

Bergische Universität Wuppertal

A STUDY OF THE LUMINOSITY ZONE OF THE ATLAS DETECTOR AT THE LHC

BACHELOR-THESIS
IN DER
EXPERIMENTELLEN TEILCHENPHYSIK

Studiengang: Bachelor of applied science
Erstgutachter: Prof. Dr. W. Wagner
Zweitgutachter: Priv.-Doz. Dr. D. Wicke
Eingereicht am: 19. Januar 2012
Vorgelegt von: Michael Schuh
Matrikelnummer: 510944

Contents

Table of contents	I
1. Introduction	1
1.1. Motivation	1
1.2. Problem description	2
1.3. Outline of thesis	2
2. Physics background and experimental setup	3
2.1. Particle physics	3
2.1.1. Fermions, the matter particles	3
2.1.2. Bosons, the force-carrier particles	5
2.1.3. Questions about the Standard Model	6
2.2. LHC	7
2.2.1. RF and bunch structure	8
2.2.2. Transverse dynamics and focusing	9
2.3. ATLAS	13
2.3.1. Luminosity	15
2.3.2. Vertex reconstruction	17
2.3.3. Vertex distribution	19
2.3.4. Vertex resolution	20
3. Run selection	20
4. Observed beamspot and beamwidth for primary vertices	23
4.1. Distribution in the ATLAS coordinate system	23
4.2. Tilt in the ATLAS coordinate system	27
4.3. Rotation of the coordinate system	29
4.4. Vertex distribution in the rotated coordinate system	32
5. Vertex resolution	34
5.1. Dimension of the vertex resolution	35
5.1.1. Single run evaluation	35
5.1.2. Evaluation for the set of investigated runs	40
5.2. Spatial distribution of the vertex resolution	42
5.3. Vertex resolution as a function of the number of tracks	46
6. Vertex resolution correction	50
6.1. Avoiding low resolution	50
6.2. The mean estimation	52

6.3. Deconvolution	53
6.3.1. Inverse filter	54
6.3.2. Wiener filter	55
7. Distribution of the pile-up events	57
8. Conclusions	60
8.1. Summary	60
8.2. Discussion	60
8.3. Future work	61
A. Appendix	63
A.1. Collections of plots for all 23 runs	63
References	XII
List of figures	XIV
List of tables	XVI
Acknowledgements	XVII
Erklärung	XVIII

1. Introduction

The LHC - Large Hadron Collider - is the world's largest particle accelerator. It is located near Geneva, Switzerland and operated by CERN, the European Organisation for Nuclear Research. Built to produce proton-proton and heavy-ion collisions at unreached levels on the energy scale and unreached event rates, the twin beam collider has run at half of its design energy of 14 TeV since 2010. Colliding particles are destroyed in the experiment, but the energy is conserved and forms new elementary particles. Four detectors, ATLAS, CMS, ALICE and LHCb, are installed along the accelerator tunnel to identify the collision products and provide insight into the large number of processes involved. They collect information about the type, pathway, energy and momentum of the collision products and gain a huge amount of data, analysed by thousands of physicists all over the world. The Department for Experimental Particle Physics at the Bergische Universität Wuppertal is involved in the ATLAS experiment, which contributes to the development of its inner detector as well as working on the analysis of the data taken by the experiment.

In 2011 the amount of data collected by the four detectors at LHC has been well increased to more than 12.4 fb^{-1} (see section 2.3.1). This has generated a lot of public interest on LHC and ATLAS as CERN has announced to continue this work throughout 2012 and expressed expectations to answer the question of the existence or non-existence of the Higgs boson by the end of 2012 [1]. The Higgs boson is the only fundamental particle, predicted by the theory of the Standard Model, that has not yet been observed by experimental physics. A proof of its existence is regarded to be an essential step in the progression of modern physics.

In order to produce the huge amount of recorded events, that are required for the statistics in the search for the Higgs boson, the transverse size of the interaction region, in which the two beamlines cross inside the ATLAS detector, has to be made as small as possible. Knowing its size and position is of great importance, because the coordinates of the average interaction point are used as parameters for the tracking process in real time during ATLAS runtime. Thus it is computed "online" and updated several times per run. The recorded data allows additional offline analysis of the structure and dynamics of the interaction region. This is what is examined in this thesis with a focus on the spatial structure and the measurement precision of the recorded data.

1.1. Motivation

Measurements of the three-dimensional distribution of the interaction vertices allow for conclusions about the transverse and longitudinal emittance and the luminosity of the LHC operations. In 2011 the successful operations of LHC and ATLAS led to an outstanding increase in luminosity production. During the year, the number of bunches per fill has been raised, the beta-function at the interaction point has been levelled down to $\beta^* = 1 \text{ m}$ and the emittance of the injected beams has improved. The evaluations for single runs can be drawn together and analysed with a view to the changes of the beam parameters. With higher luminosity, LHC does not only produce

more events classified as primary events (the “events of interest”) but also generally a higher amount of interactions per bunch crossing. This includes more events classified as pileup events, that are of secondary importance regarding the aims of the experiment. Since the number of interactions per bunch crossing is expected to rise significantly with the beam intensity, it is of interest to examine the impact of the pileup events for the analysis of the luminosity region. Knowing the spatial distribution of the precision of the data acquisition, the confidence intervals of the interaction vertices, is an opportunity to estimate a correction for the statistical results of the spatial vertex distribution. In order to follow this approach the underlying structures of the measurement precision have to be understood as well.

Due to the stable performance of the LHC, there are numerous data sets for runs with stable beams, lasting 10 hours and longer, which gives a good background to perform an offline analysis of the luminosity region.

1.2. Problem description

The spatial distribution of the interaction region in the ATLAS detector at the LHC is the main subject of this thesis. This includes the transverse and longitudinal size of the interaction region, the position of the average interaction point and the orientation of the beamlines in the detector’s coordinate system. The distribution of the measurement precision is analysed with a view on the absolute values, spatial dependencies and its relation to the track multiplicity. The known precision of the data acquisition is used to approach a correction of the given results. The findings for the single runs are compared and presented in timelines to show the evolution over time period between April 2011 and the end of June 2011.

The huge amount of data made it a challenge to automate the evaluations and to focus on the efficiency of the used algorithms. The raw compressed data for the analysed runs, organised in ROOT Trees [2], includes 563 GB. Even after extracting the required subsets of the original datasets and reducing the information content per event to the level needed for the analysis, 25 GB of data still remained to be analysed, looping over it run for run. The data organisation and visualisation was realised with ROOT, using its python interface PyROOT and python 2. Computations were made using mathematical algorithms of the python extensions SciPy and NumPy [3, 4].

1.3. Outline of thesis

A view on the physical background and technical aspects of the machines LHC and ATLAS is given in section 2. Methods of beam acceleration and focusing are detailed in section 2.2 and the emittance as the conserved parameter to determine the beam quality is defined. Section 2.3 focuses on ATLAS and on the interaction region. The detector layout is delineated and the luminosity as the key quantity to determine the productivity of the experiment is defined and discussed. In the analysis part the spatial distribution of the observed events is shown and the

beamspot position and width of the luminosity region are indicated by the visualisation of the datasets. The first analysis focuses on one-dimensional projections and neglects correlations between spatial directions. This is extended to an examination of the relations between transverse and longitudinal coordinate distributions in section 4.3, leading to an adjustment of the coordinate system to the observed beamline.

In section 5 a detailed analysis of the structure of the measurement precision is given, first by observing the absolute dimension of the confidence intervals assigned to the vertices in section 5.1, then by looking at the spatial dependencies in section 5.2. The measurement precision is also analysed as a function of the track multiplicity. In section 5 methods of vertex resolution correction will be compared to take a more in-depth look at the observable data. The basic results will also be reviewed against the distribution of the pileup events in section 7 and differences of this distribution and the distribution of the primary vertices, observed in the analysed runs, are addressed.

2. Physics background and experimental setup

LHC and ATLAS are flagships of the latest in experimental particle physics. A short picture of their purpose and concept will be given in this section, before the presentation of the analysis. The equations which count for the luminosity region are summarised in section 2.3.3.

2.1. Particle physics

The periodic table of elements, introduced by Meyer and Mendelew in 1869, illustrates a pattern, which describes the atomic structure of all known matter. Atoms are understood to be a composition of electrons, protons and neutrons, but until now the electron is the only one of these building blocks which hasn't been found to contain a further substructure. Thus the electron is classified as an elementary particle, but neither a proton nor a neutron. The theoretical system used to describe the structures on a subatomic scale beyond the periodic table is the Standard Model. Being a quantum field theory, it distinguishes matter particles, the fermions, and force-carrier-particles, the bosons.

2.1.1. Fermions, the matter particles

The Standard Model explains the protons and neutrons as a composition of three quarks, presenting two different types, namely up and down. The up-quark u has a charge of $2/3$ and the down-quark d has a charge of $-1/3$, given in units of the elementary charge e . The proton contains two up-quarks and one down-quark, resulting in its charge of $+e$, the neutron contains one up-quark and two down-quarks. The up-quark and the electron are stable particles, but the down-quark can decay into an up-quark. This is the process taking place in the decay of free neutrons. Together with the electron neutrino ν_e , the up-quark, down-quark and the electron

form the so-called “first generation” of matter particles. The neutrino is the most puzzling of these four elementary particles, because even today its mass remains undetermined, known only to be less than 2 eV. It carries no charge and shows almost no interaction with other particles. Each second any squarecentimeter on the surface of the earth perpendicular to the sun is passed by more than $6 \cdot 10^{10}$ neutrinos, that originate from nuclear processes in the sun [5]. The neutrino is very different from the other three fermions of the first generation, since is not found basically bound to atoms.

The Standard Model is also sufficient to describe the second and third generation of fermions as well. It predicts two corresponding versions of each of the four elementary matter particles of the first generation, that have the same electric charge as their first generation particle, but higher masses in higher generations. These particles are unstable and have lifetimes of the smallest fraction of a second. They decay into particles with lower masses and at the end of the decay process the stable particles of the first generation remain. All of the 12 fermions have been observed in experiments with particle accelerators. The fermions and their properties are listed in table 1. Its subdivision into quarks and leptons is based on the forces, which interact between the fermions. This is explained in the next paragraph.

Table 1: Table of fermions. This part of the Standard Model describes the matter particles. All have spin $s = 1/2$ and follow the Fermi-Dirac statistics. The particle type is referred to as flavour f , the particle mass m is given in GeV/c^2 [6].

		1st generation		2nd generation		3rd generation	
	charge	f	m	f	m	f	m
Quarks	$2/3$	up	$(1.7 - 3.1) \cdot 10^{-3}$	charm	1.3	top	173
Quarks	$-1/3$	down	$(4.1 - 5.7) \cdot 10^{-3}$	strange	0.1	bottom	4.2
Leptons	1	e	$5, 11 \cdot 10^{-4}$	μ	0.11	τ	1.78
Leptons	0	ν_e	$<2 \cdot 10^{-10}$	ν_μ	$<2 \cdot 10^{-10}$	ν_τ	$<2 \cdot 10^{-10}$

For each quark and each lepton, a corresponding antiparticle exists, which has the same mass, but the opposite charge. These antiparticles are not found as stable particles on the earth. They are observed in experiments with particle accelerators and in the secondary cosmic radiation. Altogether, there are 24 types of fermions, described by the Standard Model.

Besides the proton and neutron, other different combinations of two or three quarks have been observed, all resulting in effective charges of either 0 or ± 1 or ± 2 . They are collectively labelled under the term hadrons. Particles composed of three quarks or three antiquarks are labeled baryons. Particles that consist of one quark and one antiquark, where quark and antiquark can in general be of different flavour, for example $u\bar{c}$, are called mesons.

2.1.2. Bosons, the force-carrier particles

Adding up the masses of the up and down quarks, combined in protons and neutrons, results in masses far lower than actually observed. For the proton, the sum of the quark masses is less than 1% of its mass of 0.938 GeV. The explanation for this is the presence of the binding energy which keeps the three quarks bound to each other. According to the theory of relativity, masses can be assigned to these quantized energies, given by $E = mc^2 \iff m = E/c^2$. The Standard Model explains the bindings by the exchange of virtual force-carrier particles that are listed in table 2. The so-called strong force interacts between quarks. That is the background for the subdivision of the fermions into leptons and quarks, as leptons remain unaffected by this force. The strong force also interacts between its exchange particles, which is different to the other forces. The charge that applies to the strong interaction is described in terms of colours. Because there are three different types of colour charge present, the terminology of plus and minus used to describe the electrical charges is not sufficient. The used expressions are red, blue and green, but these colours do not correspond to the visible colors in any way.

The theory which explains the strong interaction in terms of colour is the quantum chromodynamics QCD. The virtual force-carriers that are described by QCD are the gluons. There are eight types of gluons. They have a very short interaction range and only act inside the hadrons and between them. They keep quarks bound inside hadrons and hadrons bound inside the nuclei.

Table 2: Table of bosons. All have spin $s = 1$, they follow the Einstein-Bose statistics.

bosons	notation	charge (weak-isospin)	mass in GeV/c^2	force	
photon	γ	0	0	electromagnetic	electroweak
W bosons	W^-	-1	80.4	weak	electroweak
	W^+	+1	80.4	weak	electroweak
Z boson	Z^0	0	91.188	weak	electroweak
8 gluons	g	color charged	0	strong	strong

In comparison to the strong force, the weak force affects all fermions, quarks as well as leptons, but none of the bosons. It has been unified with the electromagnetic force, described by one theory, the electroweak force. The exchange particles of the weak interaction are the W^+ , W^- and Z^0 bosons that carry a weak-isospin \vec{T} . In the process of quark flavour changing, a W^- boson is emitted, when a down-type quark changes to an up-type quark and a W^+ boson is emitted, when an up-type quark changes its flavour to a down-type. This way, the weak interaction explains the beta decay as illustrated by the Feynman graph in figure 1.

In the table of bosons, there is no Higgs boson listed, because from the present point of view it is a hypothetical particle as it has not been observed in any experiments. The Higgs boson is proposed to have spin $s = 0$ and to interact with all elementary particles that have a mass $m > 0$. It is proposed to be the theoretical explanation for exactly this property. The Standard Model includes the Higgs boson, but there are ‘‘Higgsless theories’’ which favour different arguments.

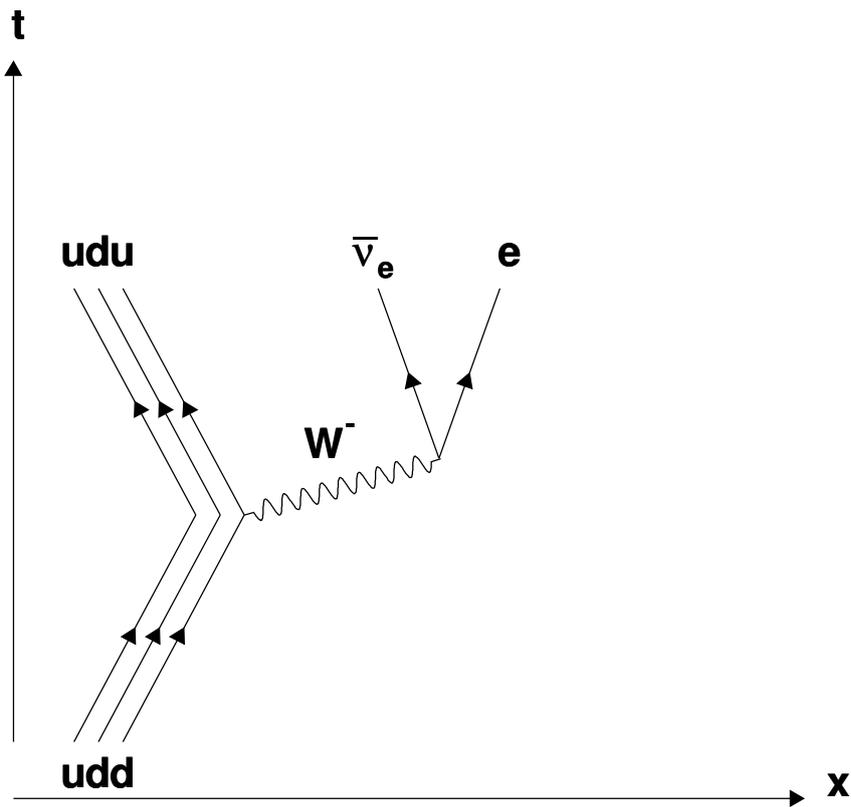


Figure 1: The W^- boson as exchange particle in the beta decay $n \rightarrow p + e^- + \bar{\nu}_e$. One of the down quarks in the neutron changes its flavour to up type, emitting a W^- boson. The boson decays into an electron and an electron antineutrino.

2.1.3. Questions about the Standard Model

The Standard Model is sufficient to describe the elementary particles that have been observed in experiments with particle accelerators. But there is still a set of open questions and a lot of reasons for further research.

- The unifying of the electromagnetic and weak force to the theory of the electroweak force calls for an unified theory that includes the strong force as well. Can the Standard Model be extended that way?
- The Standard Model does not include the gravitation in the particular force-exchange processes. On an atomic and subatomic scale it can be neglected in praxis, since it is many orders of magnitude weaker than the three other forces. But a complete theory should of course explain gravitation as well. What is the nature of the proposed force-carrier, the graviton?
- Why are there additional generations of particles and why three of them? Are there particles with even higher masses, yet unrevealed by experimental physics and theory?

- Will the existence of the Higgs boson be proved? Is the Higgs mechanism the reason for the mass of elementary particles? Can the numeric values for these masses be derived from theory?
- Astrophysics brought insight into the existence of dark matter. Can this matter be included into the Standard Model?
- How did matter become the dominant form of fermions and not antimatter?

2.2. LHC

To possibly answer the open questions about the Standard Model, an impressive effort is being made, to build experiments used for this research. In the search for the Higgs boson and for elementary particles with even higher energies, the rest energy required for their emergence has to be merged into a space small enough for the strong interaction to work, which is on a fm scale. The basic concept of the LHC accelerator is to let protons and heavy ions collide at high energies to meet exactly this condition. It is operated with proton fills in the majority of runs and the three detectors ATLAS, CMS and LHCb are designed with a focus on this operation mode. ALICE (A Large Ion Collider Experiment) focuses on the Pb^{+82} collisions to study the physics of the quark-gluon plasma created by these collisions.

The LHC has been developed to explore new ranges at the energy scale. Its design aims at a collision energy of 14 TeV. To prepare for this run mode, the collider will be shut down for a longer technical stop, after the current stage of commissioning at 7 TeV has been completed. This will not be the case before the end of 2012 [1]. The runs analysed in this thesis all have a collision energy of 7 TeV in the mass center system and are proton fills only. The protons in each beam have an energy of 3.5 TeV, what corresponds to a speed of $v = 0.99999996c$, just $4 \cdot 10^{-6}\%$ below the speed of light in vacuum.

The LHC is built in the tunnel of the former LEP electron-positron collider at CERN. The tunnel, illustrated by figure 2, has a length of 26.659 km and lies 50-175 m underneath the border of Switzerland and France. At 3.5 TeV the revolution frequency of an accelerated proton is 11.245 kHz. The beams are kept in two beam tubes and only cross inside the four experiments. The tubes have to provide an ultrahigh vacuum of $10^{-10} - 10^{-9}$ mbar [7]¹ to avoid losing protons which are destroyed in collisions with restgas particles.

¹All technical information provided in section 2.2 is based on the LHC Design Report Volume I and III. To avoid its too frequent usage, [7] will be added only, when the Design Report is cited literally.

