, an additional zero residual with the chosen uncertainty is added to the total χ^2. This effect results in a soft constraint at module level, which is not destructing the alignment solution but not linking neighbouring modules either. It is seen that the constraint does not remove the appearance of large movements, especially in the y-coordinates of SCT barrel modules. This DOF is obviously difficult for alignment with cosmic ray data. Figure 5.19 depicts the alignment corrections for all barrel SCT modules and their six DOFs for the iterations 11–25. For comparison, the alignment correction for iteration 11, but not using the constraint, is filled into the bin of iteration 10. It can be seen that, although large alignment corrections appear in the local y-translation, the modules converge eventually. The final alignment parameters are obtained by summing up all corrections within the iterations. The constraint is most effective for the local y-translation, since in the other DOFs it was tuned to affect only a small number of modules. Correcting the software defect and installing the possibility to better constrain this DOF to avoid the unphysically large movements while not switching off local y-alignment completely is subject to current investigation.

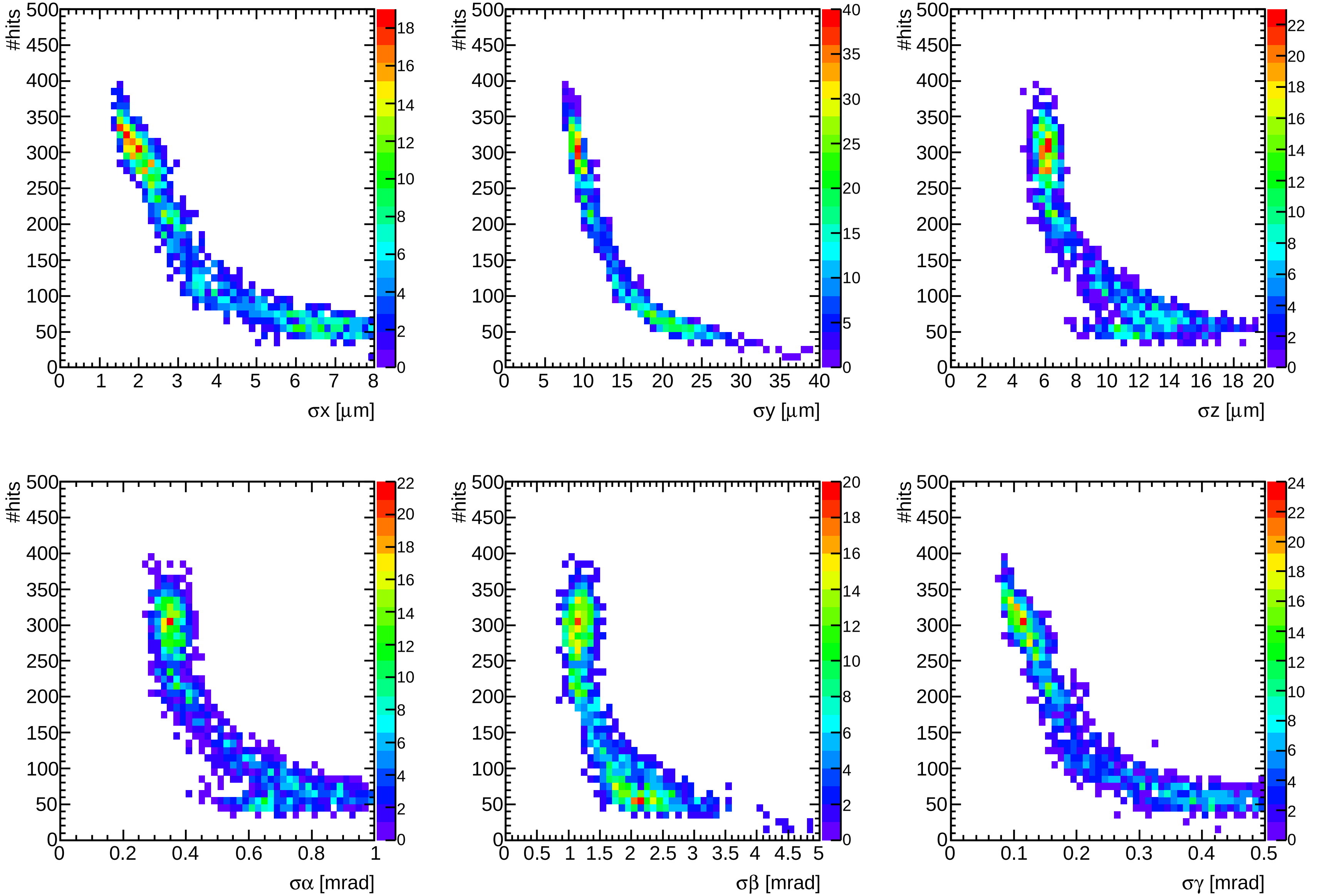


Figure 5.17: *Numbers of hits versus the uncertainties of the alignment corrections for Pixel barrel modules in iteration 11 without application of a survey constraint.*