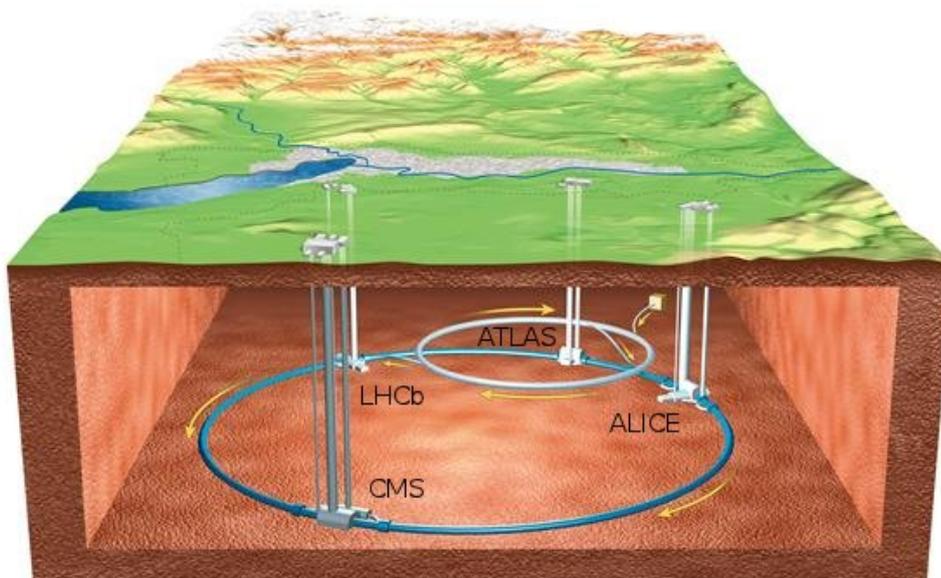


Figure 2: The Large Hadron Collider and the four detectors [8]. ATLAS is surrounded by the SPS accelerator. It has a circumference of 6.9km and is used as the injector for the LHC. The several other facilities at CERN are not shown in this image.

2.2.1. RF and bunch structure

A chain of four accelerators (Linac2 - Proton Synchrotron Booster (PSB) - Proton Synchrotron (PS) - Super Proton Synchrotron (SPS)) [7] accelerates the protons to 450 GeV, before they are injected into the LHC. A synchrotron radiofrequency (RF) of 400 MHz is used to accelerate, capture and store the injected particles. The RF has to be provided by two independent systems of cavities, because both particle beams have a positive charge. It is realized by a series of gaps and cavities, shielding the beams from the decelerating phase and exposing them to the accelerating phase.

Approximately 20 minutes after injection, the full design energy is achieved and the RF is used to store the beam and keep the particles collected into bunches. This is often compared to waveriding, since the position of the synchronous particle is at the slope of the RF waveform. Particles which arrive ahead of the synchronous particle receive a smaller amount of energy and become less accelerated; in turn particles which travel behind the synchronous particle receive a higher acceleration and catch up with the synchronous particle. The particles oscillate around the position of the synchronous particle. After acceleration the phase of the RF focuses the highly relativistic protons at the synchronous position at the falling edge of the RF. That way the RF keeps the particles in a bunch structure which is already predetermined before the injection, since SPS applies the same principle and injects bunches of about 1.2 ns in length. Because the RF has to be harmonized with the revolution frequency, the number of stable positions for synchronous particles, so-called buckets, depends on both frequencies. This is called the harmonic number, in

the LHC it is

$$f_{\text{harmonic}} = \frac{400\text{MHz}}{11.245\text{kHz}} \approx 36000 \quad (1)$$

For the adjustment of the bunch crossing frequency a minimum gap of 9 buckets is kept unfilled between used buckets, aiming at a bunch distance of 25 ns. In the 7 TeV runs examined in this thesis, the bunch spacing is 50 ns, since a maximum of every 20th bucket is filled. Counting the 3600 buckets which can be potentially filled following the 25 ns design, the bunches are assigned to unique bunchcrossing IDs. This provides the opportunity to conduct evaluations with respect to single bunch pairs. Figure 3 shows the bunch filling scheme of run 184169. The gaps of single bunch crossing IDs left out in this 50 ns run is not resolved at the used scale, but it is illustrated that an additional structure is formed, building so-called bunch trains, by leaving longer gaps between groups of consecutive bunches. The design aims at a maximum of 2808 bunch crossings per beam turn with $1.15 \cdot 10^{11}$ particles per bunch.

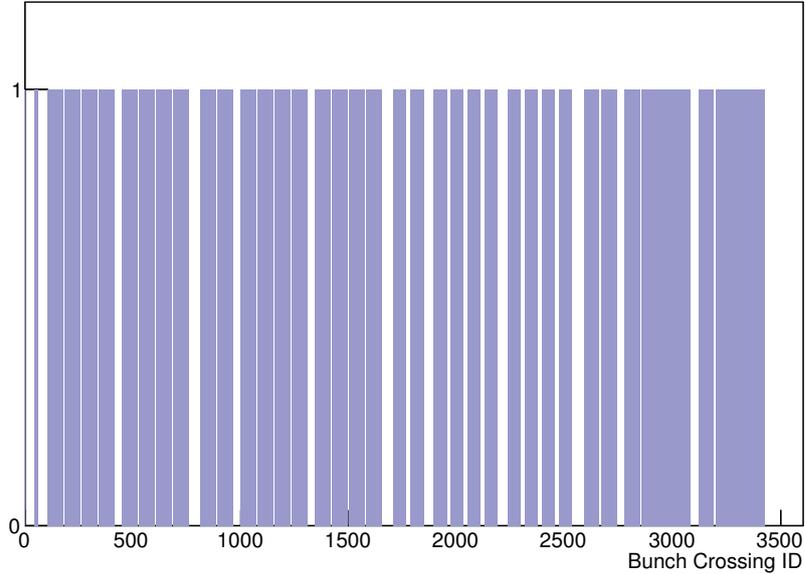


Figure 3: Bunch filling scheme of run 184169

2.2.2. Transverse dynamics and focusing

In contrast to the impression given in figure 2, the LHC is not a perfect cycle, but consists of eight arcs and eight straight sections. In the arcs, the beam is curved to the pathway by dipole magnets using the Lorentz force

$$\vec{F} = e(\vec{v} \times \vec{B}) \quad (2)$$

Because both beams have particles of positive charge, but an opposite direction of \vec{v} , they require their own magnetic fields with \vec{B} also pointing in opposite directions, too. The coordinate system, in which the directions are expressed, is moved with the beam. With the z-direction

pointing parallel to beam 2, the x-direction 90° horizontal towards the center of the collider ring and the y-direction 90° vertical, the coordinate system is right-handed and orthogonal. The magnetic field for beam 1, travelling clockwise, points in positive y-direction whereas the field for beam 2 points downwards. LHC applies the twin-bore technology to provide these fields in the same electro-magnets with a distance of only 194 mm between the two beams. The magnets provide fields of up to 8.33 T in 14 TeV runs which is achieved by using superconductivity at temperatures below 2 K. 1104 of those magnets are installed along the 8 arcs. The ensemble of magnets along the collider ring is referred to as lattice. 128 additional dipole magnets are installed along the ring at the transition points between arcs and straights. Their purpose is to suppress dispersion of the beam optics. Each of the dipole magnets is more than 14 m long and weighs about 35 tons.

Because the protons tend to diverge from the bunch due to the electromagnetic force between them, the beam has to be focused in all three spatial dimensions. In longitudinal direction, this is done by the RF, as described in section 2.2.1. In the transverse direction, this is done by quadrupole magnets leading to the same pattern, having the particles carrying out oscillations around the orbit of the synchronous particle. Their effect on the beam can be described by focusing and defocusing lenses. The transverse motion is analysed for the particle ensemble of the beam, by looking at the evolution of the envelope functions $E_x(s)$ and $E_y(s)$, where s refers to the position along the orbit. A quadrupole that focuses the beam in the x-z-plane acts as defocusing in the y-z-plane and vice versa. Because of this effect, the focusing system consists of a series of pairs of quadrupoles building a 90° -FODO-lattice. F refers to a quadrupole that focuses in the x-z-plane, D to a quadrupole that acts defocusing in that plane. O represents the straight drift sections between the quadrupoles. The oscillation of the single particle around the orbit of the synchronous particle is described by

$$x(s) = \sqrt{\varepsilon} \sqrt{\beta_x(s)} \cos(\Psi(s) + \phi) \quad (3)$$

where the amplitude is given by a constant $\sqrt{\varepsilon}$ which is given by the initial conditions of the beam in the product with a modifying factor $\beta(s)$ and changes with the quadrupole field strength along the orbit. The emittance ε is the key value in measuring the beam quality. The phase advance of the oscillation $\Psi(s)$ also depends on the quadrupole field strength, while ϕ is given by the beam conditions and acts as an initial constant. When the amplitude is large due to low focusing by the quadrupole field, $\beta(s)$ is large and the phase advance is low. In turn when the amplitude is low due to strong focusing, $\beta(s)$ is low and the phase advance is large. $\Psi(s)$ depends on the beta-function $\beta(s)$ as in equation 4.

$$\frac{d\Psi(s)}{ds} \propto \frac{1}{\beta(s)} \implies \Psi(s) = \int_0^s \frac{1}{\beta(\sigma)} d\sigma \quad (4)$$

A detailed discussion of equations 3 and 4 is given in [9]. Calculating the derivation of equation 3 gives the divergence x' .

$$x'(s) = -\sqrt{\frac{\varepsilon}{\beta(s)}} \sin(\Psi(s) + \phi) \quad (5)$$

Rearranging the equations for $x(s)$ and $x'(s)$, such that the trigonometric functions are isolated, and using that $\cos^2(\Psi(s) + \phi) + \sin^2(\Psi(s) + \phi) = 1$ gives equation 6, where $\alpha(s) = -\beta'(s)/2$.

$$\varepsilon = \frac{x^2}{\beta(s)} + \left(\frac{\alpha(s)}{\sqrt{\beta(s)}}x + \sqrt{\beta(s)}x' \right)^2 \quad (6)$$

Defining $\gamma(s) = (1+\alpha^2(s))/\beta(s)$, this equation can be modified, leading to the Courant-Snyder ellipse equation [10].

$$\varepsilon = \gamma(s)x^2(s) + 2\alpha(s)x(s)x'(s) + \beta(s)x'^2(s) \quad (7)$$

The motion of each particle describes an ellipse in phase-space, lying within the ellipse described by the envelope functions that has the area

$$A = \pi\varepsilon \quad (8)$$

The used coordinates $x(s)$ and $x'(s)$ comply with the Hamiltonian canonical coordinates p and q . The area of this ellipse and thus ε is conserved, according to the Liouville Theorem, that states that for the canonical coordinates we have

$$\int p dq = const \quad (9)$$

While the orientation of the ellipse in the phase-space is altered by the beta-functions, the emittance describes the beam quality along the whole orbit. Behind a focusing quadrupole, the ellipse is vertically oriented, with small dimension in x and larger dimension in x' . Behind a defocusing quadrupole, the ellipse is horizontally oriented with large dimension in x and smaller dimension in x' . Figure 4 shows the phase-space ellipse at the transition between a focusing and a defocusing quadrupole. Because the beam enters a focusing cell with a broader beamwidth and enters a defocusing cell with a smaller beamwidth and the strength of the dipole field increases with the distance from the orbit of the synchronous particle, the effect of passing a complete FODO-cell is a gain of focus in both transverse directions. Calculating the maximum of $x(s)$ in equation 7, for particles that move inside the orbits described by the envelope functions, results in

$$|x - \bar{x}| \leq \sqrt{\varepsilon_x \beta_x(s)} \quad (10)$$

as illustrated in Figure 4. Thus, the beamwidth around the orbit changes with the focusing quadrupole field, but knowing both values, the emittance and the beta-function, gives conclusions

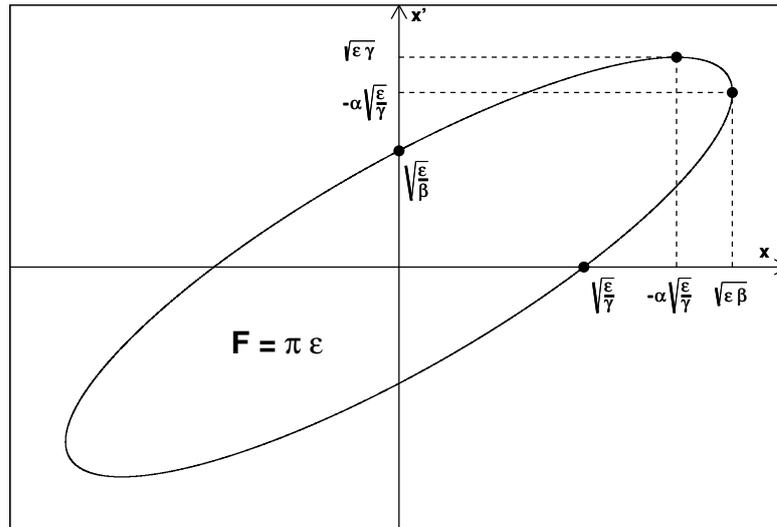


Figure 4: Phase-space ellipse of the particle motion in the x - x' -space [10].

about the beamsize along the orbit.

To minimize the size of the interaction region, the beta-function at the interaction points has to be as low as possible with regards to the limitations of the beam optics behind the interaction point. Because of the outstanding role of the beta-function at the interaction point, it is denoted to be differentiated from the beta-functions at other locations along the orbit as β^* . Figure 5 illustrates the relative beamsizes around the interaction point in the ATLAS detector. Most runs are operated with a crossing angle of $120\ \mu\text{rad}$ between the two beams. Because the beams share a common beampipe of approximately 130 m length at the interaction point of the ATLAS detector, parallel beamlines for the production of head-on collisions would, given a 25 ns bunch distance, produce 34 unwanted, parasitic collision points, leading to a quick loss of beam intensity [7]. The beamlines in figure 5 show the effects of the crossing angle orbit bumps, magnets that separate the two beams just behind the detector.

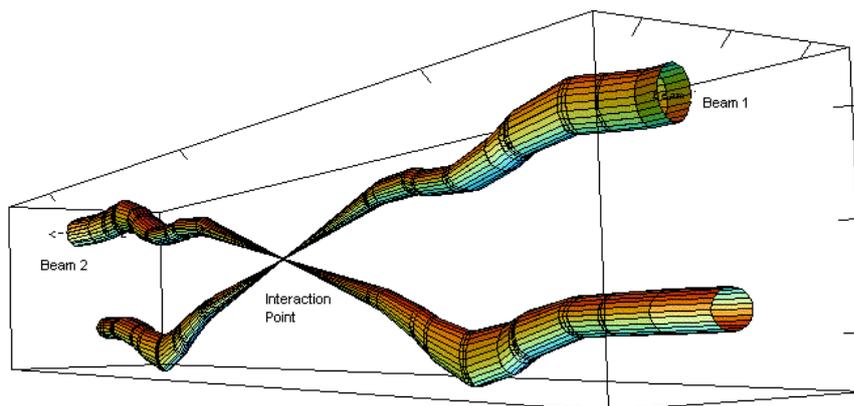


Figure 5: Relative beamsizes around the interaction point in the ATLAS Detector [11].

2.3. ATLAS

ATLAS - A Toroidal LHC AparatuS - is built at intersection point 1 of the LHC and designed as a multi-purpose detector, capable of working in both, proton- and heavy-ion experiments. Its enormous dimensions are 44 m in length, 22 m in height and 7000 tons in weight [12]². Figure 6 provides a cut-away view that illustrates the size and complexity of ATLAS. It can be subdivided into three subdetectors, built around the interaction point in the centre of the detector. In transverse direction the detector layers are installed concentrically around the beryllium beampipe and two endcaps and the endcap-magnets complete the detector in the longitudinal directions.

- The INNER DETECTOR tracks the particles and measures their momentum. It is cylindrical with a length of 7m which begins 5.5 cm away from the beam axis and ends at a distance of 108.2 cm. A solenoid magnet system surrounds the inner detector and provides a 2 T strong field which curves the tracks of charged particles. The inner detector consists of three parts, the pixel detector with 80.3 million readout channels, the silicon microstrip tracker (SCT) with 6.3 million channels and the xenon-based straw-tube transition radiation tracker (TRT) with 0.35 million channels. The inner detector reaches a resolution of 10 μm transverse and 115 μm longitudinal in the pixel-detector, 17 μm transverse and 580 μm longitudinal in the SCT and 130 μm in all directions in the TRT. The fine-grained pixel structure consists of sensor elements of $50 \times 400 \mu\text{m}^2$.
- The CALORIMETER SYSTEM absorbs most of the particles and measures their energies. It is built to detect electromagnetic particle showers caused by photons and electrons and hadronic showers, so-called jets. It determines the transverse energy of the detected parti-

²The technical information in this section is based on „The ATLAS Experiment at the CERN Large Hadron Collider” [12], published by the ATLAS Collaboration. To avoid the too frequent usage of citing this source, it will only be given, where cited literally.

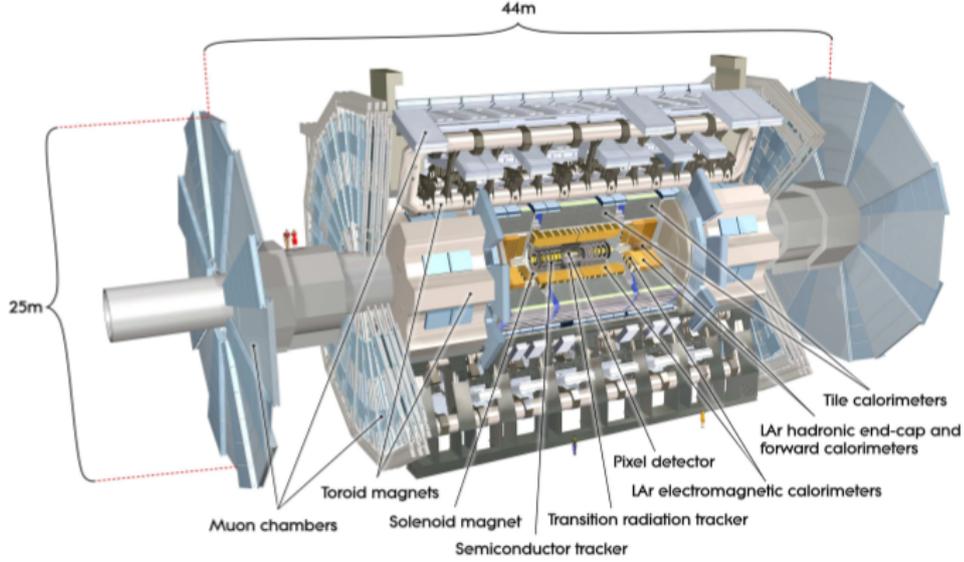


Figure 6: Cut-away view of the ATLAS Detector [12]

cles E_T and the missing transverse energy E_T^{miss} , to collect information about undetected particles. Because neutrinos cannot be expected to be absorbed in the ATLAS detector, their momentum can only be observed due to missing transverse energy. With a view to avoiding hadrons breaking through into the muon system, it absorbs energy in lead and stainless steel layers. It consists of two parts, the electromagnetic calorimeter with a granularity of min. 0.025×0.025 rad and the hadron calorimeter with a granularity of min. 0.1×0.1 rad. In the region $|\eta| < 1.475$, the inner parts are liquid-argon calorimeters located in the cryostat of the superconducting solenoid magnet, and the outer parts used for the hadronic calorimetry which are less cost-intensive iron-scintillator tile-calorimeters. In the region $|\eta| > 1.35$ liquid argon calorimeters, located in the end-cap cryostat, are used for both hadrons and electromagnetic calorimetry.

- The MUON SYSTEM measures the momentum of muons which are not expected to be stopped in the calorimeter system. It is the largest detector part, beginning 4.25 m away from the beam axis and ending in 11 m distance. It has 1 million readout channels in 4 subdetectors, including high precision tracking chambers. The muon system lies in a 0.5 T magnetic field generated by two end-cap magnets and the barrel toroid magnet. This field curves the muon tracks to achieve high precision momentum resolution.

At $z = \pm 17$ m two additional Cherenkov light detectors are installed for luminosity measurements.

The angular range of particle tracks which is covered by the detectors is usually given as pseudorapidity [12], defined by

$$\eta = -\ln \tan \frac{\Theta}{2} \quad (11)$$

where Θ is the polar angle between the track and the beam axis. The function defining η is symmetric by $\eta(\Theta + \pi/2) = -\eta(\Theta - \pi/2)$ with $\eta(\pi/2) = 0$, so the values for the pseudorapidity ranges, listed in table 3, represent intervals around $\Theta = \pi/2$. The inner detector is limited to process tracks which have an angle with the beam axis of at least 9.4° .

Table 3: Angular range covered by the subdetectors of the ATLAS Detector

detector	pseudorapidity η	minimum Θ in degrees
inner detector	± 2.5	$\pm 9.4^\circ$
EM calorimeter	± 3.2	$\pm 4.7^\circ$
hadronic calorimeter barrel and endcap	± 3.2	$\pm 4.7^\circ$
hadronic calorimeter forward	$3.1 \leq \eta \leq 4.9$	$\pm 0.9^\circ \leq \Theta \leq \pm 5.2^\circ$
muon spectrometer	± 2.7	$\pm 7.7^\circ$

2.3.1. Luminosity

The event rate generated by the LHC beams in the ATLAS detector is a key value in measuring its performance. This rate is usually given normalized by the cross section σ of the examined type of particle collision, which is set by the nature of the collision process. When examining a number of collision types, the cross sections of the single processes are added up to the resulting cross section. The normalized event rate is called instantaneous luminosity.

$$\mathcal{L} = \frac{\dot{N}}{\sigma} \quad (12)$$

For inelastic proton-proton collisions at 7 TeV, σ has been measured by the ATLAS Collaboration to be 60.3 ± 2.1 mb [13]. The common unit is the barn, having $1 \text{ b} = 10^{-24} \text{ cm}^2$. Integrating the luminosity over a time interval, like a run or an operation period, gives the total number of events normalized by the cross section. The integrated luminosity is usually given in inverse femtobarns fb^{-1} . During stable beams, for approximation in short intervals of time, the luminosity can be treated as a constant. This is why the time axis in the produced datasets is divided into so-called luminosity blocks, which have an average length of about 60 s. The luminosity falls during a run because of the loss of the protons destroyed in the bunch crossings. Collisions with rest gas particles and other effects leading to emittance growth cause further decreasing.

Figure 7 shows the luminosity for run 184169, recorded online for the luminosity blocks with stable beams [14]. The run starts with a peak luminosity of 1.26 inverse nanobarns per second, so according to equation 12, at the beginning of the run the rate of inelastic proton-proton collisions is $N = 1.26 (\text{nb} \cdot \text{s})^{-1} \cdot 60 \text{ mb} = 75.6 \text{ MHz}$. Divided by the revolution frequency of 11.245 kHz and the number of bunches $N_b = 1318$, this corresponds to an average of 5.1 collisions per bunch crossing at the beginning of the run. The luminosity of the ATLAS detector is expected to rise above $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, when the LHC will be operated at its full design energy. The expected event rate is at least 600 MHz, with 19 collisions per bunch crossing at a bunch crossing

frequency of 40 MHz [7].

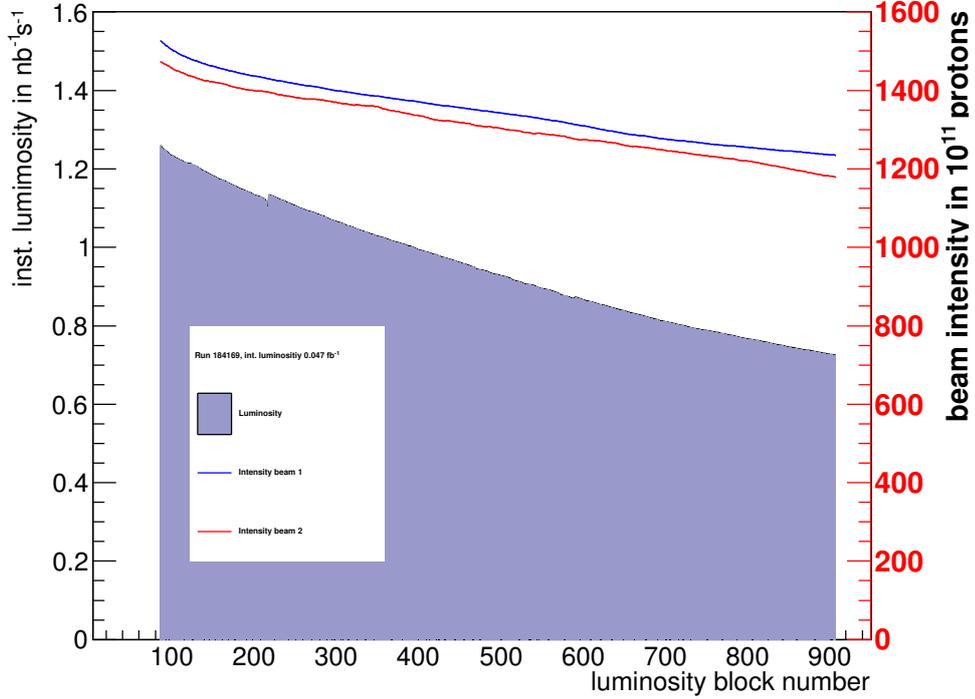


Figure 7: Online luminosity and beam intensity for run 184169, June 28 2011

To reach a higher luminosity, one has four parameters to optimize:

- revolution frequency
- number of bunches
- number of protons per bunch
- size of the luminosity region

Given that $v \simeq c$, the revolution frequency of the highly relativistic protons cannot be increased significantly. The first three of these parameters are found to be proportional to the luminosity, the latter anti-proportional [10]. The area of the luminosity region in the transverse plane can be described by an ellipse with $A = \pi \sqrt{\epsilon_x \beta_x^*} \sqrt{\epsilon_y \beta_y^*}$, using the relation between beamwidth and emittance introduced in equation 10. Thus the luminosity can be given as

$$\mathcal{L} = \frac{f N_1 N_2}{\pi \cdot \sqrt{\epsilon_x \beta_x^* \epsilon_y \beta_y^*}} \quad (13)$$

Figure 8 shows the evolving of the integrated luminosity during the proton-proton program for the high luminosity experiments ATLAS and CMS, as well as for LHCb. The luminosity

production of ALICE cannot be resolved at the used scale. It reaches about 5 pb^{-1} in the same period.

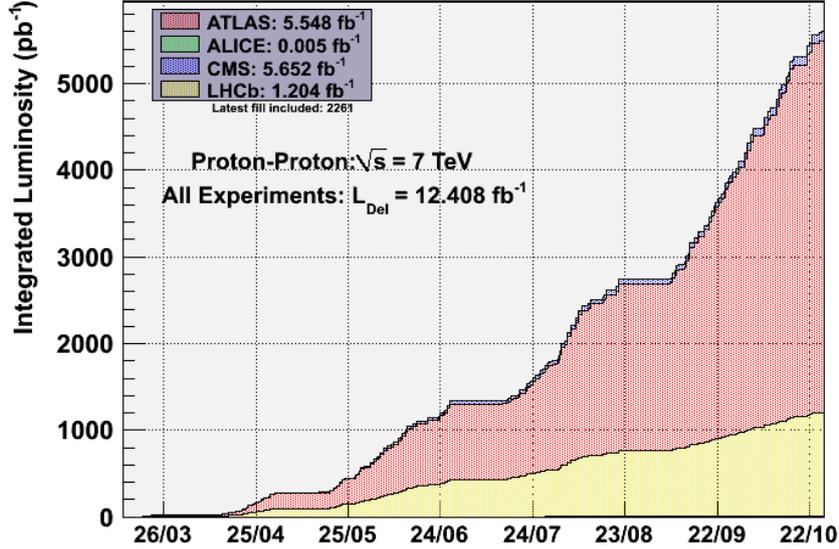


Figure 8: Luminosity production 2011 during the proton-proton physics program [15].

2.3.2. Vertex reconstruction

ATLAS is not capable of processing and storing all data that amounts from a collision rate of 600 MHz or higher. But since the majority of the events do not match the energy scale and processes that ATLAS aims to observe, it is neither necessary nor beneficial to process the whole amount of data. A multistage triggering system is in place to filter events for either storage or deletion. The Level-1-Trigger (L1) is used to filter events in real time, reducing the data rate to about 75 kHz. The Level-2-Trigger (L2) and the event filter (EF) introduce further limitations to the process and reduce the data to about 200-300 Hz. The L2 trigger and the event filter are together referred to as High Level Trigger (HLT). The challenge presented to the trigger system implementation in the ATHENA frame work is in balancing the opposing demands of achieving high resolution and saving CPU time used by the required algorithms.

The L1 trigger uses information of the calorimeters and the muon detector to trigger for a minimum of E_T and E_T^{miss} . The values can be adjusted in the trigger menu individually for different event types, an overview is given in table 4. It builds tracks from the hits in the inner detector and assigns the transverse momentum to those tracks. In normal operation it triggers tracks with a $p_T \geq 0.5 \text{ GeV}$ and in minimum bias experiments this can be levelled down to $p_T \geq 0.1 \text{ GeV}$. Since the luminosity region is of order $O(1000)$ larger in longitudinal direction than in the vertical directions, the tracks are grouped into clusters by the z -coordinates from

which they are fitted to originate [16].

The tracks cannot be assigned to starting points in the transverse plane due to the detector geometry. Thus knowledge of the beamline is of importance here in order to find the so-called perigees, the points where the tracks are closest to the beamline. The coordinate system in place here is different to the one described in the theory of the particle dynamics in section 2.2, since it is not centred at the beam but at the centre of the ATLAS detector. The basic orientation, however does match, since the positive z-axis points along the main axis of the detector in the direction of beam 2, the x-axis points horizontally to the centre of the LHC and the y-axis upwards in vertical direction.

The found clusters, so-called Region of Interests (ROI), are passed to the L2 trigger and invoke a vertex fitting algorithm performed by the HLT. For each ROI, a list of tracks sorted by momentum is processed, inside-out from the highest resolution pixel tracker to the TRT, extending the tracks through the distances between the detector layers. Tracks which do not satisfy a predefined fit quality are rejected at this stage, by setting a limit of the accepted χ^2 to 8% [17]. The L2 trigger passes the vertex position and its covariance matrix as well as information about the fit quality and the tracks to the event filter, where events are built according to the Event Data Model (EDM).

The described method performed online is called “fitting after finding”. It is necessary to

Table 4: Trigger configuration used at the end of the 2011 proton-proton program and the observed event rates [18]. The rates refer to a luminosity of $3.3 \text{ nb}^{-1}\text{s}^{-1}$. By triggering for higher energies than the L1 trigger, the HLT decreases the event rates.

	L1 trigger		event filter	
	E in GeV/c^2	\dot{N} in Hz	E in GeV/c^2	\dot{N} in Hz
single muon	11	8000	18	100
single electron	16	9000	22	55
two muons	11, 11	8000	15, 10	4
two electrons	10, 10	2000	12, 12	3
two taus	15, 10	7500	29, 20	15
2 photons	12, 12	3500	20, 20	5
single jet & E_T^{miss}	50, 35	800	75, 55	18
E_T^{miss}	50	600	70	5
multi-jets	5×10	200	5×30	9
		<75000		214

bear in mind that the number of vertex candidates is set at the L1 trigger, and therefore a track rejected by one of the fitting processes at the L2 trigger is not regarded by the processes fitting the other vertex candidates. Additional offline performed vertex fitting algorithms can provide improvements on this issue of the online performed algorithms using the “finding through fitting” method. They initially fit all tracks to one single vertex at first and iteratively reject the tracks that do not satisfy a χ^2 limit to be originating from this vertex. Thereby an additional vertex is formed by the rejected tracks and in the next step the probability for each track is computed

for both vertices, creating a third vertex and so forth. Tracks that do not enter any vertex with a probability of χ^2 higher than 1% are finally rejected. Different implementations of this method are part of the ATHENA framework [17].

2.3.3. Vertex distribution

The high rate of events allows significant statistics to be gathered regarding the size of the luminosity region. The vertex distribution ρ in each of the three spatial dimensions is expected to be Gaussian in good approximation. Equation 14 describes the three dimensional distribution for one pair of bunches with N_b particles, disregarding time-wise effects [10].

$$\rho(x, y, z) = \frac{N_b}{\sqrt{(2\pi)^3 \sigma_{x,b}^2 \sigma_{y,b}^2 \sigma_{z,b}^2}} \cdot e^{-\frac{1}{2} \left(\frac{x-\bar{x}_b}{\sigma_{x,b}} \right)^2 - \frac{1}{2} \left(\frac{y-\bar{y}_b}{\sigma_{y,b}} \right)^2 - \frac{1}{2} \left(\frac{z-\bar{z}_b}{\sigma_{z,b}} \right)^2} \quad (14)$$

In this thesis $\sigma_{i,b}^2 = \sigma_i^2$ where $i \in \{x, y, z\}$ will be set for all bunches and equation 14 is applied to the distribution of the whole data set. The distribution in the x-z-plane is described by

$$\rho(x, z, t) = \frac{N}{2\pi\sigma_x^2\sigma_z^2} \cdot e^{-\frac{1}{2} \left(\frac{x-\bar{x}}{\sigma_x} \right)^2 - \frac{1}{2} \left(\frac{z-\bar{z}}{\sigma_z} \right)^2} \quad (15)$$

Analogue equations count in the other principal planes of the coordinate system. Throughout this thesis, equations for single coordinates will be given for the coordinate x to avoid the frequent usage of the expression $i \in \{x, y, z\}$. It will be noted in the particular context, whether the evaluations count in the transverse plane only, or whether they can also be deployed for the longitudinal dimension as well.

For proton fills the beamwidth is usually given as the range $|x - \bar{x}| \leq 2\sigma$. Because the beamwidth represents the dimension of the phase-space ellipse along the x -axis, equation 10 can be used to determine the emittance by measuring the beamwidth.

$$\varepsilon_x = \frac{(2\sigma_x)^2}{\beta_x^*} \quad (16)$$

Using this definition, equation 13, that expresses the relation between the transverse beamwidth and the luminosity, can be written as

$$\mathcal{L} = \frac{fN_1N_2}{4\pi\sigma_x\sigma_y} \quad (17)$$

Since the standard deviation σ is the squareroot of the variance, it can be calculated from the vertices of the recorded events using the computational formula for the variance:

$$\sigma_x = \sqrt{\text{Var}(x)} = \sqrt{\overline{x^2} - \bar{x}^2} \quad (18)$$

This is applied by ROOT to compute the RMS along histogram axes.

2.3.4. Vertex resolution

The distribution of the recorded vertices, the observed distribution, has to be distinguished from the real luminosity region. The real vertex distribution cannot be observed directly by examining the vertex positions in the recorded data, since each entry is superimposed by the uncertainty of the vertex fitting. Thus, the observed distribution is given by the result of the convolution of the real vertex distribution with the distribution of the vertex resolution. Having the two Gaussian functions $\rho_{real}(x)$ and $\rho_{res}(x)$, describing both distributions, their convolution is given by

$$\rho_{obs}(x) = \rho_{real}(x) * \rho_{res}(x) = \left(\frac{1}{\sqrt{2\pi\sigma_{real}^2}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}_{real}}{\sigma_{real}}\right)^2} \right) * \left(\frac{1}{\sqrt{2\pi\sigma_{res}^2}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}_{res}}{\sigma_{res}}\right)^2} \right) \quad (19)$$

The convolution of the two Gaussian functions is again Gaussian and evaluates to equation 20 [19].

$$\rho_{obs}(x) = \frac{1}{\sqrt{2\pi(\sigma_{real}^2 + \sigma_{res}^2)}} e^{-\frac{1}{2}\frac{(x-(\bar{x}_{real}+\bar{x}_{res}))^2}{\sigma_{real}^2 + \sigma_{res}^2}} \quad (20)$$

Given that the vertex resolution can be assumed to distribute around $\bar{x}_{res} = 0$, as it describes a confidence interval, the observed beamspot coordinates \bar{x}_{obs} in all three dimensions are given as

$$\bar{x}_{obs} = \bar{x}_{real} + \bar{x}_{res} = \bar{x}_{real} \quad (21)$$

Hence equation 20 can be simplified to

$$f_{obs}(x) = \frac{1}{\sqrt{2\pi(\sigma_{real}^2 + \sigma_{res}^2)}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}_{obs}}{\sqrt{\sigma_{real}^2 + \sigma_{res}^2}}\right)^2} \quad (22)$$

which is a Gaussian distribution with the standard deviation given by equation 23.

$$\sigma_{obs} = \sqrt{\sigma_{real}^2 + \sigma_{res}^2} \quad (23)$$

3. Run selection

After a technical stop in January and hardware commissioning in February, LHC began recommissioning with beam at the end of March 2011. The full physics program was successfully executed until the end of October, with just a few short technical stops. Runs with heavy-ion beams were only performed over a period of three weeks in November and December. The

proton-proton program had a total duration of 229 days and 5 hours, 23.5% of that time the LHC provided stable beams for the experiments. This means that LHC provided stable proton beams for ATLAS for a total of 54 days and 16 hours. The runs had an average duration of six hours and there were a total of 49 runs lasting 10 hours or longer.

With a view to the size of the data sets, a selection of those runs has been chosen for this thesis. At the time of data preparation, the only data available was for runs earlier than July 1st. This limitation led to a selection of 23 runs, the first run starting on April 15th, the latest run starting on June 28th. An overview of the runs and their beam parameters is given in table 3. Over the analysed time period, the beam intensity has been increased by injecting more bunches, but the average intensity per single bunch was stable at $(2.36 \pm 0.12) \cdot 10^{11}$ protons. The integrated luminosity for the examined runs is 0.756 fb^{-1} .

Table 5: Selected runs and their fill parameters. Not listed are β^* , which was 1.5 m in all runs, except run 180153 with $\beta^*=2$ m. The bunch distance for all runs was 50 ns, except run 178109 for which it was 75 ns. In all runs the crossing angle was $120 \mu\text{rad}$.

run	start	duration	colliding bunches	protons beam 1	protons beam 2	average protons per bunch	peak luminosity	integrated luminosity	average transverse emittance
		h:min		10^{11}	10^{11}	10^{11}	$\text{nb}^{-1}\text{s}^{-1}$	fb^{-1}	μm
184169	Jun 28 02:23:59	13:36	1318	1540	1480	2.29	1.26	0.047	5.98
184130	Jun 27 02:32:34	19:44	1180	1390	1360	2.33	1.28	0.063	^a
184022	Jun 24 22:57:34	14:04	1180	1360	1340	2.29	1.19	0.046	5.96
183780	Jun 21 00:17:17	17:54	1041	1210	1200	2.32	1.13	0.048	6.4
183462	Jun 14 07:34:02	15:03	1042	1230	1240	2.37	1.17	0.048	6.39
183407	Jun 12 14:00:24	13:22	1042	1230	1230	2.36	1.17	0.043	6.31
183286	Jun 09 14:34:51	10:16	1042	1150	1170	2.23	1.04	0.031	6.26
183216	Jun 07 22:37:28	10:44	1042	1240	1230	2.37	1.05	0.031	6.66
183130	Jun 06 23:43:14	7:45	1042	1130	1130	2.17	0.82	0.021	5.98
183081	Jun 05 09:30:17	15:00	1042	1210	1210	2.32	1.14	0.046	6.42
183003	Jun 02 02:08:36	14:43	1042	1190	1190	2.28	1.17	0.047	6.55
182787	May 29 10:57:48	10:21	1042	1240	1250	2.39	1.28	0.038	6.37
182747	May 28 08:38:42	10:12	874	1060	1090	2.46	1.0	0.031	6.26
182486	May 24 02:19:28	10:00	874	1040	1040	2.38	1.01	0.03	6.14
182456	May 23 10:25:45	9:14	874	999	987	2.27	0.81	0.03	7.05
182424	May 21 20:40:10	9:59	874	1040	1410	2.8	0.86	0.026	7.16
180636	Apr 30 13:43:22	15:43	598	736	722	2.44	0.68	0.029	5.91
180481	Apr 28 21:45:30	13:13	598	721	708	2.39	0.67	0.025	2.33
180400	Apr 27 23:18:57	10:01	598	735	713	2.42	0.66	0.019	5.87
180225	Apr 24 21:11:28	16:18	424	484	479	2.27	0.45	0.019	5.5
180164	Apr 23 12:30:09	17:52	424	520	508	2.42	0.48	0.023	6
180153	Apr 22 22:40:05	8:53	424	485	485	2.29	0.37	0.01	4.2
179710	Apr 15 02:11:14	9:55	214	264	258	2.44	0.22	0.007	6.4

^aThe values given for the average vertical emittance are the mean values of the horizontal and vertical emittance determined by the synchrotron radiation telescopes. The one left out is of order $O(1000)$ in the vertical direction. This is very likely to be a defect, the value is neglected.