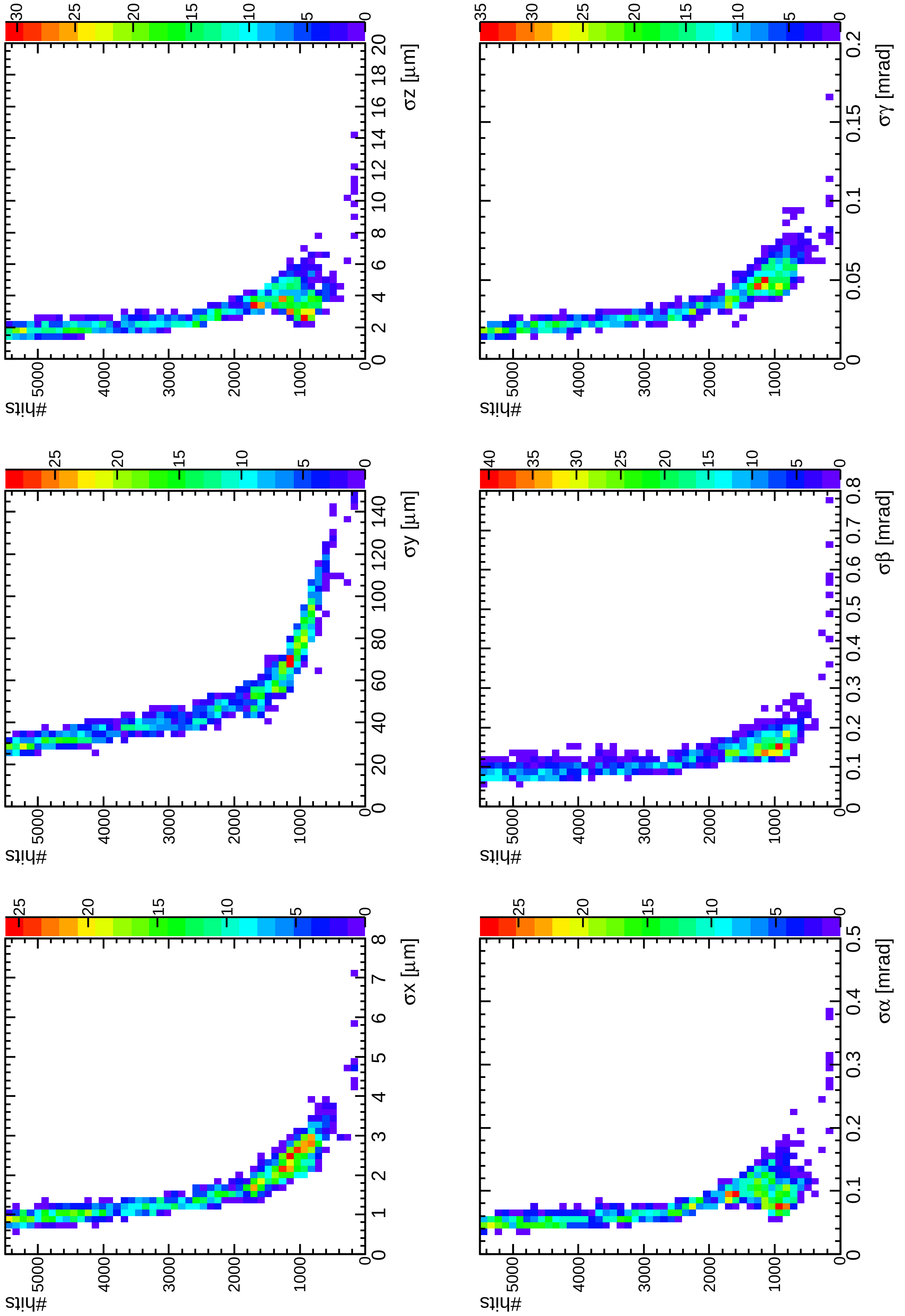


Figure 5.18: *Same as figure 5.17, but for the SCT barrel modules.*

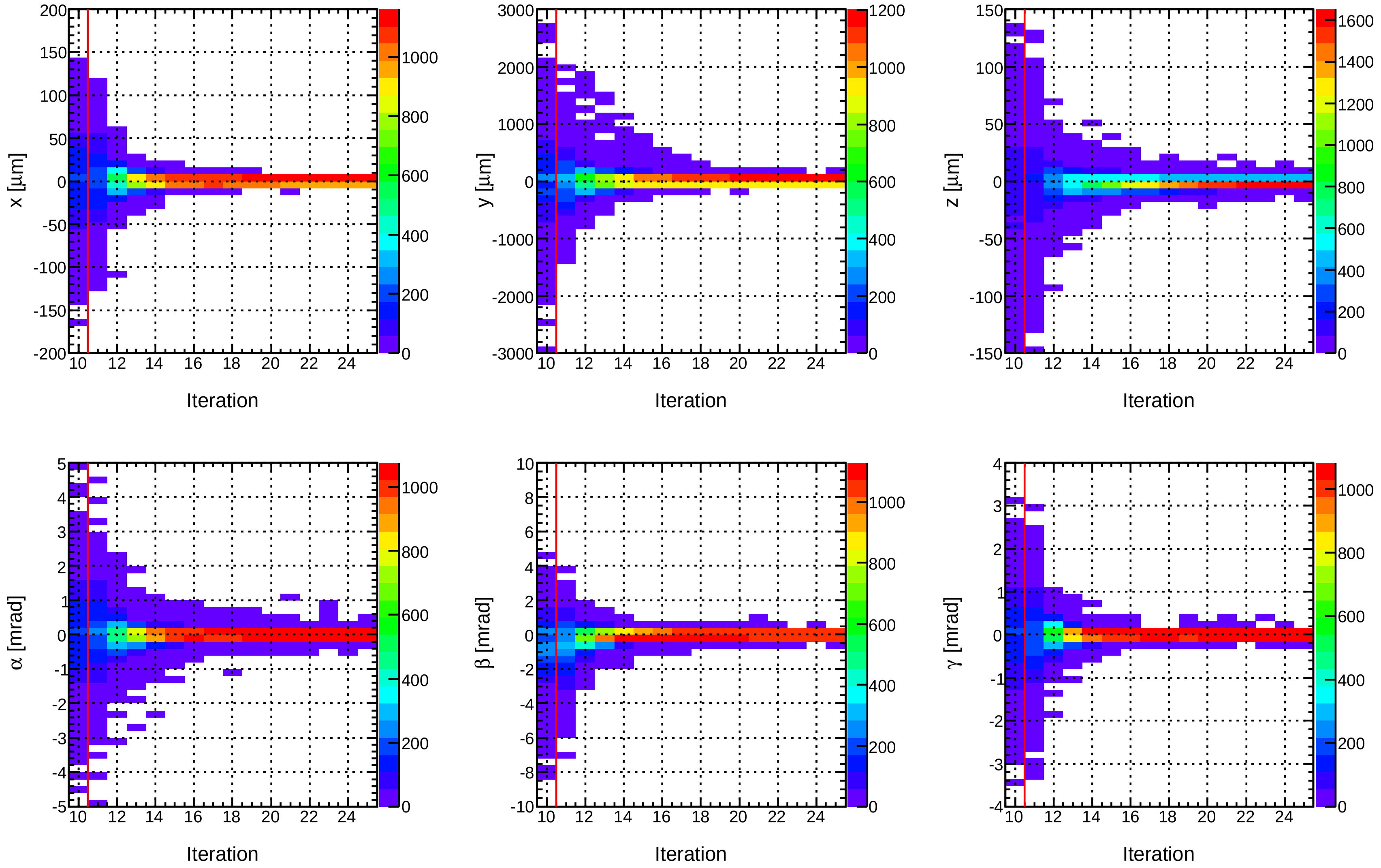


Figure 5.19: *Alignment correction occurrence within iterations 11–25 for the six DOFs of SCT barrel modules. For comparison, the alignment corrections for iteration 11, but without using the soft constraint on L32 are displayed at "iteration" 10.*

Final iterations at L2

After iteration 15, some of the SCT endcap DOFs show no sign of convergence. It is decided to fix these problematic DOFs, while trying to align as many disks as complete as possible. By looking at the alignment parameter evolution, the following DOFs were found to be diverging and were therefore fixed: α and β for SCT endcap C disks 5, 6, and 8; additionally β for disk 7 and Z for disk 6. In the SCT endcap A, Z is fixed for disks 6 and 7 and α for disk 8. The remaining 10 iterations on L32 and the final 10 iterations on L2 are performed without calculating alignment solutions for these DOFs. The software to do this has never been used in the local χ^2 alignment code before and needed to be repaired and tested. The code drops the corresponding rows and columns in the vector $\boldsymbol{b}_k$ and matrix A_k, and only solves e. g. a 3×3 system. Then it expands the solution back to 6 dimensions, filling the fixed parameters with zeros.

The evolution of alignment parameters of the Pixel structures is given in figure 5.20. After the final iterations, all endcap disks have nearly achieved convergence, and the barrel layers stay at their converged positions. The transition from L32 to L2 alignment in iteration 26 shows no kinks or irregularities in the diagrams, so it is concluded that L32 works as expected. The SCT barrel structures (figure 5.21) only show minor movements after the L32 alignment, which means that L3 and L2 alignment are consistent. For the SCT endcap A (figure 5.22), some of the disks do not converge in Z and an ongoing rotation in β can be observed for the outer disks. Figure A.3 shows the movement of the disks centrepoints in global coordinates and demonstrates that the X- and Y-movements have converged. For the SCT endcap C structures, it can be seen from figure 5.23 that most of the DOFs show a converging behaviour. The frozen α and β rotations also bring the compensating movements in X and Y to a halt. Some DOFs are not yet fully converged, especially the non-frozen α-rotations or Z-shifts for some disks. For completeness, the other figures showing the X and Y-coordinates of the layer and disk centrepoints are given in the appendix (figures A.1–A.4).

Overall, the alignment of the Pixel detector as well as the SCT barrel part works without problems. The SCT endcaps are hard to align with the given sample of tracks and many DOFs seem to not converge. It is presumed that the addition of tracks going through the SCT endcap region can improve the convergence of these structures drastically. The next evolution of the local χ^2 alignment constants will be derived using additional tracks that were triggered by the RPC and TGC triggers from the muon system, which will enhance the fraction of endcap tracks.

As summary, the settings that are used to derive the final alignment solution are again given in the following table:

iteration	level	soft mode term applied	SCT endcap DOFs fixed	survey constraint applied
1–4	L1			
5–10	L2	X		
11–15	L32	X		X
16–25	L32	X	X	X
26–35	L2	X	X	

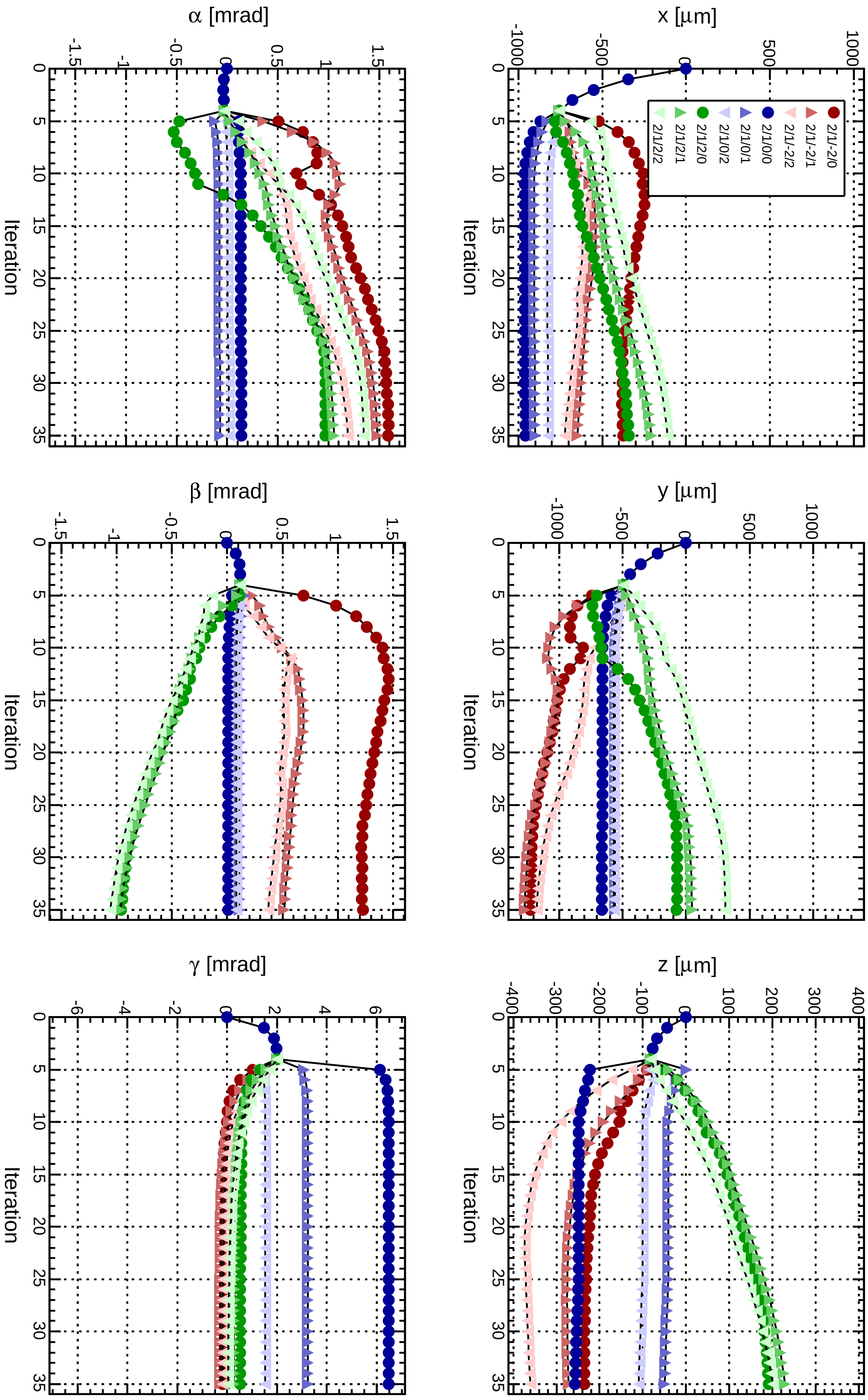


Figure 5.20: *Evolution of the six alignment parameters of the Pixel layers and disks during 35 iterations.*

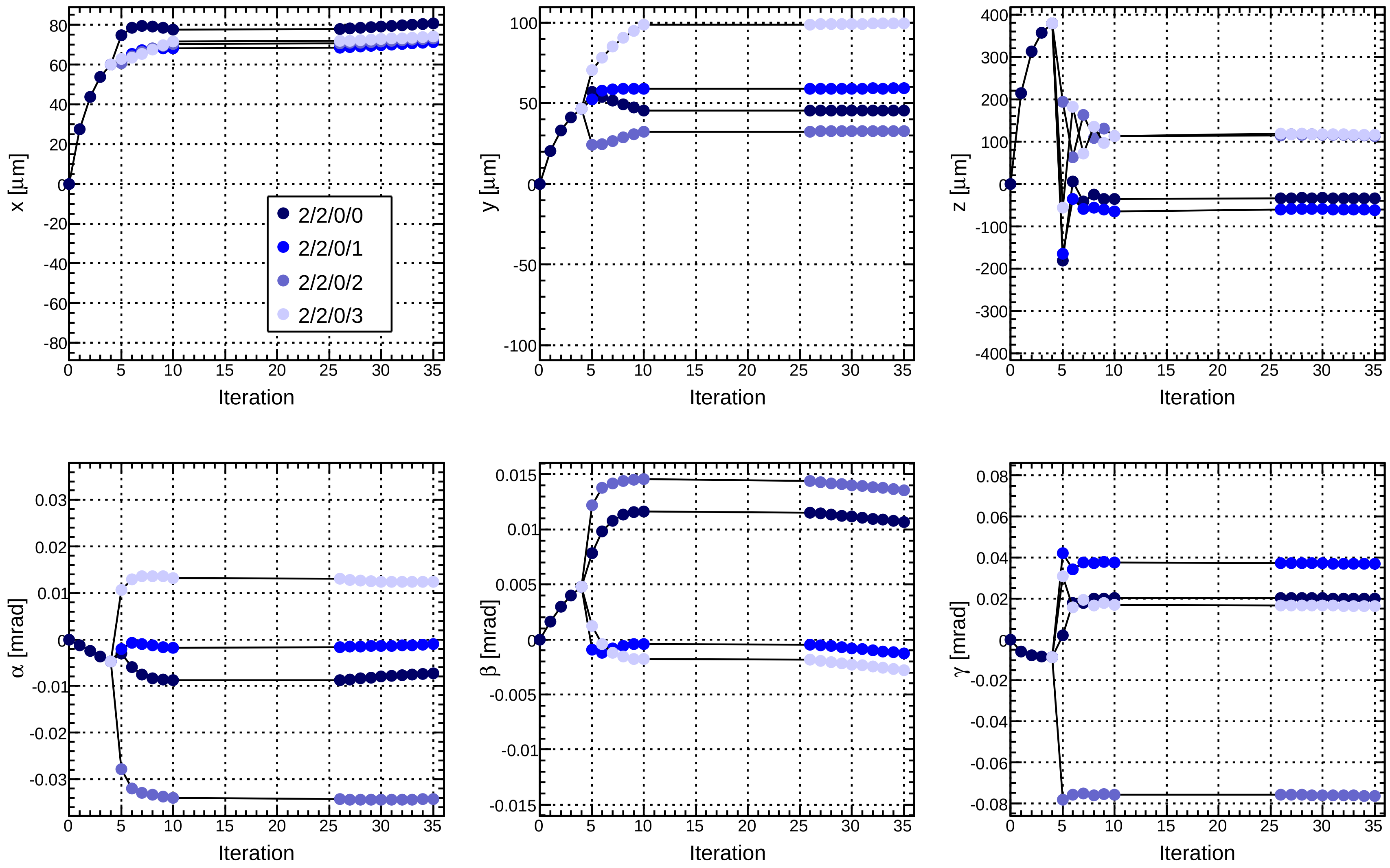


Figure 5.21: *Same as figure 5.20, but for the SCT barrel layers. Lines without markers indicate iterations where the SCT barrel is aligned on module level*