4. Observed beamspot and beamwidth for primary vertices

As a preparation for the following evaluations, the distributions of the vertex positions in all three directions are examined in this section. Their mean values represent the coordinates of the average primary interaction point. The evolving of the observed beamwidth is compared for the period covering all 23 runs. In the analysis part of this thesis, results of the evaluations for single runs are always shown for the latest run, number 184169. Figures illustrating the results are given for this run as well. Plots that show the central results for the other runs are attached to appendix. The fill parameters of the 23 runs show only little difference, so a comparability of the results for run 184169 with the other runs is expected. Parameters evaluated for the single runs thus are collected together in timeline plots, to review the found results for the whole run sequence and confirm the findings for run 184169.

4.1. Distribution in the ATLAS coordinate system

To find the average interaction point, the data for each run is used to fill a histogram, which is then fitted by a Gaussian function. Figure 9 shows that in both transverse directions as well as in the longitudinal direction, the observation and the model of a normal distribution agree well. The relative error for the calculated σ is 0.06% transverse and 0.05% longitudinal. The average beamspot error is $0.015 \mu\text{m}$ for x , $0.013 \mu\text{m}$ for y and 0.04 mm for z . Only for bins that lie close to the mean of the histograms do the fit and the data have less conformity than at the slopes and tails of the distribution.

The difference between the Gaussian distribution and the observed distribution that occurs in the bins close to the mean is observed to follow the same pattern in all examined runs. In the transverse directions, the observed distribution shows significantly more vertices found in a region of about 0.3σ around the mean than would be expected in a Gaussian vertex distribution. Opposite to this, in the longitudinal direction the values close to the mean lie slightly below the fit function. This is found to be the same in all examined runs too.

The fit results are summarised in the timeline plots in figures 10 and 11 to examine the evolving of the average beamwidth and the position of the average interaction point through the investigated sequence of all 23 runs. Only the first run shows a larger deviation of the beamspot's y -coordinate. All other runs only show a low variation of the transverse coordinates in ranges of about $\pm 20 \mu\text{m}$ in x -direction and between $+40 \mu\text{m}$ and $-20 \mu\text{m}$ in y -direction. Compared to the average beamwidth, the beamspot moved in x -direction within a range of about $2 \sigma_x$ and in y -direction within a range of about $3.5 \sigma_y$. Compared to this, the relative variation of the beamspot in z -direction is much smaller, since it only varies over a range of 15 mm , which is about a quarter of the average σ_z .

In all examined runs the form of the transverse beam projection is not perfectly circular, but an ellipse with a slightly smaller axis in y -direction, with an average of $19.2 \pm 0.6 \mu\text{m}$, than in x -direction, with an average of $21.06 \pm 0.7 \mu\text{m}$. These mean values are calculated as the weighted

4.1 Distribution in the ATLAS coordinate system

mean of the sets of the 23 average beamspot coordinates computed for the analysed runs. The found results are listed in table 6.

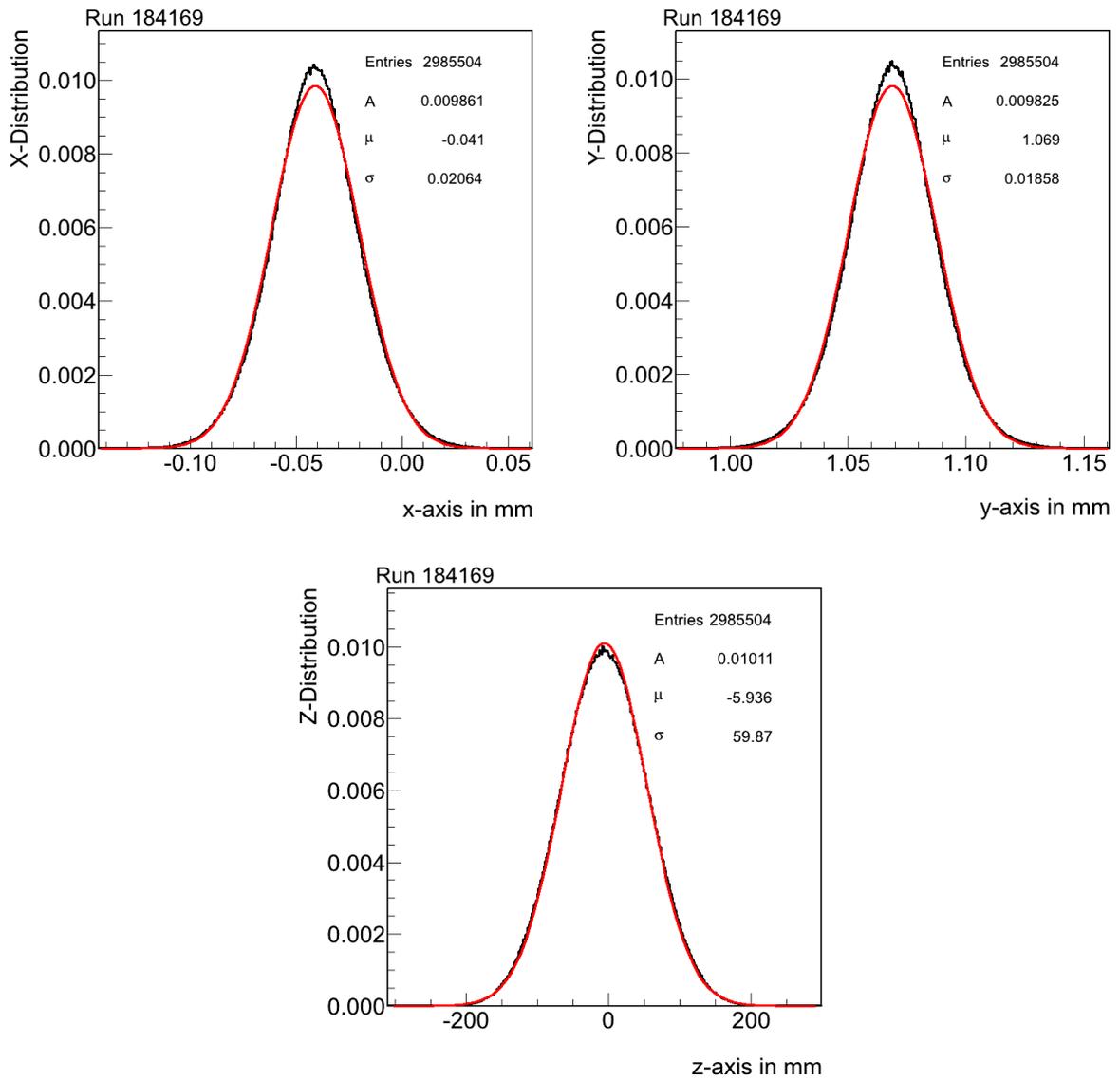


Figure 9: Spatial distribution for all primary vertices of run 184169. The transverse distributions show an accumulation of vertices close to the mean. The longitudinal distribution shows a better agreement of model and observation.

4.1 Distribution in the ATLAS coordinate system

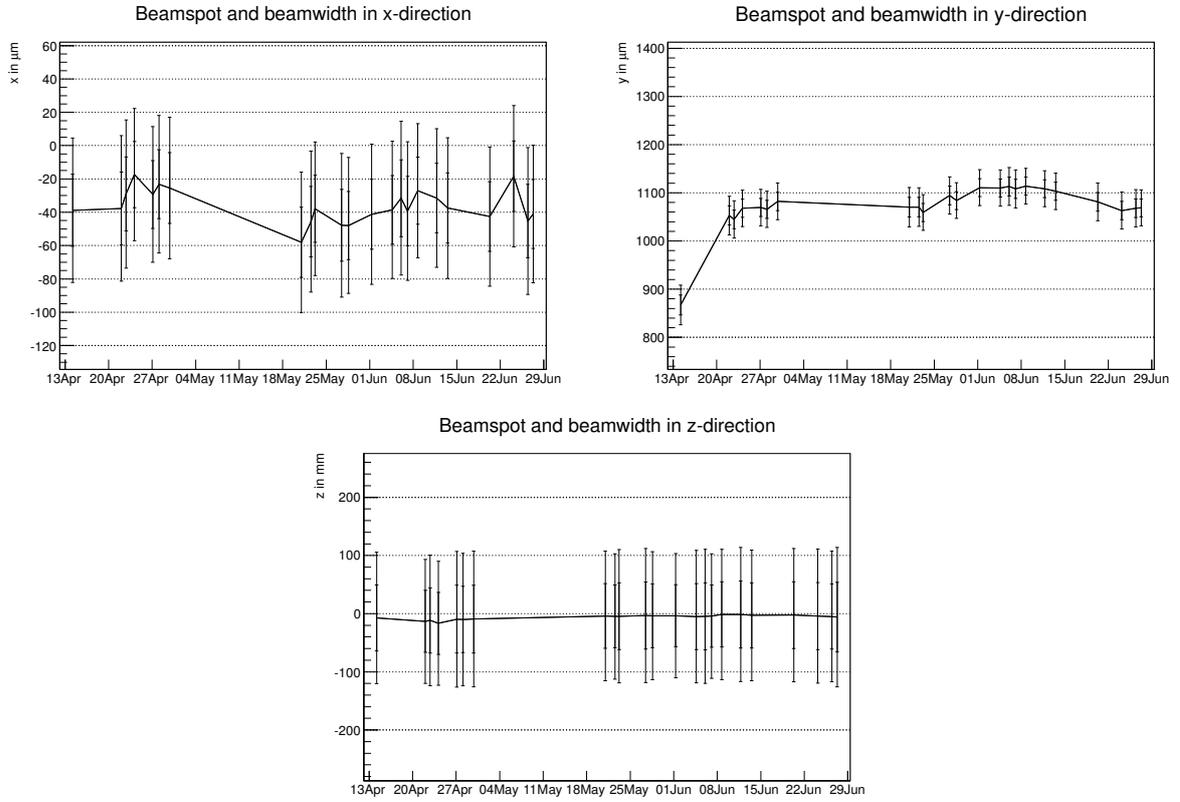


Figure 10: Spatial distribution using all primary vertices. The errorbars indicate the beamwidth 1σ range and 2σ range.

Table 6: σ for primary vertices, average over all 23 runs

Axis	average σ	average σ fit error
x in μm	21.06 ± 0.7	0.06%
y in μm	19.2 ± 0.6	0.06%
z in mm	56.3 ± 1.8	0.05%

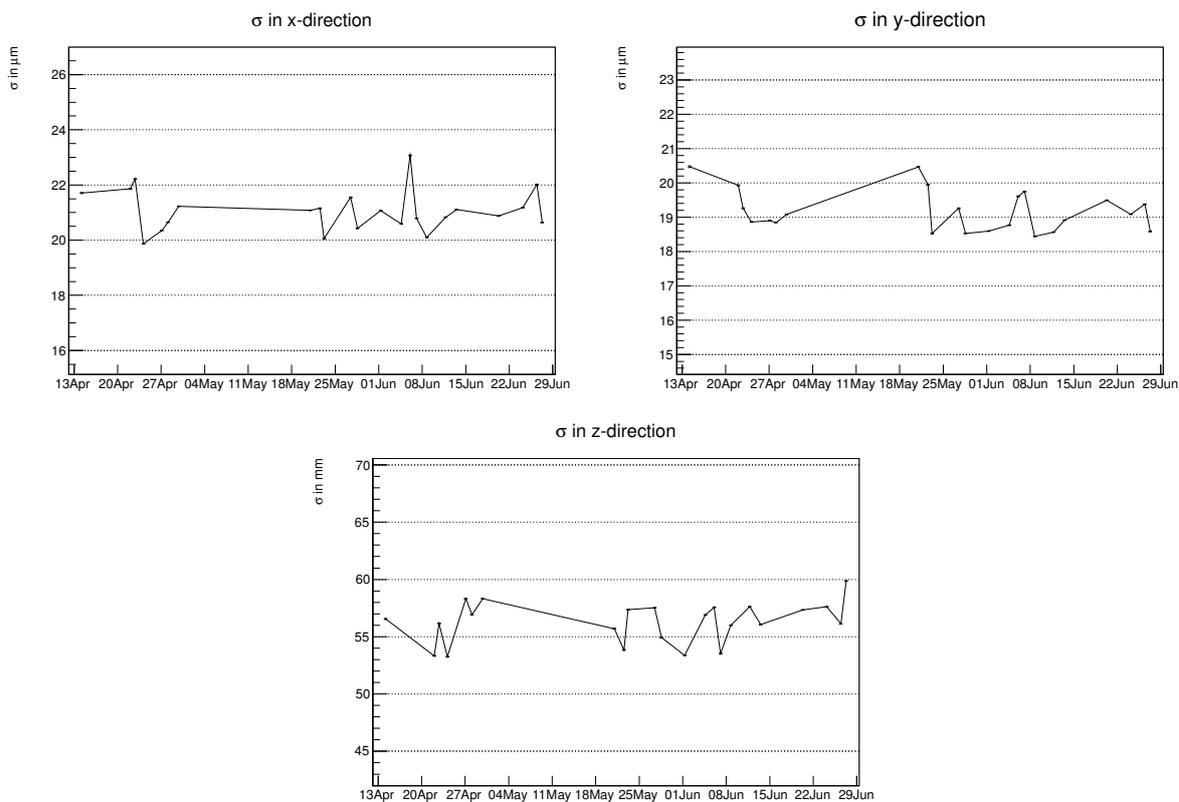


Figure 11: Variation of the observed beamwidth for all 23 runs.

4.2. Tilt in the ATLAS coordinate system

The data entries are given in the coordinate system of the ATLAS detector. But the beamline does not necessarily follow a path parallel to the longitudinal axis of this system. By looking at the transverse beamspot coordinates along the beamline it is shown in this section, that the beamline is orientated slightly downwards and to the left relative to the z-axis. Thus the transverse plane, used for the calculations in section 4.1, is not exactly orthogonal to the beam. The slope of the beamline in the transverse directions is called tilt in x-direction and in y-direction. The crossing angle of $-120 \mu\text{rad}$ has no impact on the position of the mean, because it is located between the two beams and only affects the beamwidth along the z-axis. The route of the beamspot along the z-axis is shown in figure 12.

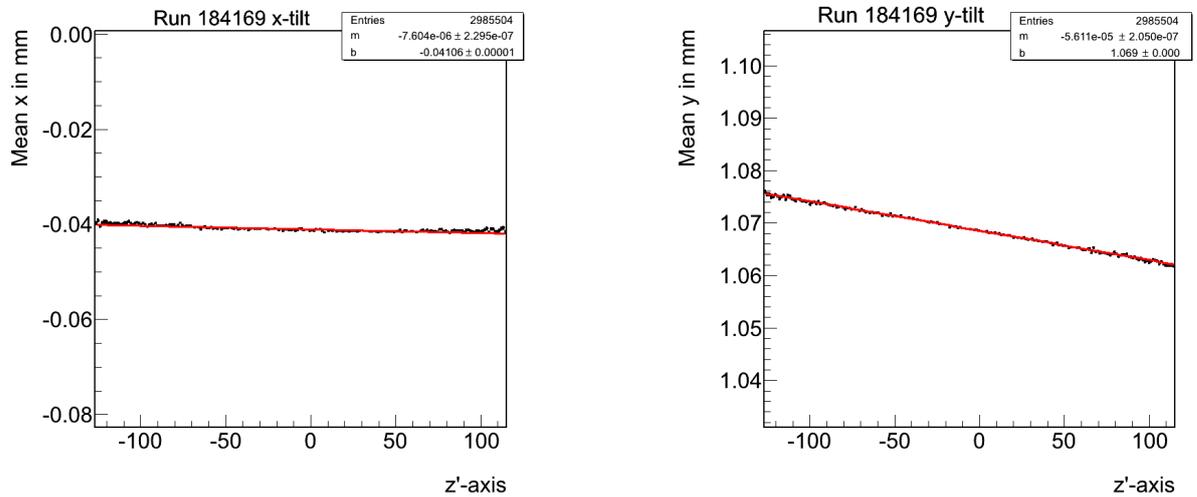


Figure 12: Observed beamline tilt of run 184169 for both transverse directions. The z-interval represents $\bar{z} \pm 2\sigma_z$, the interval on the vertical axis represents $\bar{x} \pm 2\sigma_x$, respectively $\bar{y} \pm 2\sigma_y$.

As expected, the beamspot positions in the 100 bins along the z-axis are fitted very well by the linear functions. To avoid bins with low statistics in the outer region of the luminosity zone, the range is limited to $\bar{z} \pm 2\sigma_z$. This cut along the z-axis is expected to contain more than 95% of the vertices. The vertical axes ranges are also limited to $\bar{x} \pm 2\sigma_x$ and $\bar{y} \pm 2\sigma_y$, this corresponds to the definition of the transverse beamwidth. The linear functions $x = m_x \cdot z + b_x$ and $y = m_y \cdot z + b_y$ follow the beamline projections onto the x-z-plane and the y-z-plane. The slopes m_x and m_y represent the tilt in the particular transverse direction and the products $m_x \cdot (z - \bar{z})$ and $m_y \cdot (z - \bar{z})$ calculate the additional offset from the z-axis of the ATLAS coordinate system due to tilt. In x-direction the tilt of run 184169 is $(-7.6 \pm 0.3) \cdot 10^{-6} \text{ mm}^{-1}$, which is significantly smaller than in y-direction where it is $(-56 \pm 0.2) \cdot 10^{-6} \text{ mm}^{-1}$. With such a flat pathway, the tilt in x-direction is almost negligible even on the small scale of the luminosity region. Figure 13 shows the tilt for the other runs of the examined run sequence. The values of both series vary only in a small range

4.2 Tilt in the ATLAS coordinate system

around the mean values $\bar{m}_x = (-9.4 \pm 2.4) \cdot 10^{-6} \text{ mm}^{-1}$ and $\bar{m}_y = (-51.7 \pm 3.7) \cdot 10^{-6} \text{ mm}^{-1}$.

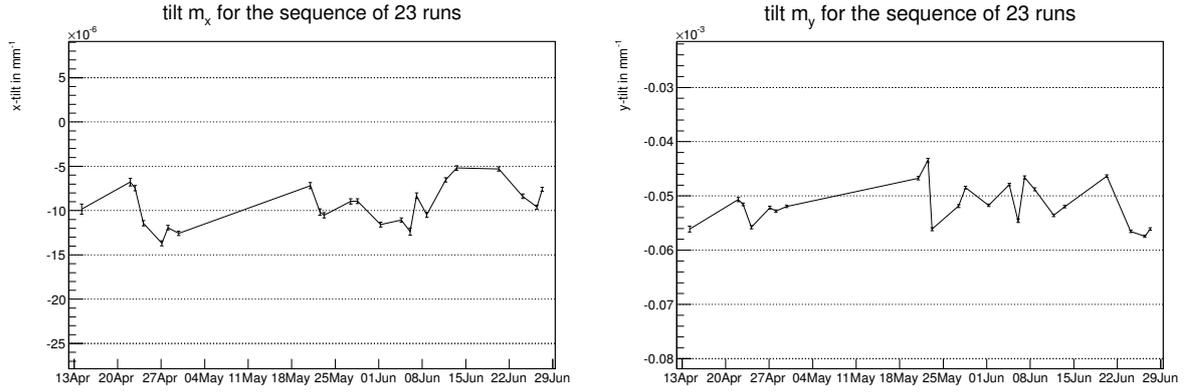


Figure 13: Observed beamline tilt in mm^{-1} for all runs. The plots show m_x and m_y found by the method used in figure 12. Both tilts are actually negative, but plotted with their absolute values to improve the readability of the plots.