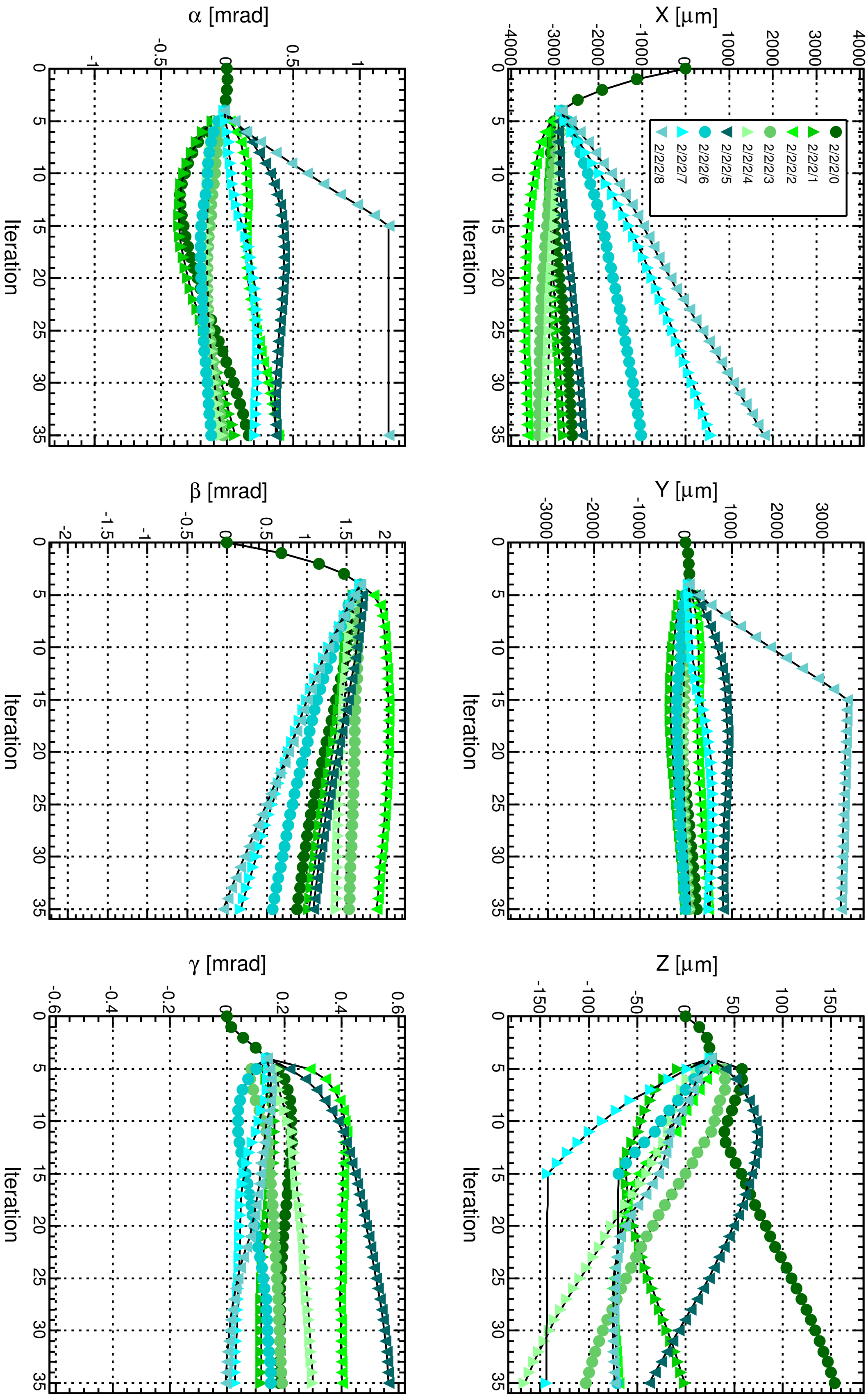


Figure 5.22: Same as figure 5.20, but for the SCT endcap A disks. Lines without markers indicate fixed DOFs for the corresponding disks.

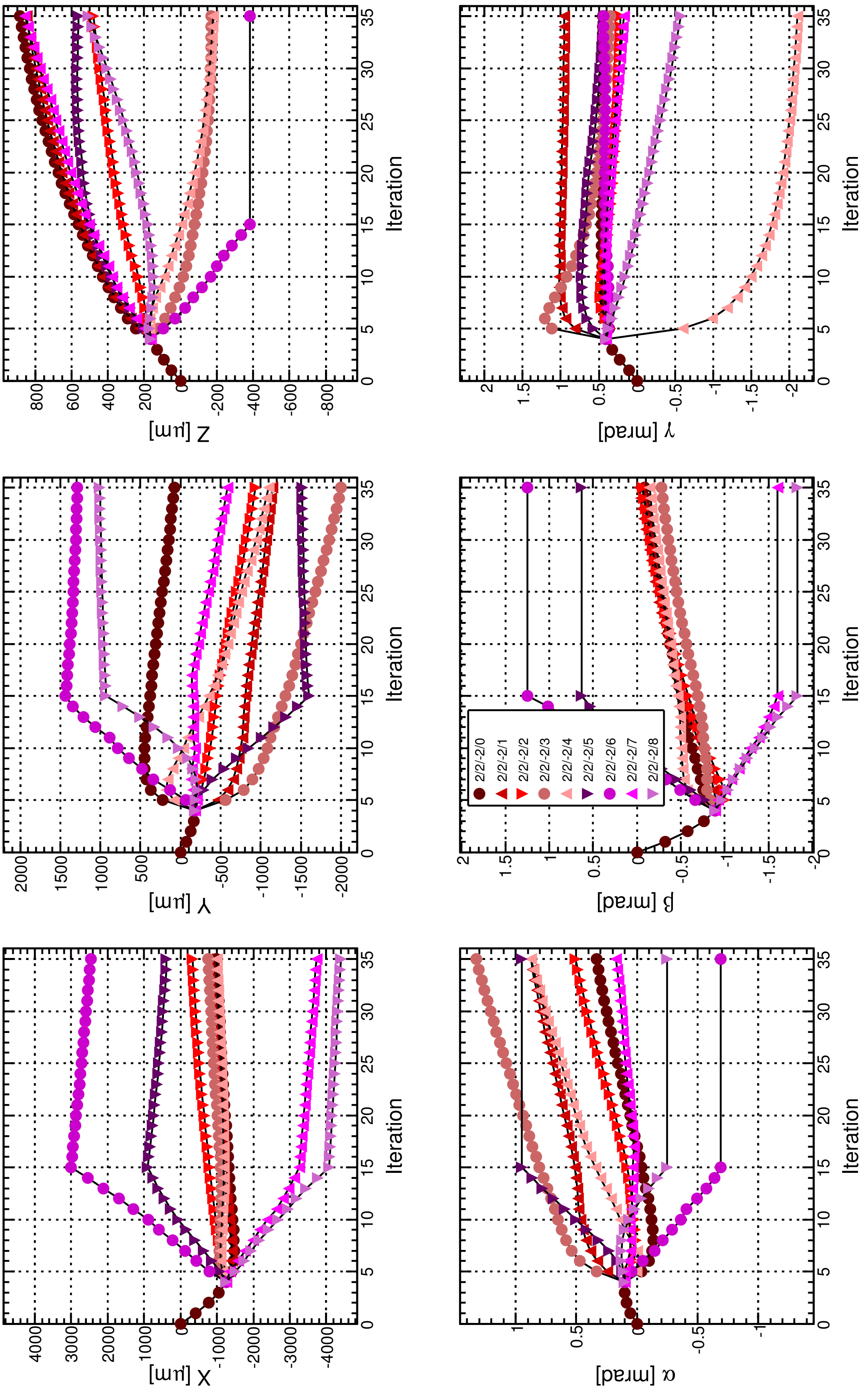


Figure 5.23: *Same as figure 5.22, but for the SCT endcap A disks.*

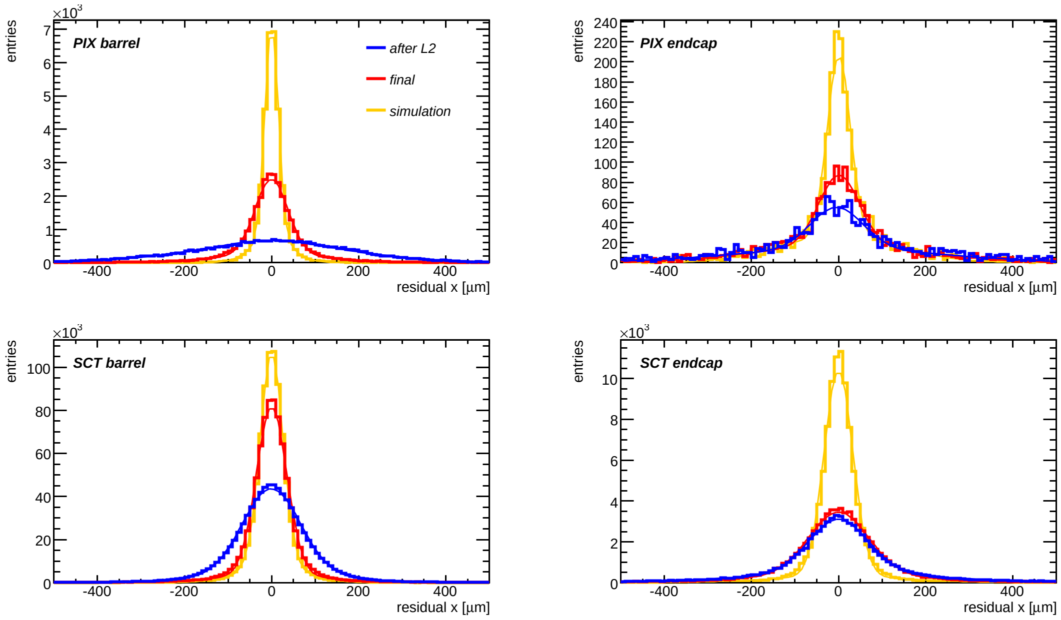


Figure 5.24: *Residual distributions for the four silicon subdetectors for x-residuals after L2 alignment and after the final alignment compared to those from the perfect alignment knowledge (simulation).*

5.3.3 Quality of the final alignment

The final alignment constants are used to derive residual distributions to infer the remaining amount of misalignment. The residual distributions are shown in figures 5.24 and 5.25, together with residuals from the simulated cosmic data based on perfect alignment. There are 940k residuals from each sample used (real data and simulation), to directly compare the residual distributions. Table 5.8 and 5.9 show the means and widths of double-Gaussian fits to the residual distributions after the 35 iterations. Comparing these to tables 5.5 and 5.6, a remaining random misalignment of (32–47) µm can be inferred for the Pixel endcaps in the x-directions. They benefit greatly from the ongoing L2 alignment. Also the y-residual widths go down and the remaining misalignment is mostly less than 100 µm. However, the L2 alignment centres the endcap disks better for the x-residuals than the y-residuals. This is because x-residuals have smaller uncertainties associated to them and no module-wise alignment is attempted, so that smaller x-misalignment is favoured over smaller y-misalignment. The SCT endcaps cannot improve on the residual widths, probably because they were reasonably prealigned on L2, and stay at (40-90) µm remaining misalignment. The barrel parts improve drastically due to the per-module alignment; the pixel barrel modules show a remaining misalignment in the order of (33–38) µm in x and around (50–75) µm in the y-direction. The SCT modules acquire a lot more hits and achieve a remaining misalignment in the order of (21–28) µm. This is the final precision on module level that can be achieved with the given dataset, since it is seen that the residual widths after 10 and after 15 iterations on L32 only differ marginally.

To test the consistency of the alignment solution, the `InDetAlignMonitoring` algorithm is used to create distributions from split tracks. Every cosmic track that has enough hits in the silicon detectors is split into an upper half and a lower half at $Y = 0$, which are

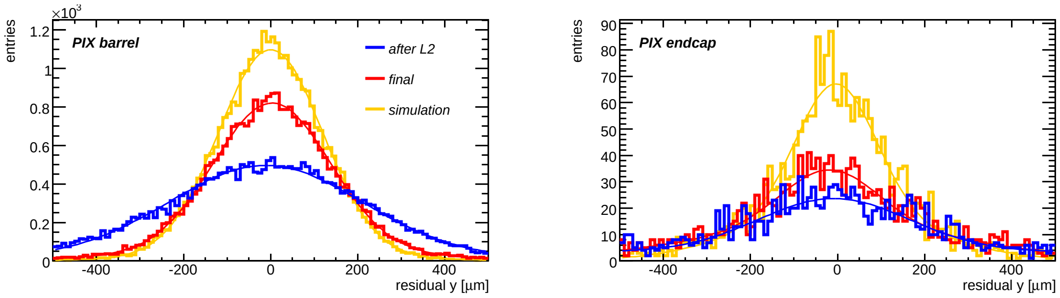


Figure 5.25: *Same as figure 5.24, but for y-residuals in the Pixel detector.*

Table 5.8: *Means and widths of double-Gaussian fits to the DOCA residual distributions of Pixel structures after 35 iterations.*

Pixel x-residuals

structure	$\mu/\mu m$	$\sigma_{small}/\mu m$	$\sigma_{large}/\mu m$
endcap C disk 0	-3.74 ± 1.85	54.69 ± 3.25	186.93 ± 11.47
endcap C disk 1	-1.89 ± 1.65	46.03 ± 1.85	209.78 ± 8.95
endcap C disk 2	-4.13 ± 1.83	46.14 ± 2.19	241.59 ± 10.44
barrel layer 0	1.14 ± 0.21	36.49 ± 0.26	147.62 ± 2.54
barrel layer 1	-0.69 ± 0.17	37.74 ± 0.23	146.24 ± 1.91
barrel layer 2	-0.92 ± 0.16	41.55 ± 0.24	149.41 ± 1.59
endcap A disk 0	2.16 ± 1.79	51.95 ± 2.43	203.86 ± 8.34
endcap A disk 1	2.02 ± 1.67	50.67 ± 2.20	202.95 ± 9.46
endcap A disk 2	2.48 ± 2.04	51.44 ± 2.74	243.99 ± 19.96

Pixel y-residuals

structure	$\mu/\mu m$	$\sigma_{small}/\mu m$	$\sigma_{large}/\mu m$
endcap C disk 0	-40.14 ± 4.21	142.74 ± 7.16	400.72 ± 32.34
endcap C disk 1	-21.39 ± 4.63	127.60 ± 8.21	403.86 ± 22.37
endcap C disk 2	-17.34 ± 5.08	134.14 ± 7.28	476.82 ± 21.94
barrel layer 0	2.00 ± 0.61	125.41 ± 0.58	435.80 ± 15.76
barrel layer 1	2.56 ± 0.49	133.11 ± 0.47	497.43 ± 14.53
barrel layer 2	3.98 ± 0.50	146.17 ± 0.55	465.43 ± 8.94
endcap A disk 0	7.31 ± 4.21	156.46 ± 5.93	441.35 ± 31.22
endcap A disk 1	12.90 ± 4.23	139.55 ± 5.87	484.93 ± 32.27
endcap A disk 2	2.81 ± 5.31	141.99 ± 6.78	543.57 ± 47.18

Table 5.9: *Same as for table 5.8, but for the SCT structures.*

SCT residuals

structure	μ/µm	σ_{small}/µm	σ_{large}/µm
endcap C disk 0	3.54 ± 0.38	69.22 ± 0.54	219.12 ± 3.87
endcap C disk 1	-0.43 ± 0.34	72.75 ± 0.49	219.65 ± 3.44
endcap C disk 2	1.79 ± 0.43	80.48 ± 0.66	233.11 ± 4.46
endcap C disk 3	-1.88 ± 0.51	84.68 ± 0.91	216.95 ± 5.94
endcap C disk 4	0.70 ± 0.50	76.85 ± 0.82	206.29 ± 5.26
endcap C disk 5	-3.64 ± 1.00	84.85 ± 1.35	270.17 ± 11.65
endcap C disk 6	0.28 ± 6.50	95.57 ± 11.13	556.79 ± 142.14
endcap C disk 7	-12.33 ± 3.04	93.13 ± 6.58	222.74 ± 22.67
endcap C disk 8	14.48 ± 4.45	97.16 ± 4.68	1045.13 ± 2508.20
barrel layer 0	-0.35 ± 0.03	33.89 ± 0.04	137.63 ± 0.91
barrel layer 1	-0.12 ± 0.03	32.41 ± 0.04	124.68 ± 0.77
barrel layer 2	0.16 ± 0.03	33.56 ± 0.04	131.25 ± 0.63
barrel layer 3	-0.18 ± 0.04	40.96 ± 0.04	157.00 ± 0.51
endcap A disk 0	0.18 ± 0.36	67.48 ± 0.49	226.95 ± 3.93
endcap A disk 1	1.14 ± 0.31	66.20 ± 0.42	210.77 ± 2.91
endcap A disk 2	-5.22 ± 0.34	62.12 ± 0.44	217.52 ± 3.12
endcap A disk 3	3.10 ± 0.35	60.09 ± 0.42	206.14 ± 3.92
endcap A disk 4	1.41 ± 0.40	63.09 ± 0.53	196.53 ± 4.15
endcap A disk 5	1.33 ± 0.58	53.98 ± 0.83	181.29 ± 4.40
endcap A disk 6	2.63 ± 0.93	59.97 ± 1.19	188.86 ± 7.08
endcap A disk 7	-2.53 ± 1.66	69.67 ± 2.32	208.40 ± 12.94
endcap A disk 8	8.12 ± 2.29	67.25 ± 3.40	214.92 ± 22.89

Table 5.10: *Means and widths of double-Gaussian fits to the track parameter difference distributions. The common mean value and the smaller of the two Gaussian widths is quoted for the final alignment and the simulated data sample. The given uncertainties correspond to one standard deviation uncertainty from the fit procedure. An uncertainty of zero denotes a value smaller than 0.005.*

Δtrack parameter	μ_{final}	σ_{final}	μ_{simul}	σ_{simul}
Δd_0 [μm]	4.25 ± 2.58	69.86 ± 4.82	-1.09 ± 0.47	41.76 ± 0.59
$\Delta\phi$ [mrad]	0.02 ± 0.03	0.60 ± 0.09	-0.01 ± 0.01	0.39 ± 0.03
Δz_0 [μm]	-12.25 ± 6.50	209.18 ± 6.80	3.94 ± 2.88	179.47 ± 3.33
$\Delta\eta$ [10^{-3}]	-0.27 ± 0.01	2.53 ± 0.02	0.13 ± 0.00	2.08 ± 0.01
$\Delta(q/p_T)$ [TeV^{-1}]	-0.14 ± 0.10	3.65 ± 0.28	0.02 ± 0.04	2.92 ± 0.12

then fitted independently from each other. Unlike for the alignment, no additional track selection is applied on the reconstructed tracks. The distributions for the upper-to-lower differences of the 5 + 1 track parameters are shown in figure 5.26. The blue histograms show the six track parameter differences between upper and lower track segment of the tracks for the L2-aligned geometry after iteration 10. The red histograms represent the final aligned geometry. The distributions are compared to the expectations obtained from the simulated dataset, shown in orange, which gives an impression of the internal consistency of the alignment. However, a direct comparison of the distributions should be taken with caution since the simulated cosmic data sample does not fully match the real cosmic data, e. g. in terms of track parameter distributions. A double-Gaussian fit with a common mean value is performed, similar to the procedure described for the residual distributions. The fit range is chosen to cover half of the parameter range of the histogram, symmetrically around its mean value. The distribution for Δp_{T} is not fitted, since it is of the form of an inverse Gaussian. The obtained values for the μ parameter and the smaller of the two σ parameters for the final alignment and simulated distributions are given in table 5.10. It is seen that the track parameter differences yield about the same centrality for the final alignment on real data as they do for the simulated dataset. For the Δd_0 and Δz_0 parameters, the mean values are about 2 standard deviations away from zero, which is also the case for the simulated perfect geometry. The mean of the $\Delta\eta$ parameter, although being remarkably close to zero, is not compatible with zero within the error reported by the fit procedure. This is also true for the simulated dataset with at least the same level of confidence, indicating that this might be an effect imminent in the track reconstruction of the used Athena release. The widths of the distributions are in general not as slim as for the simulation, while in some parameters they already approach the "desired" values, e. g. for Δz_0 or $\Delta\eta$. It remains to be seen whether the simulated values give a realistic final precision achievable in future alignment efforts.