Figure 14 shows the findings expressed by the relative size of the offset that the tilt adds to the beamspot position along a range of σ_z , compared to the observed σ in the transverse directions. Calculated over all examined runs, the beamline tilt in x-direction has an average of $-2.5 \pm 0.7\%$

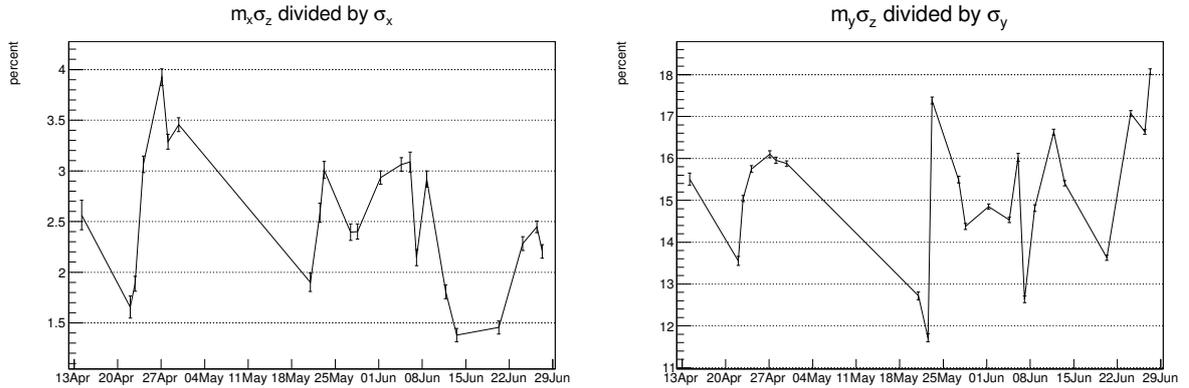


Figure 14: Observed beamline tilt for all runs, expressed by the ratio of the beam offset and the beamwidth σ in the given direction. $m\sigma_z$ is the offset that the tilt adds to the beamline along a range of σ_z . Both tilts are actually negative, but plotted with their absolute values to improve the readability of the plots.

of the half beamwidth σ_x . In y-direction, the effect is significantly stronger with an average of $-15.2 \pm 1.6\%$ of σ_y . The tilt is illustrated in form of the found beamline in the plots in figure 15. Vertices that have a greater distance from the average interaction point along the z-axis, thus add a systematic bias to the observed beamwidth, when it is calculated in the ATLAS coordinate system. This is analysed in the next section, by an adjustment of the coordinate system to the beamline and comparison of the results of the beamwidth evaluations in both coordinate systems.

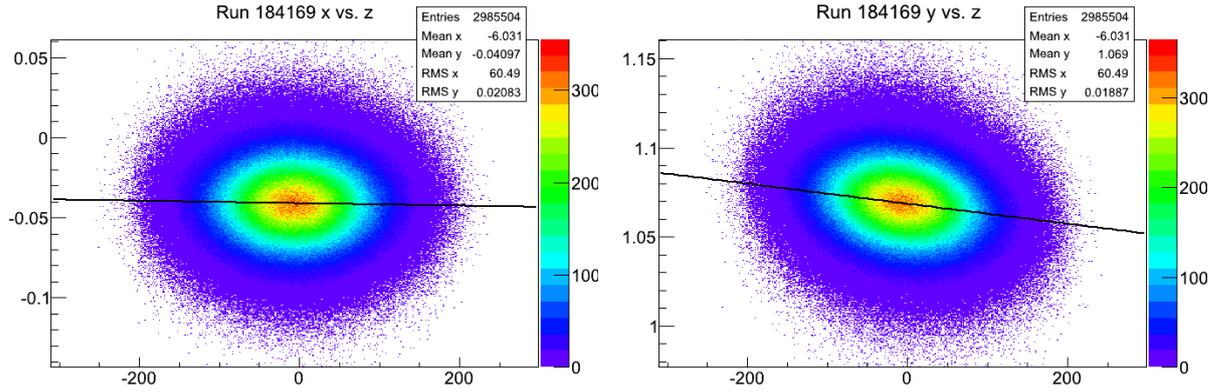


Figure 15: 2-dimensional distributions in the x-z-plane and the y-z-plane with the beamline found by bin-wise fitting. The axis intervals are set to $\pm 5\sigma$ on all axes. The z-interval is about 3000 times larger than the x- and y-intervals.

Finite binwidth effects This computation contains a less obvious bias that affects the bin-wise average beamspot position along the z-axis. The beamline found in the vertical-longitudinal plane still has a minimal deviation from the apparently somewhat steeper path one would draw by hand. The negative y-tilt leads to an orientation of the ellipse in the y-z-plane that is equivalent to the result of a clockwise rotation around the average interaction point $IP_{yz} = (\bar{y}/\bar{z})$. At any point $P = (y/z)$ in that plane, the gradient of the point density $\nabla\rho(y,z)$ is expected to be orientated towards IP_{yz} . This gives $d\rho(y,z)/dz > 0$ for $z < \bar{z}$ and otherwise $d\rho(y,z)/dz < 0$ for $z > \bar{z}$. This leads to a mirror-symmetric effect in any interval $\Delta z_0 = (z_0 - dz, z_0 + dz)$, but the method of finding the tilt relied on just such intervals. The rotation shifts points in the upper-left and lower-right quarter of the ellipse with a parallel projection on the z-component of $\nabla\rho_{y,z}(y,z)$, but in the lower-left and upper-right quarter with an anti-parallel projection. Thus for $z_0 < \bar{z}$ the point density is decreased in the upper region of the interval and is increased in the lower region of the interval, that way the mean is found too low.

In turn, for $z > \bar{z}$, the point density is increased in the upper region of the interval and decreased in the lower region of the interval and the mean is found too high. In the next section the impact of this effect on the beamwidth measurements will be addressed, by rotating the coordinate system to adjust it to the beamline and review the tilt again. Because the values for the tilt determined in this section are used for the rotation, any tilt still present after this process can be seen as based on the analysed bias. The next subsection will show that this remaining tilt is on a neglectable scale and the discussed inaccuracy of the process is small even relative to the small scale on the transverse axes.

4.3. Rotation of the coordinate system

As shown in the last section, the beamline is not exactly parallel to the z-axis, but has angles with it in both the x- and y-direction. Following the beam along a distance Δz in z-direction

leads to an additional offset from the z -axis, given by $dx/dz \cdot \Delta z = m_x \cdot \Delta z$ in x -direction and by $dy/dz \cdot \Delta z = m_y \cdot \Delta z$ in y -direction. Thus the beam points from the interaction point to the point $(m_x, m_y, 1)$. To adjust the coordinate system to the beamline, this vector is used as the z' -direction in the new coordinate system K' . Satisfying the condition that the new x' -axis is still horizontal and orthogonal to the new z' -axis, it has to be parallel to the vector $(1, 0, -m_x)$. The vector product $\vec{y}' = \vec{z}' \times \vec{x}'$ gives the new y' -direction and a right-handed coordinate system K' , with the y' -axis parallel to the vector $(-m_x m_y, 1 + m_x^2, -m_y)$ still pointing upwards. To achieve the change of basis matrix $T_{K'}^E$, that has to conserve distances and angles, the new basis vectors have to be normalized.

$$T_{K'}^E = \left(T_E^{K'} \right)^{-1} = \begin{pmatrix} \frac{1}{\sqrt{1+m_x^2}} & -\frac{m_x m_y}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} & \frac{m_x}{\sqrt{m_x^2+m_y^2+1}} \\ 0 & \frac{1+m_x^2}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} & \frac{m_y}{\sqrt{m_x^2+m_y^2+1}} \\ -\frac{m_x}{\sqrt{1+m_x^2}} & -\frac{m_y}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} & \frac{1}{\sqrt{m_x^2+m_y^2+1}} \end{pmatrix}^{-1} \quad (24)$$

Since $T_E^{K'}$ is an orthogonal and normalized matrix, $\vec{x} \mapsto T_E^{K'} \cdot \vec{x}$ is an isometry and the inverse is simply calculated by transposition. Following this, the coordinates in the system K' are calculated by equation 25. The process of basis change is equivalent to a three-step rotation of the ATLAS coordinate system (Euler's rotation theorem).

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+m_x^2}} & 0 & -\frac{m_x}{\sqrt{1+m_x^2}} \\ -\frac{m_x m_y}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} & \frac{1+m_x^2}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} & -\frac{m_y}{\sqrt{(1+m_x^2)(1+m_x^2+m_y^2)}} \\ \frac{m_x}{\sqrt{m_x^2+m_y^2+1}} & \frac{m_y}{\sqrt{m_x^2+m_y^2+1}} & \frac{1}{\sqrt{m_x^2+m_y^2+1}} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (25)$$

To speed up the computation of the coordinate transformation, the change of basis matrix can be simplified by neglecting all terms of order $O(m^2)$, because m is at least of order $O(10^{-4})$ smaller than the coordinate entries in x and y . This simplified matrix is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & -m_x \\ 0 & 1 & -m_y \\ m_x & m_y & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (26)$$

This basis is still orthogonal between z and the transverse directions, but the angle between the x' -axis and the y' -axis, now in the rows of the matrix in equation 26, has to be neglected following the same argument, since their product $\vec{x}' * \vec{y}' = m_x m_y$ is of order $O(m^2)$. The determinant is $1 + m_x^2 + m_y^2$ which can be approximated to be almost 1 following the same argument. It renders

4.3 Rotation of the coordinate system

the process much more readable, as it is clear at first sight that

$$\begin{aligned}x' &= x - m_x z \\y' &= y - m_y z \\z' &= z + m_x x + m_y y\end{aligned}\tag{27}$$

Taking into account, that $m_x x + m_y y$ is again small at the scale of the z -distribution, one can further approximate by $z' = z$. What remains is the substance of the process, which is subtracting the offset brought in by the tilt from the transverse coordinates.

However, the coordinate transformation for the further evaluation was programmed using NumPy [4] by implementing the exact equation 25, to avoid bringing new uncertainties into the process. Figure 16 shows the distributions in the x' - z' -plane and in the y' - z' -plane. The tilt is removed as the ellipses appear to be orientated symmetrically around the z' axis, of which a parallel through the beamspot is added to the plots. This line represents the projection of the found beamline in the plots in figures 12 and 15. The repetition of the tilt-finding algorithm analogue to the procedure

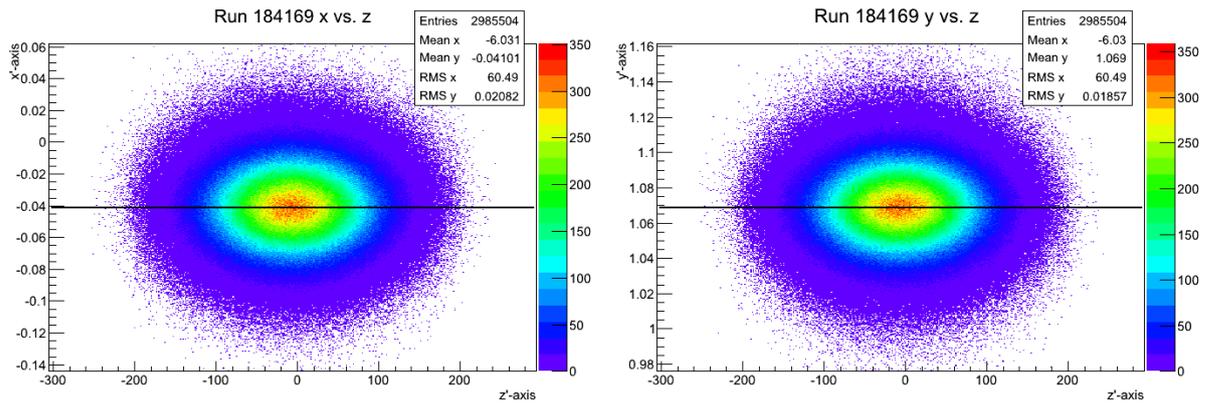


Figure 16: 2-dimensional distributions in the x' - z' -plane and the y' - z' -plane after the rotation of the coordinate system. A horizontal line shows the new position of the beamline that was fitted before and shown in figure 15.

in section 4.2 shows that in both directions the average tilt taken over all runs is reduced by $(99.987 \pm 0.004)\%$ of the value found before the rotation. In both transverse directions the results of the algorithm determine the tilt in the coordinate system K' to add an offset to the beam in order of $O(10^{-4})\%$ of the transverse σ along a z -range of σ_z . The plots for the other runs are shown in the appendix. The beamline in the y - z -plane leads to the same observation in all runs. Before the rotation, the optical impression is that it is apparently somewhat misaligned, but after rotation this effect is no longer noticeable. In all of the following sections, evaluations will be processed using the data sets transformed into the rotated coordinate system.

4.4. Vertex distribution in the rotated coordinate system

The effect of the coordinate transformation on the beamwidth determination is reviewed in this subsection by applying the same algorithms as in section 4.1. The results are not expected to differ significantly, but the found beamwidth is expected to be systematically smaller in the transverse directions, since these directions are adjusted to the minor ellipse axes. In the longitudinal direction, the result is expected to be found systematically higher than in the ATLAS coordinate system, because this direction is adjusted to the major ellipse axis. The minor and major axes correspond with the ellipse diagonals of minimal and maximal length. The effect on the average beamspot described in section 4.2 is mirror-symmetric, thus the result for the beamspot position, after re-transformation into the ATLAS coordinate system, is expected to be the same.

Figure 17 shows the results for this analysis for run number 184169. After the rotation the

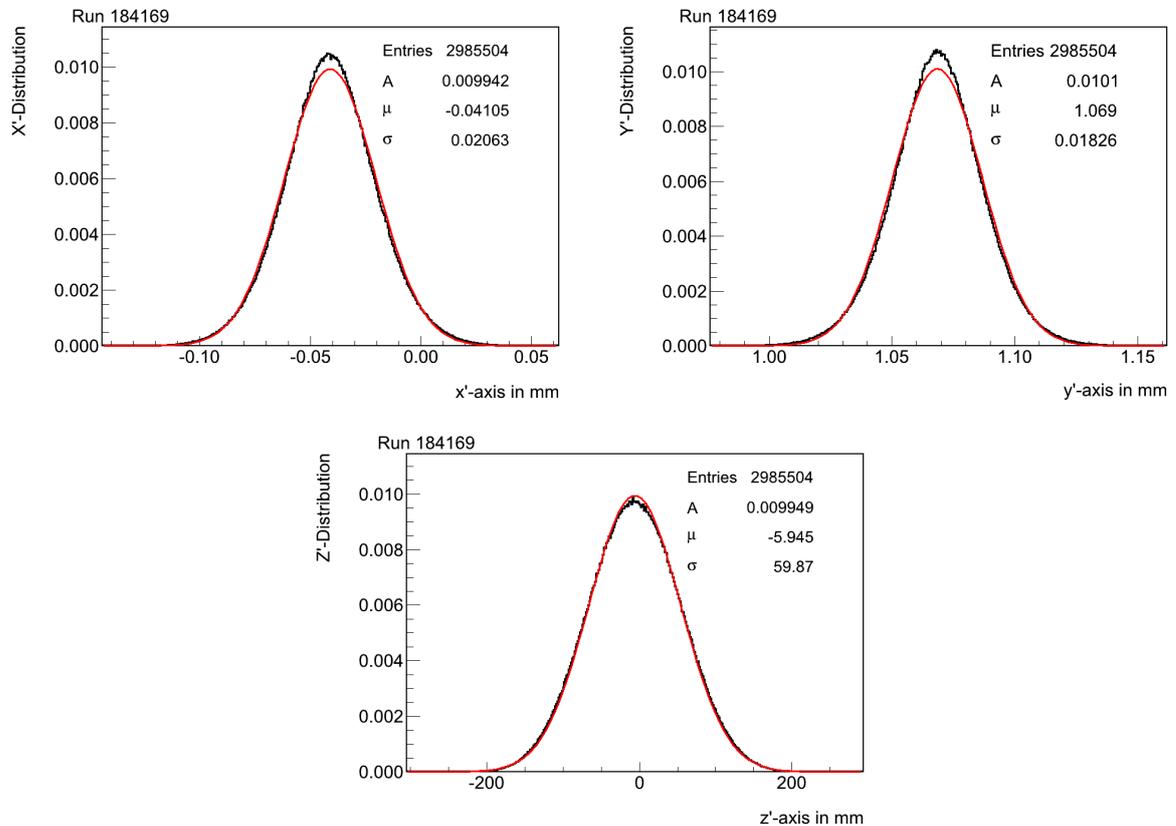


Figure 17: Observed vertex distribution in the adjusted coordinate system

values for the beamwidth in the transverse directions are smaller and the value in the longitudinal direction has not changed significantly. Figure 18 summarises the gained correction in percentage of the beamwidth measured in the ATLAS coordinate system. In x-direction the average change over all 23 runs is $-0.036 \pm 0.095\%$. The error is the sum of the standard deviation of the mean, which is only 0.017% and the propagated error for the beamwidth fits that is 0.78%, which exceeds the value of 0.036%. Thus the decreasing of the beamwidth in x-direction is negligibly

4.4 Vertex distribution in the rotated coordinate system

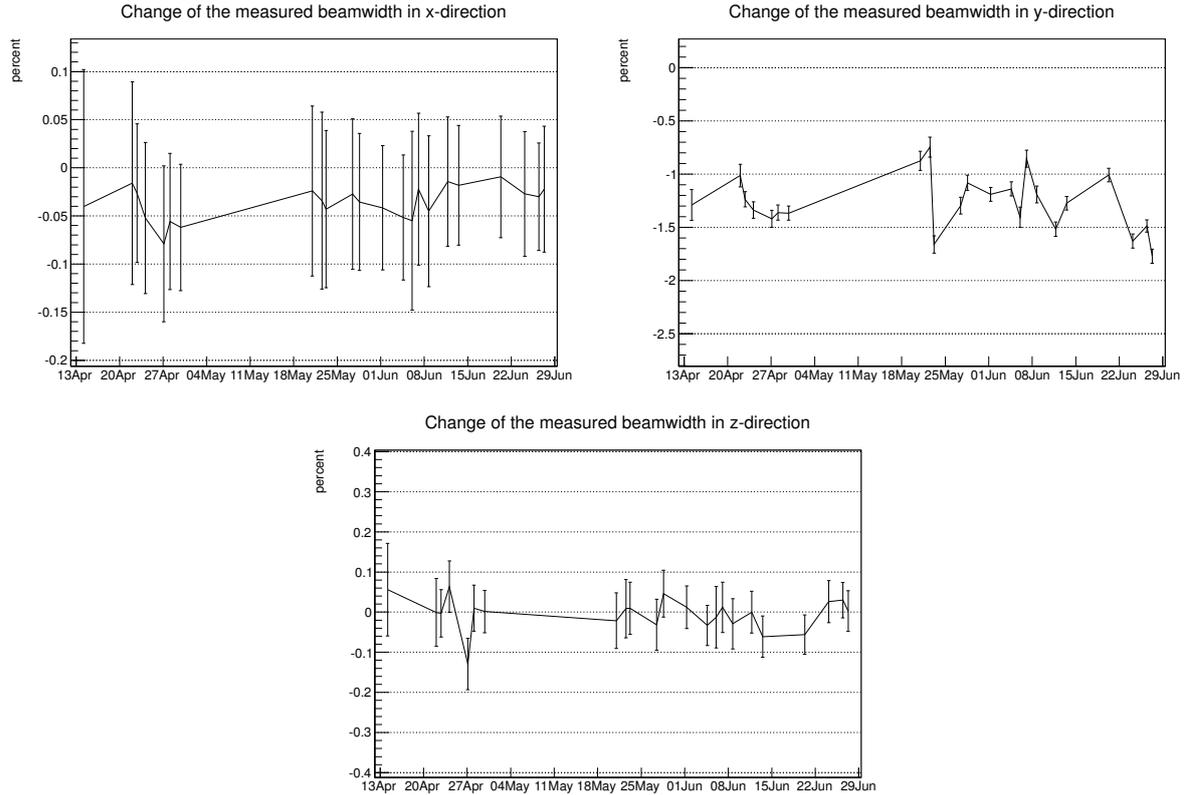


Figure 18: Change of the measured beamwidth after rotation in percent of the beamwidth measured in the ATLAS coordinate system.

small and cannot be confirmed within the fit precision.