With the distributions being centred, or at least as central as the simulated distributions, it is concluded that the obtained alignment geometry is internally consistent. This also shows that no systematic deformations affecting upper and lower detector half differently are present, which cannot be inferred by centred residual distributions alone. The alignment on module level greatly improves the achievable track resolution, as seen by the differences between blue and red histograms in figure 5.26. The widths of the distributions with the final alignment are not as small as for other alignment sets produced, which in-

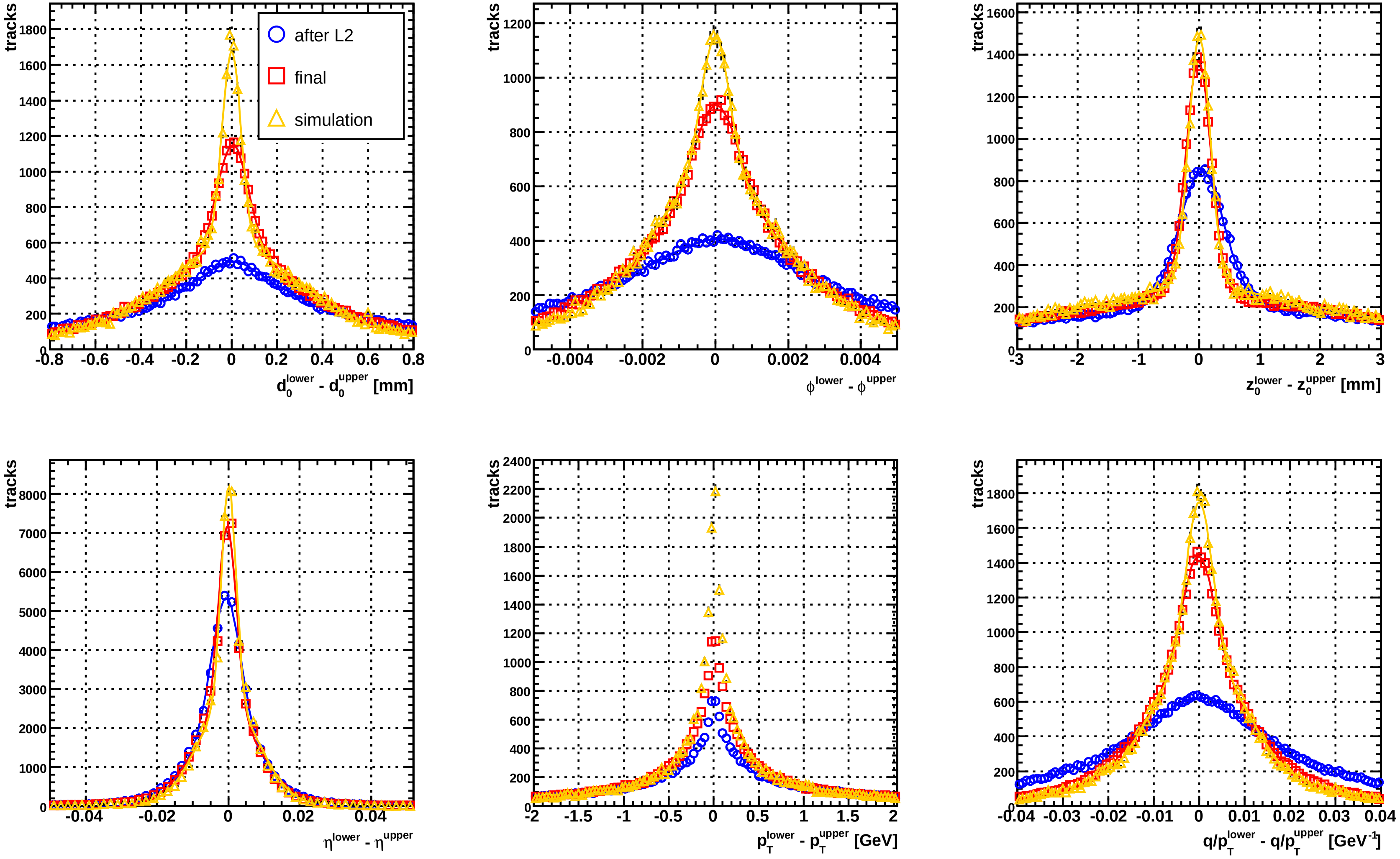


Figure 5.26: *Distributions of track parameter differences from split tracks for different alignment sets. For the variable $\Delta(q/p_T)$, tracks measured without magnetic field are taken out of the histogram. The distributions are normalised to the tracks statistics available for the simulated dataset.*

dicates the need for additional hit statistics available to the algorithm. An accurate Pixel barrel alignment plays the most important role for the precise reconstruction of perigee track parameters (but also the SCT barrel alignment is important, since SCT-only tracks are present in figure 5.26). For collision data, the use of beam and vertex constraints will additionally contribute to the exactness of the alignment solution in the inner parts of the Inner Detector.

The comparison between the previous local χ^2 alignment results [53] and the presented alignment is not a simple task, since some of the prerequisites of the production changed in the meantime. The previous effort had a sample of tracks at hand that was reconstructed using a baseline alignment for initial data taking for all three Inner Detector subdetectors, including the TRT. These prereconstructed tracks were used to deduct residuals using the unaligned, nominal geometry by stripping TRT hits and a subsequent refit. In contrast, the current alignment starts from a complete bottom-up approach without any prior alignment knowledge, i. e. initial track reconstruction including pattern recognition is performed from the nominal geometry, with the TRT detector being switched off in the software. Therefore, the hit-to-track association might suffer from a loss of hits that is not recovered during the alignment. However, the presented results show that bootstrapping the alignment works within the algorithm and leads to a consistent solution. It was seen that the obtained L1 alignment parameters are compatible with those obtained by the global χ^2 algorithm [94], which raises confidence that a sensible aligned geometry can be obtained without prior alignment knowledge. The alignment of the SCT endcaps proves to be a lot more challenging due to diverging DOFs not present before, which eventually have to be fixed during the alignment. This is attributed to the lack of TGC and RPC-triggered tracks and their addition will be one of the first tests that will follow the current alignment production. However, it is seen that the widths of the residual distributions of endcap hits compare well to the previous results, since the presented measures to tackle endcap alignment prevent disks with few hits from worsening the overall alignment, while allowing endcap disks with lots of hits to converge to appropriate positions. The widths of the residual distributions, obtained from the smaller of the two widths in a double-Gaussian fit, is $(69.4 \pm 0.4)\,\mu\text{m}$ for the SCT endcap residuals, compared to $(64.5 \pm 0.4)\,\mu\text{m}$ when using the previous alignment solution. For the Pixel endcaps they are $(47.1 \pm 2.2)\,\mu\text{m}$ compared to $(53.8 \pm 2.4)\,\mu\text{m}$ for x-residuals, and $(147.2 \pm 5.9)\,\mu\text{m}$ compared to $(148.1 \pm 7.5)\,\mu\text{m}$ for y-residuals. This shows that especially the Pixel endcap residuals are comparable to previous results. In the barrel, the residual distributions are worse predominantly for the Pixel barrel, as seen from widths of $(35.3 \pm 0.4)\,\mu\text{m}$ and $(135.2 \pm 0.8)\,\mu\text{m}$ compared to $(28.0 \pm 0.2)\,\mu\text{m}$ and $(121.6 \pm 0.6)\,\mu\text{m}$ for the x- and y-residuals. The SCT barrel residuals show a slight difference of $(32.8 \pm 0.1)\,\mu\text{m}$ compared to $(29.6 \pm 0.1)\,\mu\text{m}$.

The most striking difference that is conceived as a cause for the wider residual distributions is the lower number of hits present in the current alignment production compared to the previous effort. It is seen that the hit quality selection rejects 36 % of pixel hits and 21 % of SCT hits, reducing the number of pixel hits from about 113 % to about 70 % of the value previously available. Initially, there is a higher number of pixel hits available despite the smaller amount of tracks, since the improving detector geometry during alignment enables the pattern recognition to find more hits on the reconstructed tracks. The rejection of many pixel hits might overall degrade the pixel alignment more than the inclusion of many hits of moderate quality does. The applied selection criteria of the `HitQualSelTool` will therefore be revisited for subsequent alignment efforts to determine whether the inclusion of more hits is beneficial to the overall alignment precision.

5.4 Conclusions

The determination of alignment parameters of the silicon part of the ATLAS Inner Detector was performed using the local χ^2 alignment algorithm on cosmic ray data taken in autumn 2008. The alignment was done at disk level for the silicon endcaps and at module level for the barrel for all six degrees of freedom of the structures. Additional constraints were implemented to make the alignment solution more stable against non-converging DOFs, most likely induced by the special track geometry of the cosmic ray data, i. e. the tracks are almost parallel to the endcap disks. Ten iterations at L1 and L2 and 15 iterations at the newly introduced L32 were performed, followed by 10 final iterations at L2. After the first 10 iterations, the remaining random misalignments were estimated to be of the order of (50–80) µm for the Pixel endcaps and SCT modules. After aligning the barrel modules, 23 µm of misalignment is left on average in the SCT barrel and about 35 µm and (50–75) µm for the Pixel barrel x- and y-directions, respectively. The pixel endcaps achieve an improvement in the residual distributions and have well below 50 µm of random module misalignment left, whereas the SCT disks stay at (40–90) µm random misalignment. A comparison to a simulated dataset reveals not only these numbers, but also a comparison to the optimal resolution of track parameters. The track parameter resolution for data was determined in a fit to upper and lower track segments and both segments are found to be largely compatible in all five track parameters, with the optimal track parameters from simulation showing the same level of offset. The achieved resolution is still worse than for the simulated case. Further improvement of the residual and track parameter widths can only be expected with more data and better illumination of the detectors, e. g. by using tracks from additional trigger streams and/or relaxing the quality requirements imposed on the silicon hits.

Alignment using the local χ^2 approach is still a rather manual task, so that further developments before collision data taking should improve the automatic running capabilities of the local χ^2 algorithm. For example, the change of the total χ^2 can serve as a criterion to declare the alignment solution converged or to switch to a finer alignment level.

Several new constraints were implemented to deal with non-converging and problematic DOFs in the cosmic alignment. These were a soft constraint to limit movements of single structures, and a pseudo survey constraint to constrain module-to-module distances within a given uncertainty. The mechanism to freeze only a subset of DOFs of a module was repaired and applied to certain α- and β-rotations as well as Z-shifts of the disks 5–8 of both SCT endcaps. The convergence behaviour of the SCT endcaps was found to be significantly worse than for previous efforts and the future addition of tracks triggered by the muon system is deemed highly beneficial. In the future, the combination of tracks from particle collisions and from cosmic rays will stabilise many of the observed diverging DOFs. The strategy, which constraints to use, and how to use them, needs to be revisited then. The fact that there are several means of constraints at disposal, if problematic DOFs are encountered, is promising.