In y-direction the effect is stronger and its negative sign can be confirmed within the propagated error interval by $-1.27\% \pm 0.33\%$. In z-direction, the propagated error has an average of 0.062% , which exceeds the change of the measured beamwidth for every run. The sign of the beamwidth change alternates and the values have order of $O(10^{-2} - 10^{-4})$. The found average is even found negative, but within the errors without significance. With an average of $-0.004 \pm 0.1\%$ the effect of the coordinate transformation in z-direction has to be seen as approximately zero. Table 7 compares the measured beamwidth in all three spatial directions before and after the rotation of the coordinate system.

Figure 19 shows the effect of the coordinate transformation on the computation of the area of the transverse beam projection. Since both ellipse axes in the transverse plane are measured as smaller after the rotation, the average size of the ellipse of $A = 4\pi\sigma_x\sigma_y$ is smaller for all runs. In the ATLAS coordinate system it is measured to be $(5.08 \pm 0.3) \cdot 10^{-3} \text{ mm}^2$, in the new system K' it is measured to be $(5.01 \pm 0.3) \cdot 10^{-3} \text{ mm}^2$. The effect is low with an average correction of $-1.3 \pm 0.26\%$ which basically originates from the correction of the tilt in the y-direction. This correction is larger than the precision intervals of the ellipse areas for the single runs by a factor 10-30. The average value carries the standard deviation of the mean calculation which dominates over the applied correction.

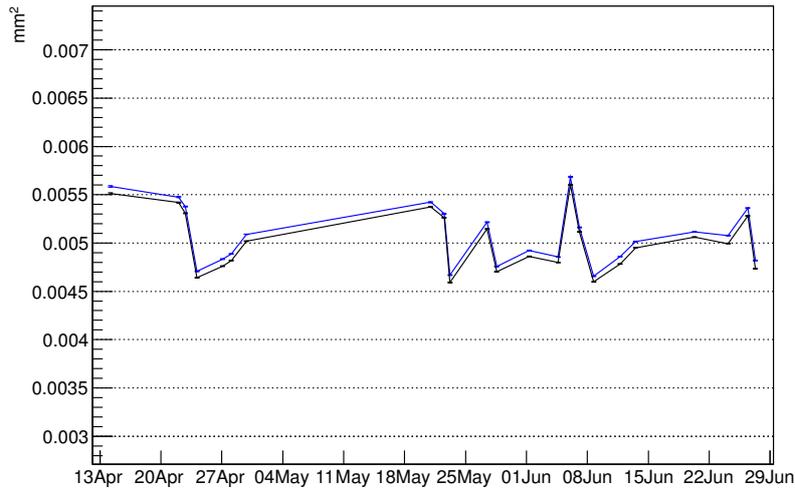


Figure 19: Area of the transverse beam projection with beamwidth 2σ in both transverse directions. The black line shows the area after the rotation, which is smaller for all evaluated runs with an average of $-1.3 \pm 0.26\%$.

Table 7: Beamwidth σ for the primary vertex distributions. Comparison of the values in the rotated coordinate system and in the ATLAS coordinate system. Average over all 23 runs.

	average σ ATLAS	average σ in K'	average σ fit error in both systems
x in μm	21.06 ± 0.7	21.05 ± 0.7	0.06%
y in μm	19.2 ± 0.6	18.9 ± 0.6	0.06%
z in mm	56.27 ± 1.8	56.27 ± 1.8	0.05%

5. Vertex resolution

The beamwidths discussed in section 4 were calculated by using only the recorded vertex positions while ignoring the resolution of the fitted vertices. But each vertex position is a result of the tracking process and carries an intrinsic uncertainty in all three spatial coordinates. Keeping this in mind, a vertex cannot be seen as a sharp resolved point, but as a centroid with given confidence intervals in all three directions. The picture of the luminosity region, drawn from the results of section 4, contains significant blurring due to the resolution of each single vertex. In this section the distribution of this blurring in the luminosity region is analysed and methods to correct the beamwidth from the effect of the vertex resolution are compared.

In the analysed data the vertex resolution is given for each vertex and each coordinate as the covariance computed from the track perigees that are assigned to the vertex. Following equation 28 this is related to a standard deviation assigned to each single vertex.

$$\sigma_{res_x} = \sqrt{Var(x_{trk})} = \sqrt{Cov(x_{trk}, x_{trk})} \quad (28)$$

Throughout this section, with a view to providing better comparability with the scale of the luminosity region, the term ‘‘vertex resolution’’ always refers to the standard deviation σ_{res} . Although the terminology could be confusing at some point, the resolution is referred to as ‘‘lower’’ in case of higher values of σ_{res} and as ‘‘better’’ or ‘‘higher’’ in case of lower values of σ_{res} . When calculating averages for this standard resolution, throughout all following sections this will be calculated as

$$\bar{\sigma}_{res} = \sqrt{\bar{Var}} = \sqrt{\frac{1}{N} \sum_{i=1}^N \sigma_i^2} \quad (29)$$

When speaking of ‘‘correlation’’ between different coordinates or between the resolution and the observed beamwidth throughout this section, this will always refer to the Pearson correlation defined by

$$corr(X, Y) = \frac{cov(X, Y)}{\sigma_x \sigma_y} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sigma_x \sigma_y} \quad (30)$$

The correlation factor is +1 in the case of total correlation and -1 in the case of total anticorrelation. A correlation factor of $corr(X, Y) = 0$ does not prove the independence of X and Y . Additionally the Spearman ρ -coefficient is used as a second indicator. This coefficient disregards the values of the single data entries and only correlates the rank within the sorted sample.

5.1. Dimension of the vertex resolution

In this subsection the distribution of the absolute values of the vertex resolution is evaluated. They are far from being on a neglectable scale compared with the observed beam width in the transverse plane. The dimension of the longitudinal resolution is dominant in the ellipsoid of each vertex, but is still of the same order as the dimension of the transverse resolution. While the size of the luminosity region is of order $O(10^3)$ larger longitudinal than transverse, this makes the relative resolution $\sigma_{res,z}/\sigma_{obs,z}$ of order $O(10^{-3})$ smaller than $\sigma_{res,x,y}/\sigma_{obs,x,y}$.

5.1.1. Single run evaluation

The analysis of run 184169 shows that the average dimensions of the confidence intervals in the transverse directions are $Mean(\sigma_{res,y}) = 16.1 \pm 3.9 \mu\text{m}$ and $Mean(\sigma_{res,x}) = 15.5 \pm 3.6 \mu\text{m}$. The average size of the transverse projections of the confidence ellipsoids thus almost covers a range of one sigma of the beamwidth. Figure 20 is shown to illustrate the relative size of the confidence ellipsoids on the scale of the measured beamwidth, since this illustrates the relative resolution better than the variance.

But since the means above are calculated as $Mean(\sigma_{res}) = \frac{1}{N} \sum_{i=1}^N \sigma_i$ they will not be used as the average resolution $\bar{\sigma}_{res}$ in the following analysis. This value is instead computed from the average variance according to equation 29.

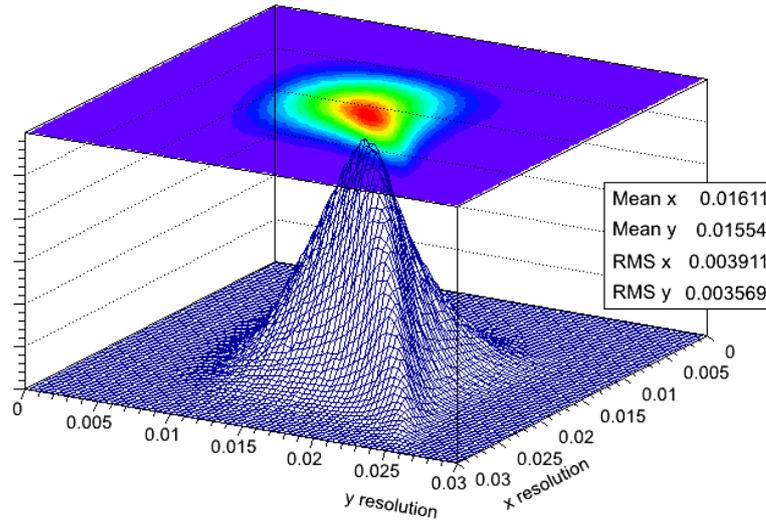


Figure 20: Distribution of the absolute value of the vertex resolution in the transverse directions. The plot shows the histogram with a view towards the origin. The ridge of the vertices with a good approximation of $\sigma_{res,x} \approx \sigma_{res,y}$ is clearly visible.

To stay on the scale of the beamwidth, the square roots of the average variances are given here as $\bar{\sigma}_{res,x} = 16.6 \pm 0.4 \mu\text{m}$ and $\bar{\sigma}_{res,y} = 15.9 \pm 0.3 \mu\text{m}$. The correlation factor between the both transverse variances is +0.37. Compared with the average beamwidths $\sigma_{x,y}$ -ranges given in table 7, the average resolution amounts to as much as $78.9 \pm 0.2\%$ in x-direction and $82.0 \pm 0.2\%$ in y-direction. The distribution in figure 20 shows a clearly visible ridge along the diagonal that corresponds to $\sigma_{res,x} \approx \sigma_{res,y}$ with a slight trend for lower values in $\sigma_{res,y}$.

Figure 21 shows the vertex resolution distribution along the three axes. It confirms that the resolution distribution in both transverse directions is approximately the same. The distribution along the z-axis can be plotted on a micrometer scale, too. The means of these distributions are not shown in figure 21, because the average resolution is computed using equation 29.

In figure 22, the relation between $\sigma_{res,y}$ and $\sigma_{res,x}$ is analysed with a histogram that shows the ratio between the vertical and the horizontal resolution. This analysis shows that the two values can be expected to be found with $\sigma_{res,y} \approx \sigma_{res,x}$ on average, but a broader range of transverse profiles of the resolution ellipsoid can be found in the distribution. The y-resolution can amount to less than half or more than double the x-resolution, although these are rare events. The median of this distribution is 0.96, the maximum is reached at 0.95. The mean is not sufficient to describe the peak, since the range bound at zero from below but has no upper bound. Both means, the one of $\sigma_{res,y}/\sigma_{res,x}$ and the one of $\sigma_{res,x}/\sigma_{res,y}$ evaluate to more than +1. The median of course divides the set in half and evaluates to 0.961 and 1.039 respectively. There is a clear trend towards better resolution in y-direction. The histogram shape of the distribution in figure 22 can be found along the slices of the x-axis bins in the resolution distribution in figure 20 and explains that the ridge in that distribution follows the pathway of $y \approx 0.96 \cdot x$.

5.1 Dimension of the vertex resolution

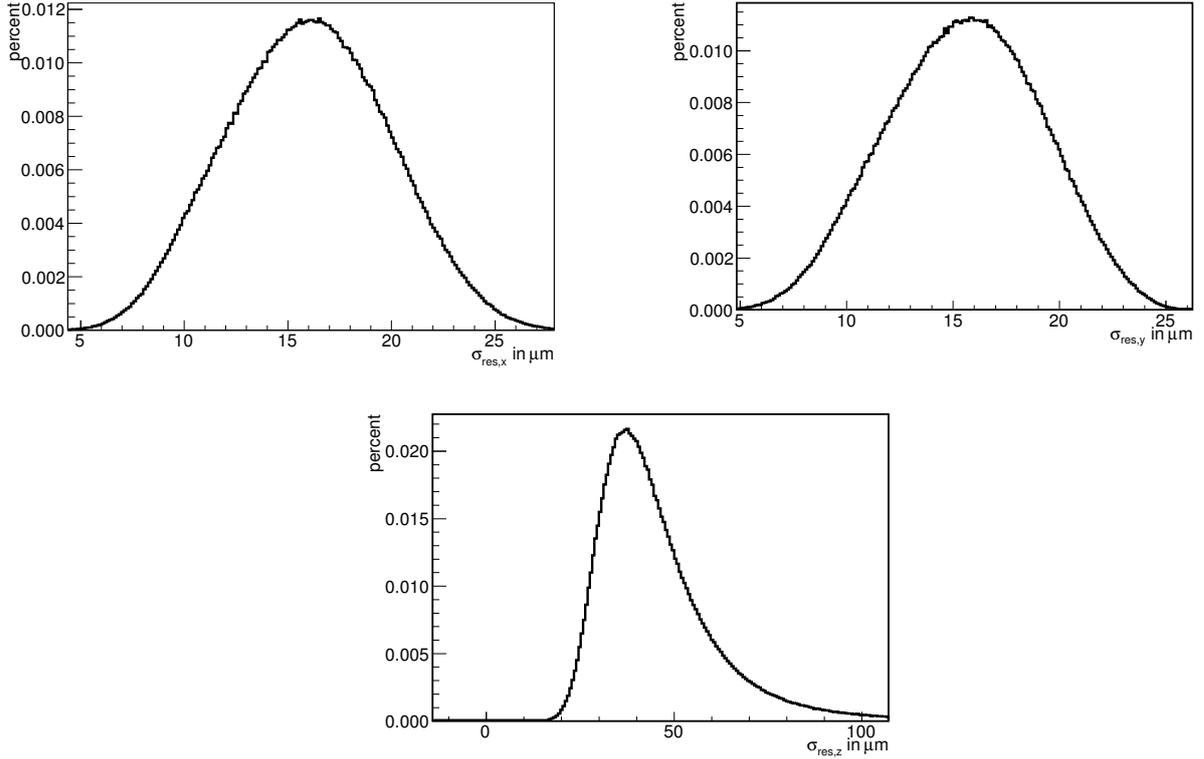


Figure 21: The distribution of the vertex resolution of run 184169 for the three spatial directions. The transverse distributions don't have significant differences. The longitudinal distribution can be plotted on a micrometer scale.

Because its distribution is close to symmetry, the transverse resolution can be reviewed versus the longitudinal resolution as one conflated parameter, computed by equation 31. This avoids doubling the discussion at this stage by diversifying it into both transverse directions.

$$\sigma_{res,x} \cdot \sigma_{res,y} = \sigma_{res,x,y} = \sqrt{Var_{res,x} \cdot Var_{res,y}} \quad (31)$$

This combination of the both transverse resolutions is the radius of a circular region around the vertex that represents the confidence region of the vertex fit in the transverse plane in micrometers.

Figure 23 shows the distribution of the absolute value of the vertex resolution by comparing its longitudinal and transverse dimensions. But analogue to the analysis in the transverse plane, the mean values differ from the expectation value for the standard deviation, which is computed from the square root of the average variance again. The correlation factor between both axes has a value of +0.78. A clear-cut ridge shows the strong dependence of the transverse resolution on the resolution along the beam axis. The resolution in the transverse directions can be expected to be better for vertices with a good resolution along the beam axis and vice versa. The average resolution in z-direction is $43.5 \pm 11.2 \mu\text{m}$. This is a factor 3 higher than the average transverse resolution $\overline{\sigma_{res,x,y}} = \sqrt{\overline{\sigma_{res,x}} \cdot \overline{\sigma_{res,yx}}}$ which is $15.5 \pm 3 \mu\text{m}$. Its effect on the observed longitudinal

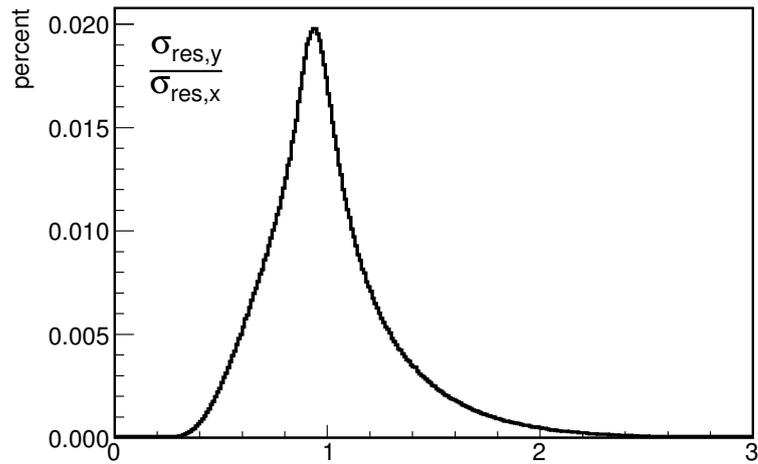


Figure 22: The distribution of the ratio of the vertical and horizontal resolution has its median at 0.96. The maximum is reached at 0.95. Only few vertices have a resolution that is higher in one direction by a factor two than in the other direction.

size of the luminosity region is neglectable, since it amounts to only about $0.8 \cdot 10^{-4}\%$ of the beamwidth presented in table 7. This makes an attempt to correct the longitudinal beamwidth from the vertex resolution dispensable.

Figure 24 shows the profile of the transverse versus longitudinal resolution. The bin-wise means are following the path of the ridge observable in the plot in figure 23. For values lower than the mean in z-direction, the dependence is stronger and even almost linear, for values higher than the mean the slope decreases. This graph indicates that the determination of the transverse beamwidth gains precision when the analysis is limited to vertices which have a good resolution along the z-axis.

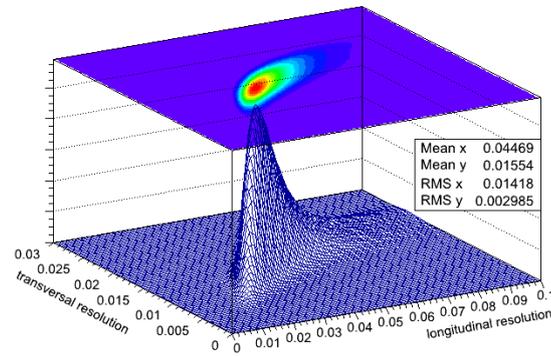


Figure 23: Transverse vs. longitudinal vertex resolution. The distribution has a sharp peak and a clear-cut ridge. The transverse resolution is a strong function of the longitudinal resolution. The plot shows the standard deviation at the scale of the luminosity region, the means differ from the average resolution that is computed using equation 29.

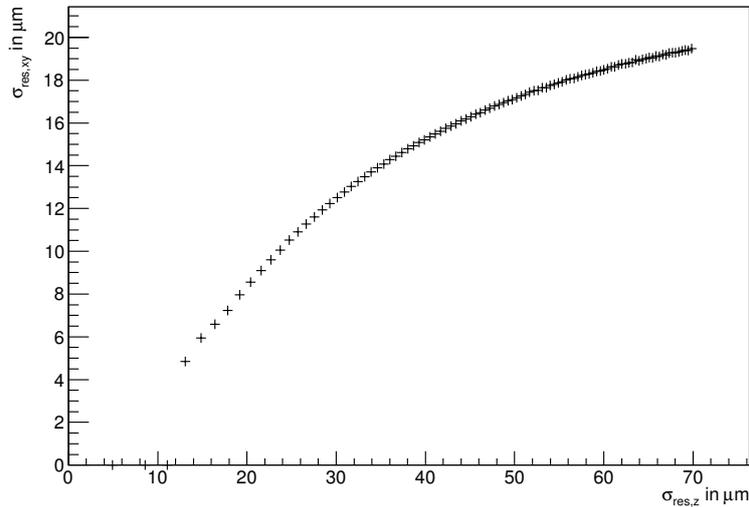


Figure 24: The transverse resolution depends on the longitudinal resolution with which it is monotonically increasing, but with decreasing slope. The bins above $\sigma_{res,z}=100 \mu\text{m}$ contain only 1.8% of the events. Shown is the square root of the bin-wise average variance $\overline{\sigma_{res,x,y}}$.

5.1.2. Evaluation for the set of investigated runs

The findings of the evaluations of the other runs are very consistent with the evaluation of run 184169. They are collected and analysed in the following timeline plots with a view on the correlation between the vertex resolution and the observed transverse beam projection. Through the run sequence there is no clear trend indicating a systematic change with the beam parameters. However, while the beamwidth changes within a small range, the resolution is rather expected to show a more stable development, since the tracking algorithm remains the same and operates on single events which can be expected to be distributed according to comparable rates of event types in all examined runs.

A dependence of the observed beamwidth on the average resolution is obvious, based on the convolution discussed in section 2.3.4. In turn a dependence of the average resolution could be based on the start parameters of the tracking process which involves the average beamspot and beamwidth computed online during ATLAS operations. As described in [16], the HLT is updated several times per run, triggered by either a change of the observed beamwidth of 10% or more, a change of the beamspot position of 10% of the beamwidth or by a reduction in any of the errors of 50% or more.

Figure 25 shows the observed beamwidth (one σ) and the resolution for all three axes. As expected, the resolution moves in a smaller range than the beamwidth. Both series have a strong correlation in the transverse plane, but a lower correlation in the longitudinal direction. The values are listed in table 8.

The average resolution calculated for all runs amounts to $79.3 \pm 1.5\%$ in x-direction and $85.2 \pm 1.7\%$ in y-direction compared to the average beamwidth 1σ . In z-direction the resolution is less than $10^{-3}\%$ of the length of the luminosity region and thus it is not relevant for the beamwidth determination. The observed correlation can be assumed to be based on the feedback process of the HLT, since the small variations in the resolution would be expected to cause only small changes of the observed beamwidth. Thus the more probable relation is the mapping of the online determined beamwidth into the resolution set.

Table 8: Average resolution, beamwidth and ranges defined by $\sigma_{max} - \sigma_{min}$. The standard deviation is included in the errors. The propagated fit errors of σ_{obs} are included as well, while the average resolution per run is the calculated mean taken from the data sets.

	unit	σ_{res} -range	σ_{obs} -range	average $\bar{\sigma}_{res}$	$\bar{\sigma}_{obs}$	$corr(\sigma_{res}, \sigma_{obs})$	Spearman ρ
x	μm	0.89	3.1 ± 0.02	16.7 ± 0.2	21.05 ± 0.7	0.91	0.95
y	μm	0.79	2.3 ± 0.02	16.1 ± 0.2	18.9 ± 0.6	0.95	0.97
z	mm	0.56	7.0 ± 0.05	0.0435 ± 0.0004	56.27 ± 1.8	0.22	0.18

Figure 26 shows the results discussed for figures 20 and 22, where a trend towards better resolution in y-direction than in x-direction could be observed. This follows the same pattern in all 23 runs where both the size of the luminosity region and the size of the resolution ellipsoid

5.1 Dimension of the vertex resolution

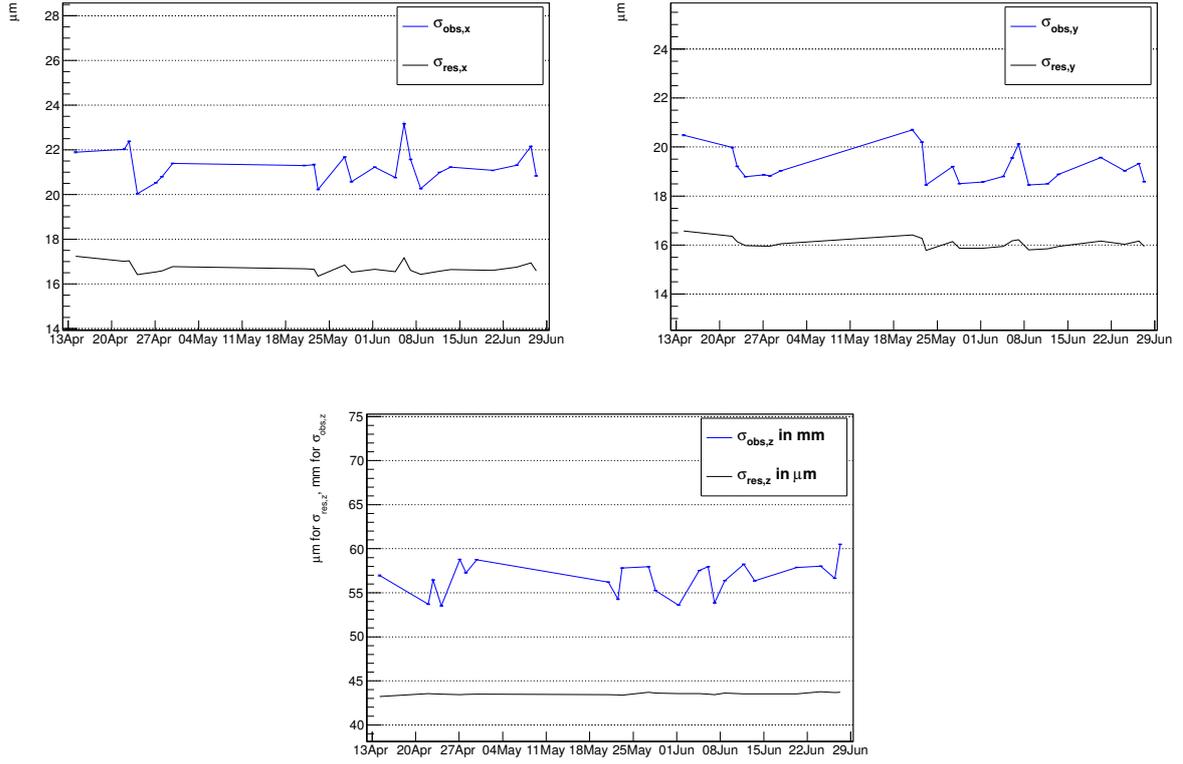


Figure 25: Change of the measured beamwidth after rotation in percent of the beamwidth measured in the ATLAS coordinate system.

are smaller in the vertical plane than in the horizontal plane. The correlation factor between the two series is 0.95 and the correlation is observable at first sight on the plot. The calculation of the Spearman ρ -coefficient evaluates to 0.92. The ratio $\sigma_{obs,y}/\sigma_{obs,x}$ has an average of $96.2 \pm 0.3\%$, while the average resolution in y-direction is evaluated to be $\sigma_{res,y} = (0.904 \pm 0.01) \cdot \sigma_{res,x}$. The difference of the beamwidths in both directions is explainable due to the beam focusing, as described in the discussion of the form of the phase-space ellipse in section 2.2.2. The HLT could be expected to treat both directions as basically equal. But the clear trend towards better resolution in the vertical direction indicates that the estimation of the perigees is influenced by the information about the beamwidth. The resolution is calculated from the perigees according to equation 28. The analysis in this section focused only on the absolute values of the resolution, the next section reviews how the resolution quality distributes within the luminosity region.

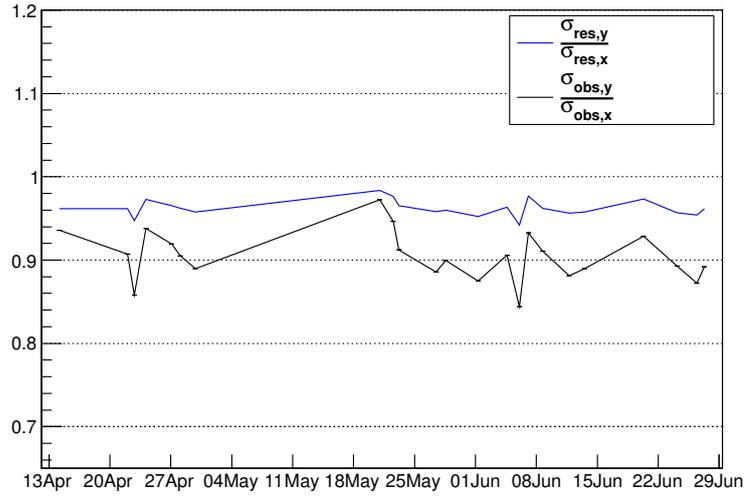


Figure 26: Comparison of the transverse beamwidth ratio $\sigma_{obs,y}/\sigma_{obs,x}$ with the median of $\sigma_{res,y}/\sigma_{res,x}$. The blue line shows the median analogue to figure 22, the black line shows the ratio of the beamwidth results from section 4. A strong correlation is observable.

5.2. Spatial distribution of the vertex resolution

In the following analysis the findings of the last section are brought into place, which is the luminosity region for the vertex ensemble and the particular vertex position for the single event. It is shown that the average transverse resolution, calculated for area intervals in the 2-dimensional projections, is not uniformly distributed. In the transverse plane the probability of finding a vertex with a lower resolution is higher close to the beamspot than in the outer zones of the interaction region. Along the beamline, the distribution of the longitudinal resolution is close to uniform distribution and does not show similar effects. Because the longitudinal resolution is neglectable in the attempt to correct the beamwidth from the intrinsic vertex resolution, it is not analysed in detail here. The focus is set on the transverse resolution.

To evaluate the distribution of the average local resolution, the events of run 184169 are used to fill two 2D-histograms. One is filled with weights $w_{event} = 1$, so that it counts the events that are sorted into each bin respectively into each area interval $A_{i,j}$. The same binning is applied to both histograms with equidistant bin ranges along both axes, so that $|A_{i,j}| = \Delta x \cdot \Delta y = const.$ The second histogram is filled with the same events with weights $w_{event} = \sigma_{res,event}^2$. That way a bin with $N_{i,j}$ entries holds the bin content of $\sum_{k=0}^{N_{i,j}} \sigma_{res,k}^2$. Dividing this histogram bin-wise by the first histogram gives the average variance in the area $A_{i,j}$. With a view on the comparability with the scale of the luminosity region the square root of the average variance is calculated and set as the bin content to show the average local resolution. Some bins in the edges of the coordinate window have no entries and the division evaluates to NaN, in which case the bin content is set to zero.

Figure 27 shows the distribution of the average local resolution in the transverse projection

divided into two plots which show the distribution of the local average transverse resolution and the local average longitudinal resolution. Since it is indicated from the trend shown in figure 24 that the transverse resolution depends on the longitudinal resolution, the similarity of both plots is expected. Both distributions show that the average local resolution is lower close to the

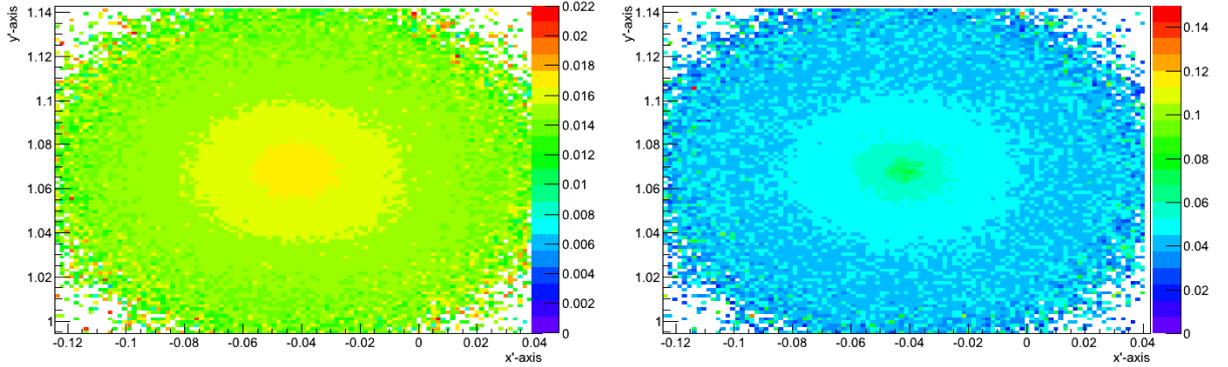


Figure 27: The left plot shows the average local transverse resolution $\bar{\sigma}_{res,x,y} = \sqrt[4]{\bar{Var}_x \cdot \bar{Var}_y}$, the right plot shows the average local longitudinal resolution $\bar{\sigma}_{res,z}$, both calculated as averages within the bin. The distributions are symmetrical to the beamline, the mean evaluates to the same coordinates as in the unweighted distribution. The average resolution in the mean region is lower than in the outer regions. The axes are set to $\bar{x}, \bar{y} \pm 4\sigma_{x,y}$.

beamline and systematically increases with distance. Outliers in the region further afar than 3σ tend to have a lower resolution and bring discontinuity into the mix. Both plots are symmetrical in the transverse directions. It is observed that with a standard deviation of $\bar{\sigma}_{res,x,y} \approx 17 \mu\text{m}$ the local transverse resolution in the area around the beamspot is noticeably lower than in the zones with a distance from the beamline of 2-3 times $\sigma_{obs,x,y}$, where it is $\bar{\sigma}_{res,x,y} \approx 13 \mu\text{m}$. Similarly, the average local longitudinal resolution amounts to about $60 \mu\text{m}$ in the region close to the beamline and $40\text{-}50 \mu\text{m}$ in a distance of $2\text{-}3 \sigma_{x,y}$.

Figure 27 shows that the symmetry found in the distribution of the local average transverse resolution is the result of a superposition of the two orthogonal symmetric distributions in the particular transverse directions. The observed effect underlying this superposition is the increase of the resolution with distance from the beamline along the given axis. Figure 28 shows the distribution of the average local resolution for both transverse axes in the transverse plane. Figure 29 shows this effect along the length of the interaction region. There is no noticeable change along the beamline.

Looking at the plots in figures 28 and 29, the visual effect is stronger in the y-direction, where both the size of the luminosity region and the average standard deviation based on the resolution are smaller than in x-direction. To examine the amount of the variation for both transverse axes, this is examined further by the calculation of the average local resolutions $\bar{\sigma}_{res,x}$ as a function of the x-coordinate and $\bar{\sigma}_{res,y}$ as a function of the y-coordinate. The result of this analysis is shown in the profile plots in figure 30.

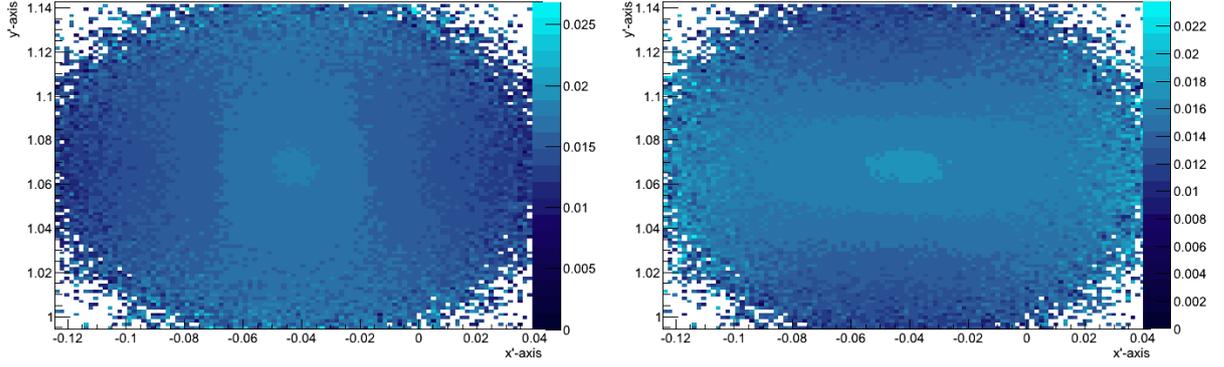


Figure 28: The left plot shows the average local transverse resolution $\bar{\sigma}_{res,x}$, the right plot shows the average local transverse resolution $\bar{\sigma}_{res,y}$, both calculated as averages within the bin. The symmetry with the beamspot observable in figure 27 is replaced by a symmetry along the particular axis.

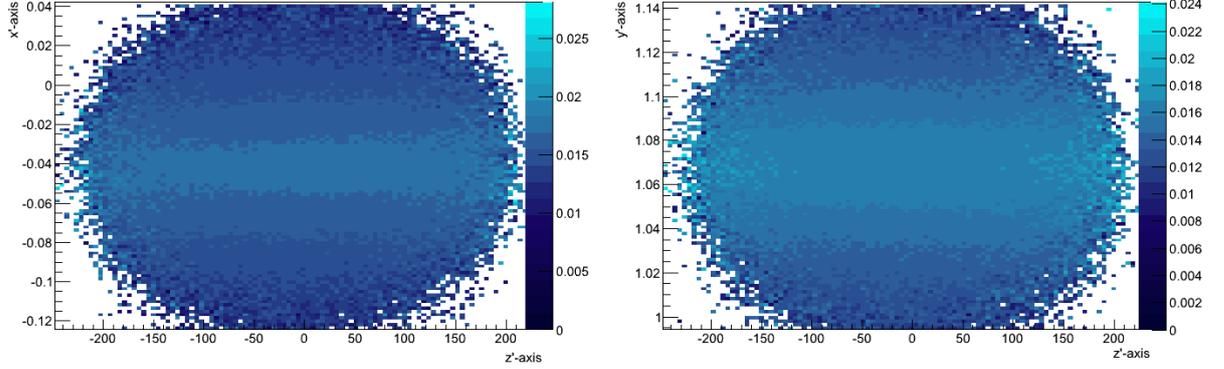


Figure 29: The left plot shows the average transverse resolution $\bar{\sigma}_{res,x}$ in the x-z-plane, the right plot shows the average local transverse resolution $\bar{\sigma}_{res,y}$ in the y-z-plane, both calculated as averages within the bin. The symmetry with the particular transverse direction observable in figure 28 is observed along the whole longitudinal range of the luminosity zone.

In both transverse directions the local average resolution varies over a range of about $4 \mu\text{m}$, symmetric with respect to the beamline. The dependence is in good approximation linear except from the turning area, where the positive slope turns into the negative slope. The found variation can be characterised descriptively, by dividing the $4 \mu\text{m}$ difference on the resolution axis by the $4\sigma_{obs}$ range on the coordinate axis. This leads to a change of the local transverse resolution of approximately one micrometer per one sigma.

But the parameter gained with more precision is the slope of the linear fit functions. It gives an opportunity to follow the effect through the examined run sequence. The same behaviour of the x-resolution along the x-axis and the y-resolution along the y-axis is observed for all runs, with profiles all analogue to figure 30. These profiles are fitted by linear functions in the ranges $0.5\sigma_{obs} \leq |x - \bar{x}| \leq 3\sigma_{obs}$ and the slope is calculated as the mean of the absolute values,

5.2 Spatial distribution of the vertex resolution

to conflate the two values taken from the left and right fit. The result of this computation is shown in figure 31.

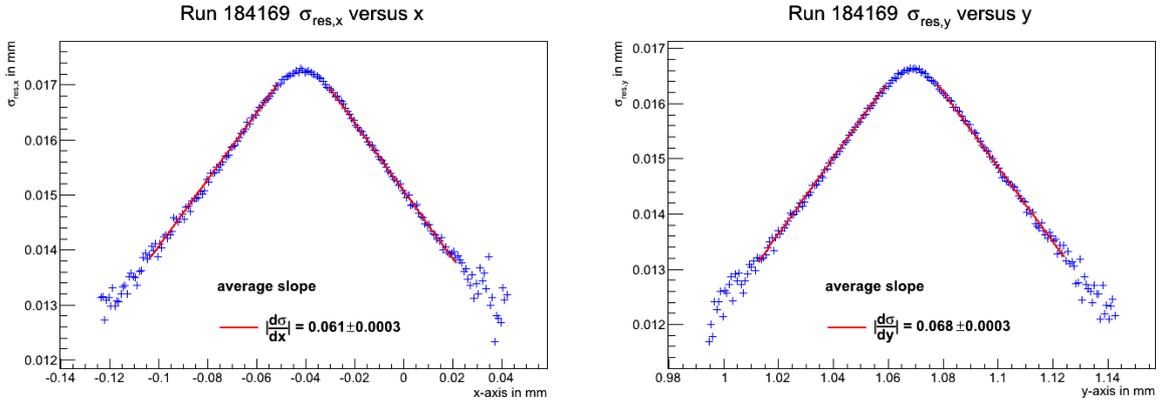


Figure 30: Average local resolution $\bar{\sigma}_{res,x}$ along the x-axis and $\bar{\sigma}_{res,y}$ along the y-axis. The linear fits cover the ranges $0.5\bar{\sigma}_{obs} \leq |x - \bar{x}| \leq 3\bar{\sigma}_{obs}$.