With the presented alignment solution, the silicon modules are aligned to a few tens of micrometers in the barrel part and well below 100 µm in the endcaps. The resulting detector geometry will be taken as a baseline to search for possible changes in the Inner Detector geometry during the last winter shutdown when analysing the new cosmic ray data from 2009. Additional track statistics will improve the alignment further and beam halo, beam-gas and eventually collision tracks will provide the required statistics and illumination to align the full detector, including all endcaps.

Chapter 6

Conclusions

The LHC successfully went into operation this year, providing proton-proton collisions of unprecedented energy for the exploration of Standard Model and beyond Standard Model physics. The ATLAS detector was built to utilise the full potential of the LHC for discovering new phenomena in high energy physics. The top quark serves as a valuable tool to investigate and test the Standard Model, and constitutes the background for many physics processes beyond the Standard Model. Within this thesis, several ways to reduce W+jets and combinatorial background to top quark mass analyses in the l+jets top quark pair decay channel were investigated. The $p_{\mathrm{T}}^{\mathrm{max}}$ method for top quark mass reconstruction was shown to choose a large fraction of wrong three-jet combinations. After a cut-based top quark pair decay selection procedure, the contribution from W+jets physics background can be reduced further by using the merging scales of an exclusive k_{T} jet algorithm as event shape variables. A single merging scale provides limited separation power, while using several of them in a multivariate method can increase sensitivity to signal-background separation. A Fisher discriminant on the first ten merging scales was investigated and found to yield 66.0 % additional W+jets background rejection, while keeping 63.5 % of $t\bar{t}$ events. While this method does not bias the correctly reconstructed jet triplets, its influence on the mass reconstruction from the total mass distribution needs to be taken into account.

Several kinematic variables were tested for their separation power against both W+jets and combinatorial background. Six of them were found to provide good separation without strong correlation to the top quark mass and were further used in a Fisher method. The resulting discriminant achieves a suppression of 87.1 % of W+jets events, plus 69.4 % on combinatorial background events, while keeping 74.8 % of correct jet triplets. The fraction of total background in the signal region around the peak can therefore be reduced from 67.7 % to 54.6 %. This method can be used as a further enhancement of the correctly found jet triplets in the reconstructed mass distribution to feed the template-based top quark mass analysis that is in preparation at the MPP top group.

A very accurate precision of the tracking devices is needed to fully exploit the track reconstruction capabilities of the ATLAS detector. This will amongst others enable a reliable b-quark induced jet identification and be very beneficial to top quark analyses, since it provides a way to suppress the predominant background processes. However, identification of secondary vertices, which is needed for b-identification to work, is crucially dependent on the accurate alignment of the Pixel detector. Also both electron and muon reconstruction algorithms base their four momenta measurements on tracks reconstructed by the Inner Detector.

In this thesis, Inner Detector tracks measured from cosmic ray muons were used in the local χ^2 alignment algorithm to align the silicon modules of the Inner Detector. This cosmic alignment on real data is meant as a baseline alignment for the start of the LHC. Several improvements of the method targeted at alignment using cosmic data were implemented and discussed: alignment on “level 32” provides module-by-module granularity in the barrel parts of the Inner Detector, while aligning disk-by-disk in the endcap parts. This way, a module-by-module alignment can be attempted for the well-illuminated barrel without the need of additional iterations of the algorithm to reconnect barrel and endcap geometries. A soft constraint was implemented to be able to constrain movements of large uncertainty on the alignment parameters. It was used to prevent the outer endcap disks from acquiring false large movements and thereby worsening the alignment solution of the other detector elements. Another constraint, the pseudo survey constraint, was meant to be used to constrain barrel modules with large alignment parameter uncertainties (i. e. usually a low number of hits) relative to their neighbouring modules. This can also provide a baseline alignment of modules which are not taking data at the moment. After a total of 35 iterations with the local χ^2 algorithm, the remaining random misalignments in the sensitive coordinate were found to be in the order of 35 μm (25 μm) for the barrel modules of the Pixel (SCT) detector and well below 50 μm (100 μm) for the endcap modules of the Pixel (SCT) detector. The quality of the alignment solution was judged by comparing residual distributions to simulated distributions and also by using a self-consistency measure, which reconstructs tracks split into upper and lower half of the detector. It was found that none of the five reconstructed parameters deviates more than the simulated data from a consistent alignment and that the presented procedure is able to produce a consistent detector geometry without any prior alignment knowledge. This initial geometry knowledge enables the Inner Detector to quickly be aligned to the needed accuracy once collision data is available, so that early physics analyses involving particle tracks can be performed.

Appendix A

Additional figures

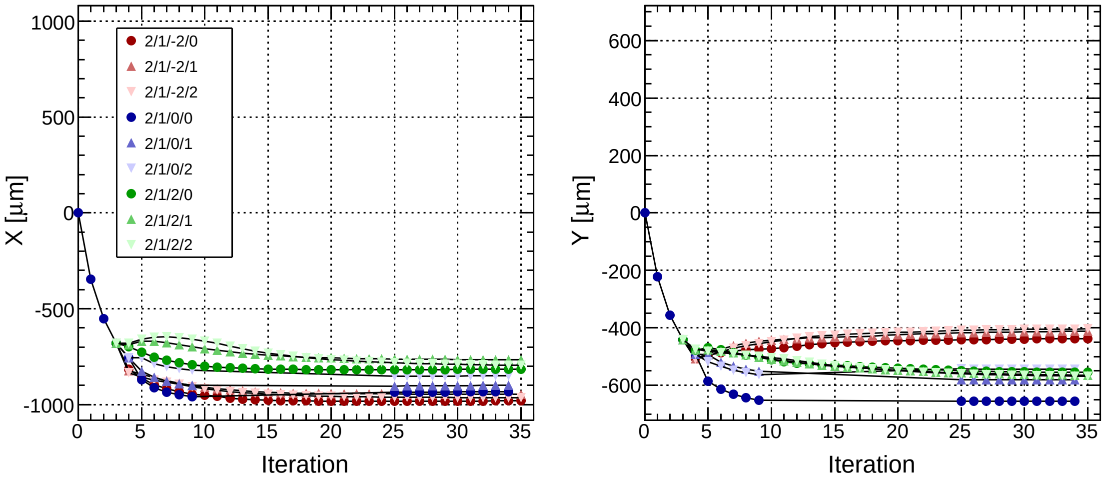


Figure A.1: *Global X- and Y-positions of the Pixel strucures during 35 iterations.*

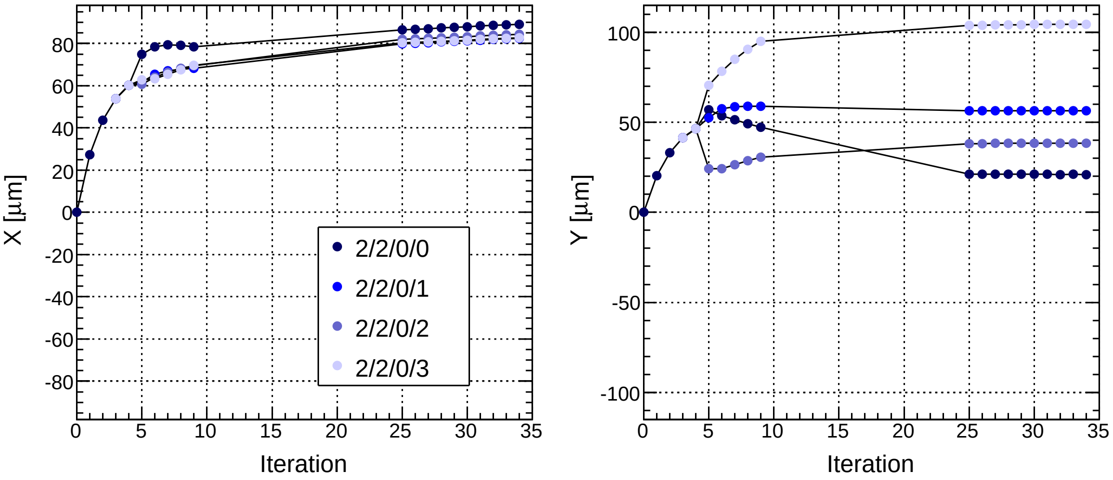


Figure A.2: *Same as figure A.1, but for the SCT barrel layers.*

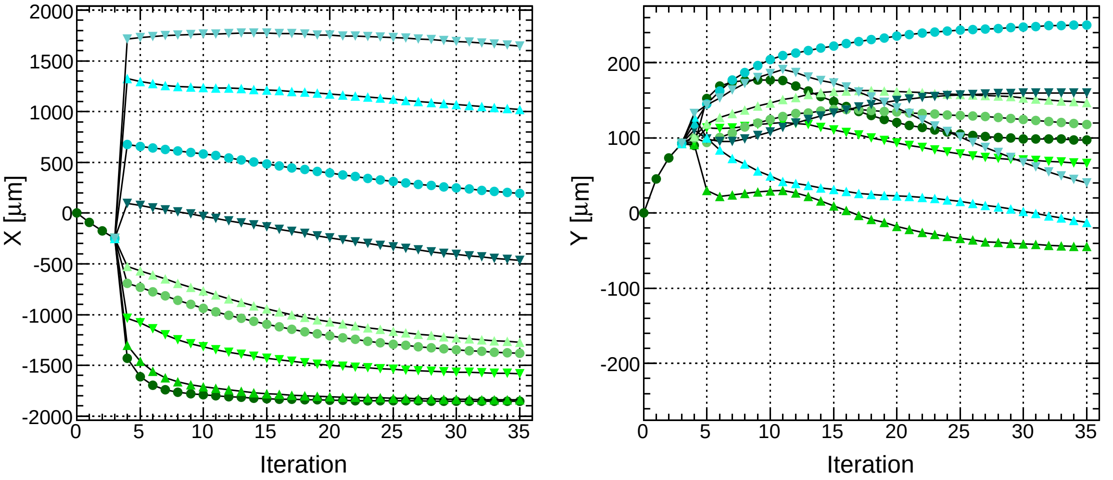


Figure A.3: *Same as figure A.1, but for the SCT endcap A disks. The legend is the same as for figure 5.22.*

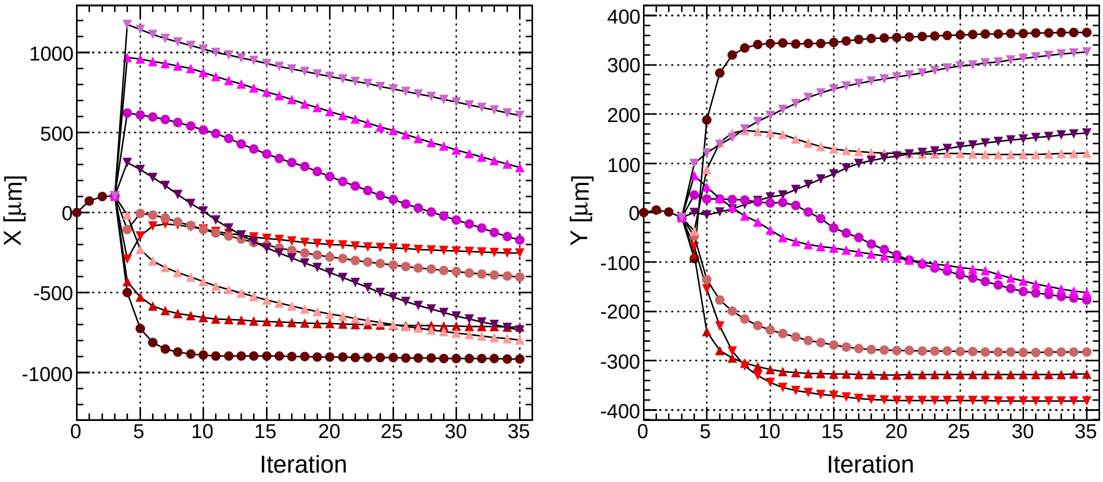


Figure A.4: *Same as figure A.1, but for the SCT endcap C disks. The legend is the same as for figure 5.14.*

Appendix B

List of Abbreviations

ALICE	A Large Ion Collider Experiment
AOD	Analysis Object Data file
ATLAS	A Toroidal LHC ApparatuS
CDF	Collider Detector at Fermilab
CERN	European Organisation for Nuclear Research (Conseil Européenne pour la Recherche Nucléaire)
CKM	Cabibbo-Kobayashi-Maskawa (matrix)
CL	Confidence Level
CMS	Compact Muon Solenoid
CSC	Cathode Strip Chambers
CTB	Combined TestBeam
DOCA	Distance Of Closest Approach
DOF	Degree Of Freedom
EC	End-Cap
EMEC	ElectroMagnetic End-cap Calorimeter
ESD	Event Summary Data file
FCal	Forward Calorimeter
FE	Front-End electronics
HEC	Hadronic Endcap Calorimeter
ID	Inner Detector
L1	Level-1 trigger / Level-1 alignment
L2	Level-2 trigger / Level-2 alignment
L3	Level-3 alignment
L32	Level-32 alignment
LAr	Liquid Argon
LHC	Large Hadron Collider
LO	Leading Order
MDT	Monitored Drift Tubes
MPP	Max Planck institute for Physics
NDF	Number of Degrees of Freedom
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
PDF	Parton Distribution Function
PMNS	Pontecorvo-Maki-Nakagawa-Sakata (matrix)
QCD	Quantum ChromoDynamics
QED	Quantum ElectroDynamics
ROC	Receiver Output Characteristics
ROOT	Data analysis framework
RPC	Resistive Plate Chambers
SCT	SemiConductor Tracker
SM	Standard Model of particle physics
TDR	Technical Design Report
TGC	Thin Gap Chambers
TMVA	Toolkit for MultiVariate Analysis
TRT	Transition Radiation Tracker

List of Figures

2.1 Top-antitop pair production processes 6
2.2 Single top production processes 7

3.1 LHC ring 10
3.2 Sketch of the ATLAS detector 12
3.3 Overview of the Inner Detector 13
3.4 Overview of the calorimeters 17
3.5 Overview of the muon system 18

4.1 Reconstructed m_t 28
4.2 Mean of d_{Merge} values versus lepton flavour 29
4.3 The d_{Merge} values at various clustering steps 31
4.3 . . . *continued.* 32
4.4 Rejection vs. efficiency for cutting on different d_{Merge} variables 32
4.5 Reconstructed m_t with requirement on d_{Merge} 33
4.6 Change of m_t when hardening d_{Merge} requirement 35
4.7 The d_{Merge} input variables for TMVA 37
4.7 . . . *continued.* 38
4.8 Fisher discriminator output 38
4.9 ROC curve for Fisher method and several d_{Merge} variables 39
4.10 Reconstructed m_t with requirement on $F(\boldsymbol{d})$ 40
4.11 Change of m_t when hardening requirement on $F(\boldsymbol{d})$ 41
4.12 Efficiency and purity curves 42
4.13 Selection of d_{Merge} variables for combinatorial background 43
4.14 Selection of d_{Merge} variables for total background 44
4.15 Kinematical variables separating signal and background 47

4.16 Discriminator output for background rejection based on kinematic variables 48
4.17 Reconstructed m_t with requirement on $F(\boldsymbol{k})$ 49
4.18 Change of reconstructed m_t when hardening requirement on $F(\boldsymbol{k})$ 50
4.19 Discriminator output for background rejection including combinatorial test events 51
4.20 ROC curves for background rejection based on kinematic variables 52
4.21 Reconstructed m_t with requirement on $F(\boldsymbol{k})$ 52
4.22 Change of reconstructed m_t when hardening requirement on $F(\boldsymbol{k})$ 53
4.23 Efficiency and purity curves for the Fisher method on kinematic variables . 54

5.1 Sketch of the four systematic deformations 62
5.2 Module movements for the systematic deformations 64
5.3 Z mass resolution with Curl misalignment 66
5.4 Cosmic ray track parameters 69
5.5 Pixel alignment parameter evolution without soft mode term 71
5.6 SCT barrel alignment parameter evolution without soft mode term 72
5.7 SCT ECA alignment parameter evolution without soft mode term 73
5.8 SCT ECC alignment parameter evolution without soft mode term 74
5.9 SCT ECA X- and Y-position without soft mode term 75
5.10 SCT ECC X- and Y-position without soft mode term 75
5.11 Pixel alignment parameter evolution with soft mode term 79
5.12 SCT barrel alignment parameter evolution with soft mode term 80
5.13 SCT ECA alignment parameter evolution with soft mode term 81
5.14 SCT ECC alignment parameter evolution with soft mode term 82
5.15 The x-residuals after L2 alignment 83
5.16 The y-residuals after L2 alignment 83
5.17 Alignment correction uncertainties in the Pixel barrel without survey constraint 90
5.18 Alignment correction uncertainties in the SCT barrel without survey constraint 91
5.19 Local y alignment corrections for the SCT barrel modules 92
5.20 Final Pixel alignment parameter evolution 94
5.21 Final SCT barrel alignment parameter evolution 95
5.22 Final SCT ECA alignment parameter evolution 96

5.23 Final SCT ECC alignment parameter evolution 97
5.24 The x-residuals after final alignment 98
5.25 The y-residuals after final alignment 99
5.26 Track parameter differences from split tracks 102

A.1 Pixel structure X- and Y-positions during 35 iterations 109
A.2 SCT barrel X- and Y-positions during 35 iterations 110
A.3 SCT ECA X- and Y-positions during 35 iterations 110
A.4 SCT ECC X- and Y-position during 30 iterations 111

List of Tables

2.1 Top Quark production cross sections 7
2.2 Expected branching fractions for decaying top quark pairs 8

3.1 Precisions of the Inner Detector modules 15

4.1 Signal and background samples used 26
4.2 Expected numbers of events for a luminosity of $146\,\mathrm{pb}^{-1}$ 27
4.3 Area under the ROC curve for different d_{Merge} variables 33
4.4 Linear correlation coefficients for the selected kinematic variables 46

5.1 Uncertainties of the alignment corrections for L2 structures 78
5.2 Soft mode constraint uncertainties 83
5.3 Pixel residual means and widths after L2 alignment 84
5.4 SCT residual means and widths after L2 alignment 85
5.5 Pixel residual means and widths from simulation 86
5.6 SCT residual means and widths from simulation 87
5.7 Pseudo survey constraint uncertainties 89
5.8 Pixel residual means and widths after 35 iterations 99
5.9 SCT residual means and widths after 35 iterations 100
5.10 Track parameter difference fits 101

Bibliography

[1] G. Altarelli, *The standard model of particle physics*, (2005), CERN-PH-TH-2005-206.

[2] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **B47** (1973) 365.

[3] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343.

[4] S. Bethke, *The 2009 World Average of $\alpha_s(M_Z)$*, (2009), MPP-2009-149, hep-ph/0908.1135v2.

[5] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[6] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[7] A. Salam, Elementary particle theory, p. 367, Almqvist and Wiksell, Stockholm, 1968.

[8] S. A. Goldstone, J. and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965.

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[10] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[12] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP **7** (1958) 172.

[13] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870.

[14] J. R. Incandela et al., *Status and Prospects of Top-Quark Physics*, Prog. Part. Nucl. Phys. **63** (2009) 239.

[15] F. Abe et al. (CDF Collab.), *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626.

[16] S. Abachi et al., *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632.

[17] (Tevatron Electroweak Working Group Collab.), *Combination of CDF and D0 Results on the Mass of the Top Quark*, (2009), FERMILAB-TM-2427-E, 0903.2503v1.

[18] C. Amsler et al., *Review of Particle Physics*, Phys. Lett. **B667** (2008) 1 .

[19] V. M. Abazov et al. (D0 Collab.), *Experimental discrimination between charge 2 e/3 top quark and charge 4 e/3 exotic quark production scenarios*, Phys. Rev. Lett. **98** (2007) 041801.

[20] S. Moch and P. Uwer, *Heavy-quark pair production at two loops in QCD*, Nucl. Phys. Proc. Suppl. **183** (2008) 75.

[21] J. Campbell and F. Tramontano, *Next-to-leading order corrections to Wt production and decay*, Nucl. Phys. **B726** (2005) 109 .

[22] Z. Sullivan, *Understanding single-top-quark production and jets at hadron colliders*, Phys. Rev. **D70** (2004) 114012.

[23] T. Aaltonen et al. (CDF Collab.), *First Observation of Electroweak Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092002.

[24] V. M. Abazov et al. (D0 Collab.), *Observation of Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009) 092001.

[25] L. Evans and P. Bryant (eds.), *LHC Machine*, JINST **3** (2008) S08001.

[26] K. Aamodt et al. (ALICE Collab.), *The ALICE experiment at the CERN LHC*, JINST **0803** (2008) S08002.

[27] R. Adolphi et al. (CMS Collab.), *The CMS experiment at the CERN LHC*, JINST **0803** (2008) S08004.

[28] A. A. Alves et al. (LHCb Collab.), *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[29] G. Aad et al. (ATLAS Collab.), *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[30] M. Bajko et al., *Report of the Task Force on the Incident of 19th September 2008 at the LHC*, CERN, Geneva, (2009), CERN-LHC-PROJECT-Report-1168.

[31] *ATLAS public webpage*, `http://atlas.ch/`.

[32] (ATLAS Collab.), *ATLAS Detector and physics performance: Technical design report*, CERN-LHCC-99-14.

[33] T. Golling, *Alignment of the Silicon Tracking Detector using Survey Constraints*, CERN, Geneva, (2006), ATL-INDET-PUB-2006-001. ATL-COM-INDET-2006-002. CERN-ATL-INDET-PUB-2006-001.

[34] D. Froidevaux, *private communication.*

[35] (ATLAS Collab.), *ATLAS central solenoid: Technical Design Report*, Technical Design Report ATLAS, CERN, Geneva, (1997), CERN-LHCC-97-021.

[36] J. P. Badiou et al. (ATLAS, Collab.), *ATLAS barrel toroid: Technical Design Report*, Technical Design Report ATLAS, CERN, Geneva, (1997), CERN-LHCC-97-019.

[37] (ATLAS Collab.), *ATLAS end-cap toroids: Technical Design Report*, Technical Design Report ATLAS, CERN, Geneva, (1997), CERN-LHCC-97-020.

[38] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, CERN, Geneva, (2008), ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003.

[39] M. Aharrouche et al., *Response uniformity of the ATLAS liquid argon electromagnetic calorimeter*, Nucl. Instrum. Meth. A **582** (2007) 429 .

[40] C. Issever, K. Borras, and D. Wegener, *An improved weighting algorithm to achieve software compensation in a fine grained LAr calorimeter*, Nucl. Instrum. Meth. A **545** (2005) 803 .

[41] T. Barillari et al., *Local Hadronic Calibration*, CERN, Geneva, (2008), ATL-LARG-PUB-2009-001. ATL-COM-LARG-2008-006.

[42] G. C. Blazey et al., *Run II jet physics*, (2000), FERMILAB-CONF-00-092-E, hep-ex/0005012v2.

[43] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, JHEP **05** (2007) 086.

[44] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993) 3160.

[45] S. Catani et al., *Longitudinally-invariant k[perpendicular]-clustering algorithms for hadron-hadron collisions*, Nucl. Phys. **B406** (1993) 187 .

[46] S. Catani, Y. Dokshitzer, and B. Webber, *The k[perpendicular]-clustering algorithm for jets in deep inelastic scattering and hadron collisions*, Phys. Lett. **B285** (1992) 291 .

[47] W. Bartel et al. (JADE Collab.), *Experimental studies on multijet production in e^+e^- annihilation at PETRA energies*, Zeitschrift für Physik C **33** (1986) 23.

[48] S. Catani et al., *New clustering algorithm for multijet cross sections in e^+e^- annihilation*, Phys. Lett. **B269** (1991) 432 .

[49] S. V. Chekanov, *Jet algorithms: A mini review*, (2002), Proc. of the 14th Topical Conference on Hadron Collider Physics, ANL-HEP-CP-02-103, hep-ph/0211298.

[50] M. Cacciari and G. P. Salam, *Dispelling the N**3 myth for the k(t) jet-finder*, Phys. Lett. **B641** (2006) 57.

[51] S. Kluth, *Tests of quantum chromo dynamics at e+ e- colliders*, Rept. Prog. Phys. **69** (2006) 1771.

[52] G. Aad et al. (ATLAS, Collab.), *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics*, CERN, Geneva, (2009), CERN-OPEN-2008-020.

[53] R. Härtel, *Studies on an initial top quark mass measurement at ATLAS in the lepton+jets $t\bar{t}$ decay channel and alignment of the Pixel and SCT subdetectors*, PhD thesis, TU München, (2009), MPP-2009-134.

[54] S. Hassani et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl. Instrum. Meth. A **572** (2007) 77 , Frontier Detectors for Frontier Physics - Proceedings of the 10th Pisa Meeting on Advanced Detectors.

[55] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029.

[56] G. Corcella et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.

[57] G. Corcella et al., *HERWIG 6.5 release note*, (2002), CERN-TH/2002-270, hep-ph/0210213v2.

[58] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. **D78** (2008) 034003.

[59] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004.

[60] S. Agostinelli et al., *G4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[61] J. Allison et al., *Geant4 developments and applications*, IEEE Transactions on Nuclear Science **53** (2006) 270.

[62] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.

[63] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)*, Phys. Rev. D **74** (2006) 114017.

[64] M. L. Mangano et al., *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP **01** (2007) 013.

[65] *The Athena framework*, `https://twiki.cern.ch/twiki/bin/view/Atlas/AthenaFramework`.

[66] *MPP top mass reconstruction software repository*, `http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/groups/MPP/ARA_Examples_Top/`.

[67] M. Abramowitz and I. Stegun (eds.), *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, Washington, 10th edition, (1972).

[68] E. W. Weisstein, *Chebyshev Polynomial of the First Kind.*, `http://mathworld.wolfram.com/ChebyshevPolynomialoftheFirstKind.html`, From MathWorld–A Wolfram Web Resource.

[69] R. Wanke, *Special practical problems in data analysis - The art of determining systematic errors*, `https://indico.desy.de/materialDisplay.py?contribId=12&materialId=slides&confId=1149`, talk at Helmholtz Terascale Statistics Tools School.

[70] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals Eugen. **7** (1936) 179.

[71] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, POSACAT **040** (2007), software available from `http://tmva.sourceforge.net/`.

[72] H. Voss, *Multivariate Analysis Techniques*, `https://indico.desy.de/materialDisplay.py?contribId=1&materialId=slides&confId=1149`, talk at Helmholtz Terascale Statistics Tools School.

[73] M. H. Seymour and C. Tevlin, *A comparison of two different jet algorithms for the top mass reconstruction at the LHC*, JHEP **11** (2006) 052.

[74] R. Brun et al., *ROOT Users Guide 5.16*, (2007), software available from `http://root.cern.ch/`.

[75] L. Sachs and J. Hedderich, *Angewandte Statistik*, Springer, Berlin, 12th edition, (2006).

[76] G. Cortiana et al., *Template Method for an early Top-Quark Mass Measurement in the* $t\bar{t} \rightarrow$ *lepton+jets Channel with ATLAS Data*, ATL-COM-PHYS-2009-603 (in preparation).

[77] P. Weigell, Constrained kinematic fitting for a top quark mass determination in the electron + jets channel at atlas, Diploma thesis, TU München, (2009), MPP-2009-177.

[78] R. Härtel, Iterative local χ^2 alignment approach for the ATLAS SCT detector, Diploma thesis, TU München, (2005), MPP-2005-174.

[79] T. Göttfert, Iterative local χ^2 alignment algorithm for the ATLAS Pixel detector, Diploma thesis, Universität Würzburg, (2006), MPP-2006-118.

[80] S. M. Gibson, *The ATLAS SCT alignment system and a comparitive study of misalignment at CDF and ATLAS*, PhD thesis, University of Oxford, (2004).

[81] P. Brückman, A. Hicheur, and S. Haywood, *Global chi2 approach to the Alignment of the ATLAS Silicon Tracking Detectors*, CERN, Geneva, (2005), CERN-ATL-INDET-PUB-2005-002.

[82] A. Bocci and W. Hulsbergen, *TRT Alignment For SR1 Cosmics and Beyond*, CERN, Geneva, (2007), CERN-ATL-INDET-PUB-2007-009.

[83] T. Cornelissen, *Track Fitting in the ATLAS Experiment*, PhD thesis, Amsterdam Univ., (2006), CERN-THESIS-2006-072.

[84] D. Brown et al., *Internal alignment of the BABAR silicon vertex tracking detector*, (2006), Proceedings of the 1st LHC Detector Alignment Workshop.

[85] J. Alison, B. Cooper, and T. Göttfert, *Production of Residual Systematically Misaligned Geometries for the ATLAS Inner Detector*, CERN, Geneva, (2009), ATL-INDET-INT-2009-003.

[86] (ATLAS Collab.), *Studies into the Impact of Inner Detector Misalignments on Physics*, CERN, Geneva, (2009), ATL-PHYS-PUB-2009-080.

[87] *Approved plots wikipage*, `https://twiki.cern.ch/twiki/bin/view/Atlas/ApprovedPlotsID#Impact_of_Global_Systematic_Misa`.

[88] S. Pataraia, *Studies of $t\bar{t}$ production in pp collisions at the Large Hadron Collider with the ATLAS experiment*, PhD thesis, TU München, (2009), MPP-2009-181.

[89] *`Chi2AlignAlg` in the Athena software repository*, `https://svnweb.cern.ch/trac/atlasoff/browser/InnerDetector/InDetAlignAlgs/InDetLocalChi2AlignAlgs/trunk`.

[90] C. Guyot et al., *Data Quality Status Flags and Good Run Lists for Physics Analysis in ATLAS*, CERN, Geneva, (2009), ATL-COM-GEN-2009-015.

[91] O. Brandt and P. Bruckman de Rentstrom, *Hit Quality Selection for Track-Based Alignment with the InDetAlignHitQualSelTool in M8+*, CERN, Geneva, (2009), ATL-COM-INDET-2009-015.

[92] P. F. Åkesson et al., *ATLAS Tracking Event Data Model*, CERN, Geneva, (2006), CERN-ATL-SOFT-PUB-2006-004.

[93] T. G. Cornelissen et al., *Updates of the ATLAS Tracking Event Data Model (Release 13)*, CERN, Geneva, (2007), ATL-SOFT-PUB-2007-003.

[94] J. Schieck, *ATLAS Inner Detector Alignment with Cosmics*, `http://indico.cern.ch/getFile.py/access?contribId=9&sessionId=1&resId=0&materialId=slides&confId=50502`, talk at 3rd LHC detector alignment workshop.

Acknowledgments

The final words in this thesis belong to those who supported me during the last years and who went with me on this scientific endeavour.

First I want to thank Prof. Bethke for giving me the opportunity to work in his group and conduct the thesiswork at a marvellous experiment. Just as well, I thank my supervisor Richard Nisius, whose attentive reading and to-the-point questions bring the scientific discussion to a new level. Thank you for the scientific input to and careful scrutiny of my work.

A big thank-you goes to the senior group members Stefan Kluth and Jochen Schieck, who were always willing to help and give input. Thank you Jochen also for your great job as convenor of the alignment group, whose extensive work now pays off with the first ATLAS data.

What makes the time spent at the institute worthwhile are of course the many colleagues and office mates that are all truly unique people. Thank you for all help and discussions, insights into foreign languages and cultures, funny evenings and remarkable get-togethers, and fruitful collaboration. Thanks to Roland, Giorgio, Anna, Philipp, Petra, Nabil, Sven, Emanuel, Andreas, Sophio, Michael, Silke, Martin for providing scientific input, having always a technical tip at hand or lending an ear to my relevant and less relevant problems. I especially wish a big success to Giorgio Cortiana with his new duty as convenor of the top mass subgroup. And of course thanks to everybody else at the MPP that I learned to love as a great place to work during the last four years.

A warm thank-you is due for all people at the ATLAS collaboration, who make the incredible happen. I owe a special thanks to the alignment group, but also every other member of the ATLAS community. Keep up the good spirit, which made the arrival of real data at ATLAS a big success! I really enjoyed being part of the collaboration.

Finally, my gratitude is with my whole family. Most important, I have to mention Kathi; without you I could have never kept up. Thank you for all your love, support and patience during this time. I would like to dedicate this work to my father. I know he would have liked to be with us for a while longer. I will never forget all you have done for me.