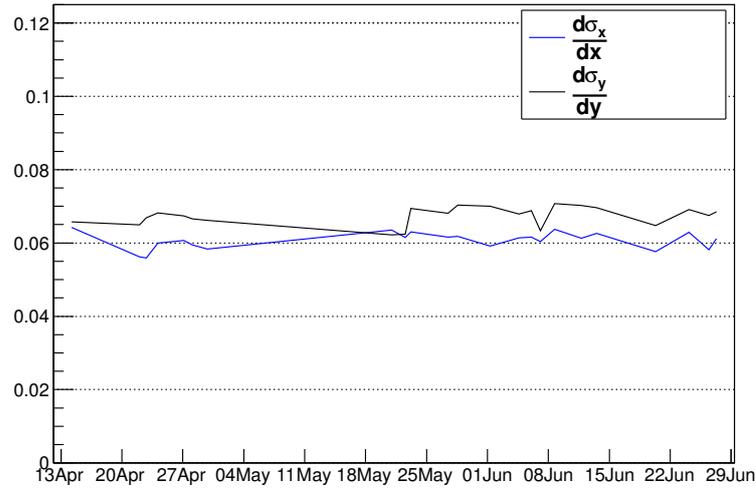


Figure 31: Range of the local transverse resolution for the examined run sequence. The effect is observed in all runs. The correlation between the two sequences is +0.23.

The two sequences are not correlated strongly, but an anticorrelation with the observed beamwidth can be stated with $corr(d\bar{\sigma}_{res,x}/x, x) = -0.31$ and $corr(d\bar{\sigma}_{res,y}/y, y) = -0.86$ which suggests that the effect is stronger for runs with lower transverse emittance. Because of this anticorrelation the slope of the linear trend of lower resolution with less distance to the beamspot gets larger with a smaller beamwidth, thus the effect of the beamwidth on the range of the resolution decrease gets somewhat compensated. Of course the found slopes can be multiplied with any range along the transverse axis in order to find the increase of the average local transverse resolution. The average slope in x-direction for the sequence of the examined runs is 0.061, the

slope for the y-direction is 0.067. Along both axes this gets an increase of the resolution of 1.3 micrometer per sigma.

In longitudinal direction there is no comparable effect. Figure 32 illustrates that the longitudinal resolution is flat along the z-axis until it reaches the 3-4 σ_z -zone, where it spreads out and scatters strongly because of statistical fluctuations. Of course the effect on the data sample is low, since it is expected to affect less than 0.27% of the events.

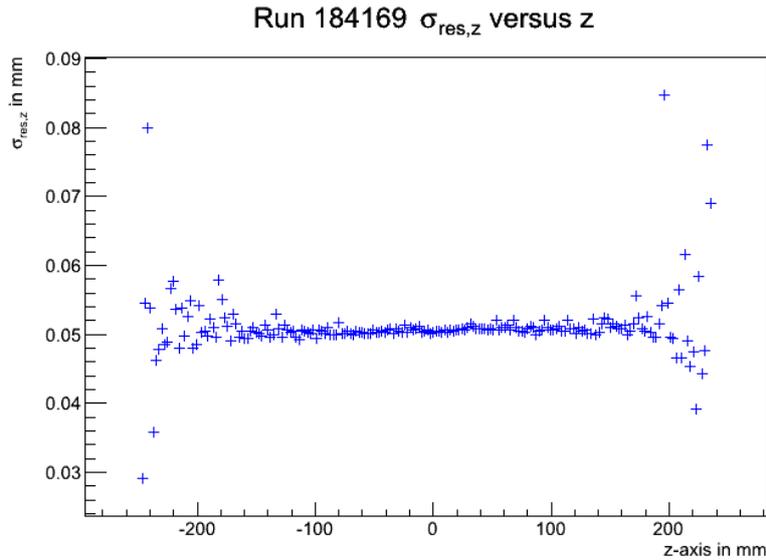


Figure 32: Local longitudinal resolution variation along the z-axis. In longitudinal direction there is no comparable effect. The resolution is uniformly distributed. In the outer zones of the interaction there are outliers with lower resolution.

The reason for the higher probability that a vertex with lower resolution is located close to the mean than in some distance cannot be definitely determined here. Since the plots show the resolution dependent on the location, both parameters could influence the distribution. A look at the spatial distributions in figure 17 recalls that the observed distribution varies from the suggested normal distribution around the mean. Together with the findings in this section, it is imaginable that the tracking algorithm weighs the information about the position of the beamspot higher in relation to the perigees, when the resolution is lower and thus such vertices are shifted towards the beamline.

5.3. Vertex resolution as a function of the number of tracks

The vertex resolution has been examined with a view on the absolute values in section 5.1 and with a view on the spatial structures in section 5.2. In this section the resolution is reviewed as a function of the track multiplicity and the meaning of the findings for the observed vertex distribution is discussed. The resolution is not distributed uniformly along the track multiplicity axis, but is clearly a function of the number of tracks that are fitted into the vertex.

The distribution of the track multiplicity has a Poissonian shape and is very consistent throughout all examined runs. In figure 33 this distribution is shown for run number 184169. The bin range covering the RMS interval holds 70.1% of the events, 13.8% have fewer than 23 tracks and 16.1% have at least 63 tracks. A small number of 223 vertices have even more than 150 tracks and are sorted into the overflow bin of the histogram.

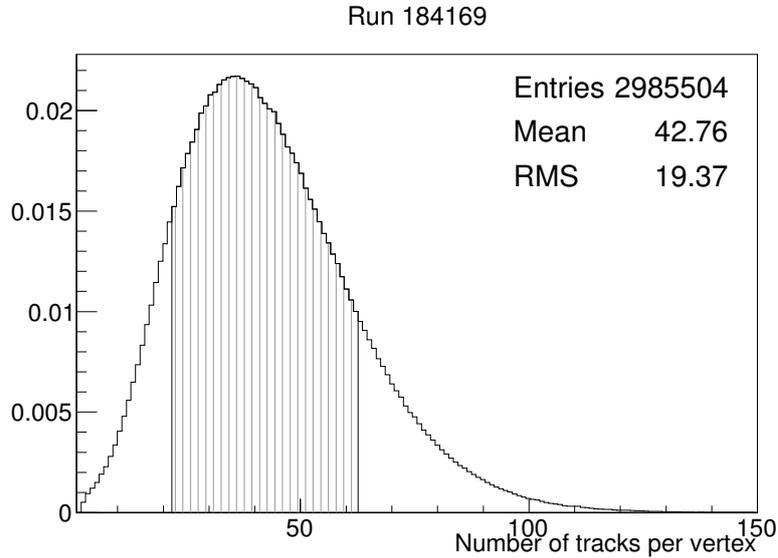


Figure 33: Distribution of the track multiplicity. The filled area represents the bins covering the RMS-range.

In figure 34 the resolution is shown as the bin-wise computed square root of the average variance and the beamwidth is shown as the bin-wise standard deviation of the vertex distribution. The pathway of the resolution is monotone decreasing and slowly approximating a lower boundary. This is observed in all examined runs. Seen from the other way, one can observe that following the track multiplicity axis from the distribution mean of 42 tracks in the direction of a descending number of tracks, not only does the resolution get lower, but the vertices also appear to close ranks, showing a descending beamwidth. This is consistent with the findings of section 5.2 where it has been indicated that vertices with a lower resolution have a higher probability to be found close to the beamline. The distribution of the beamwidth along the track multiplicity axis reaches a plateau in the range of 70 to 100 tracks, but this is located in the tail of the track distribution, where the bins contain less than 20% of the vertices.

It appears reasonable that the resolution gets lower with less tracks and thus less information about the vertex. But the proton-proton collisions cannot be expected to systematically produce less tracks in the region close to the mean, thus the observation that the beamwidth decreases with fewer tracks confirms the indication that vertices with lower resolution are located closer to the beamline systematically. This suggests that applying a cut to the data set by disregarding vertices with a low number of tracks might flatten the distribution of the average local resolution shown in the figures in section 5.2 and even lead to a better agreement of the observed vertex

distribution and their Gaussian fits shown in figure 17. A brief look at distributions obtained with an improved vertex selection is presented in section 6.

In the upper domain of the track multiplicity axis it is observed in all runs that the positions

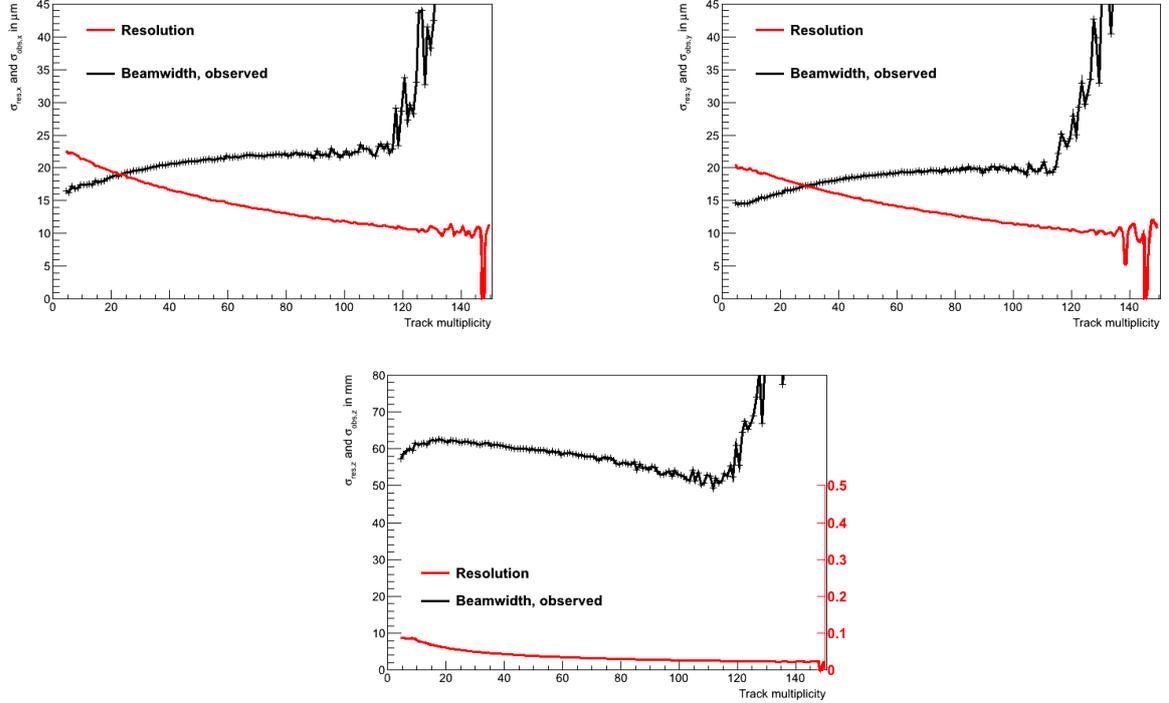


Figure 34: Profile plots for the observed beamwidth and the vertex resolution as functions of the number of tracks per vertex. Evaluation for run 184169. The plots for the x-direction and the y-direction have a μm -scale, the plot for the z-direction has a mm-scale.

of just the vertices that are assigned to a good resolution show noticeable deviations from the distribution mean, as shown in figure 35. This corresponds to the dominant visual impression in figure 34, where a considerable beamwidth increase occurs in the upper bins. But it has to be kept in mind that in this domain the statistics are getting worse as in the bins above 100 tracks there are altogether less than 1% of the events; in the bins above 120 tracks this decreases down to about 0.1% of events. However, besides the cut at a lower boundary, this suggests to determining the beamwidth by also applying a cut at 90-100 tracks. For run number 184169, this cut disregards about 2% of the events.

Figure 34 shows the same discontinuity for high numbers of fitted tracks in the bins with low statistics for the longitudinal coordinate, too. The observed length of the luminosity region is about 10 mm higher for bins with a low track multiplicity. But different from the transversal directions, it is observed that the function of the length of the luminosity region is basically monotonically decreasing with a higher track multiplicity. The functions in all three directions are found to follow the same pattern in all examined runs. This is likely to be caused on an effect brought into the process due to the reconstruction, since it is less probable, that the real beamwidth is a function of the number of tracks.

5.3 Vertex resolution as a function of the number of tracks

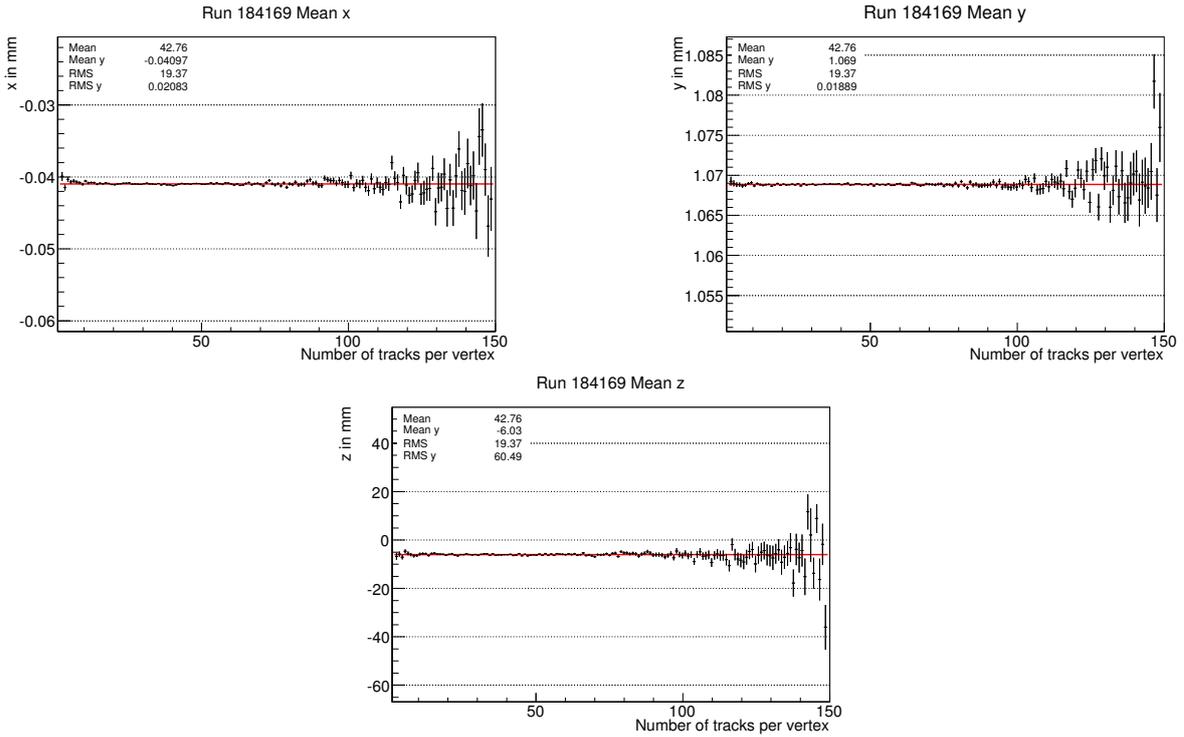


Figure 35: Trend of the bin-wise beamspot position vs. track multiplicity. The range on the y-axis represents the σ -range found by evaluating the whole data set, the red line represents the beamspot found in the same evaluation.

As an attempt to trace the vertices in the upper domain of the track multiplicity axis they are filled into a histogram of their own. The result, shown in figure 36, indicates, that the discontinuity discussed for the plots in figure 34 may be caused by unprecise fitting and mean finding due to the low statistics in the upper bins, as there is no broader beamwidth to be confirmed.

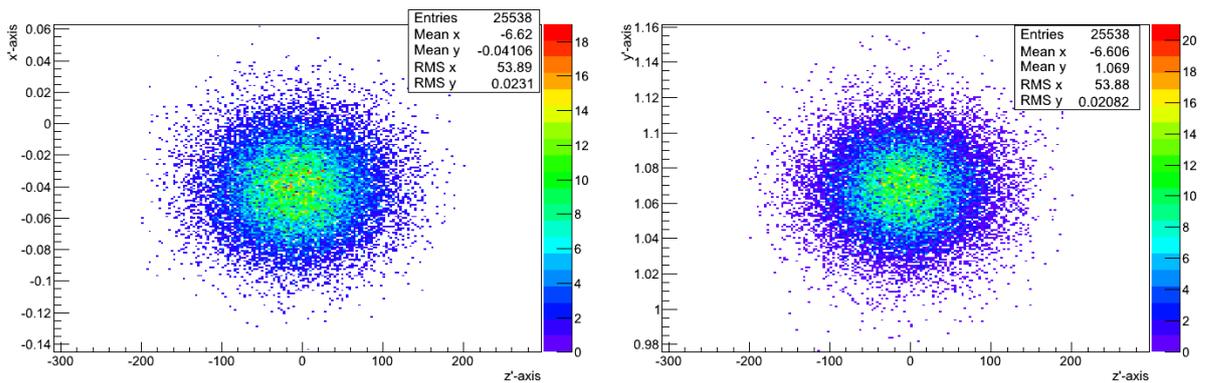


Figure 36: Histogram of the points with a track multiplicity of $N_{trks} \geq 100$

6. Vertex resolution correction

In this section the findings of section 4 and 5 are used to adapt methods to estimate the real size of the luminosity region. Of course, it is not possible to erase the limitations due to measurement precision from the data set entirely, thus a rest uncertainty will remain in any case, regardless of the strategy used. But the high relative size of the standard deviation brought into the process by the resolution suggests that a statistical estimation can achieve an improvement, since the real beamwidth appears to differ noticeably from the observed beamwidth. In section 6.1 a brief look at the idea of selecting qualified subsets of the vertex population is given. The method of subtracting the resolution quadratically from the observed beamwidth is addressed in section 6.2 and basic strategies of data denoising are applied to the problem in section 6.3.

6.1. Avoiding low resolution

A logical approach to improve the observation is to apply a sample design strategy with respect to minimizing the blurring that is brought into the analysis by vertices which are assumed to disturb the picture more than others. The most intuitive strategy is to avoid vertices that feature a low intrinsic resolution. As a consequence of the analysed spatial distribution of the resolution examined in section 5 the rather counter-intuitive expectation is that this cut leads to a broader observed beamwidth. In addition to this selection criterion the suggested cut in the lower domain of the track multiplicity axis should be dispensable, since it strongly correlates with the low resolution. The one in the higher domain of the track multiplicity axis is not expected to show a noticeable effect on the outcome, since it affects only a small number of events.

Thus the first series of evaluations presented here selects vertices such that the resolution in all three spatial dimensions is tested to be lower than the average resolution in that particular direction plus one times the RMS of the set of resolution values in that direction. All vertices that exceed this boundary in any of the components are rejected from the sample. For run 184169 this cut rejects vertices with $\sigma_{res,x} > 20 \mu\text{m}$ or $\sigma_{res,y} > 19.1 \mu\text{m}$ or $\sigma_{res,z} > 66.6 \mu\text{m}$.

The results agree with the expectations as demonstrated in figure 37. The distinctness varies from run to run, some still show a smaller deviation from a perfectly normal distribution some are really well fitted. The better agreement of the model and the observation after the selection indicates that in the sample including the low resolved vertices the underlying reason for the difference between both is likely not to be the model, but in that case the observation.

Of course a sample design strategy is an arguable issue, since it is always likely to include an unexpected bias, but in this case a bias is removed. This is examined here by looking at the distributions of a sample that selects by the track multiplicity. This second selection examined rejects all events with a track multiplicity outside of the RMS-range of the particular run. The result is shown in figure 38. The accumulation of events around the mean is slightly reduced, but the effect is significantly smaller than the one of the selection deciding by resolution. By just

6.1 Avoiding low resolution

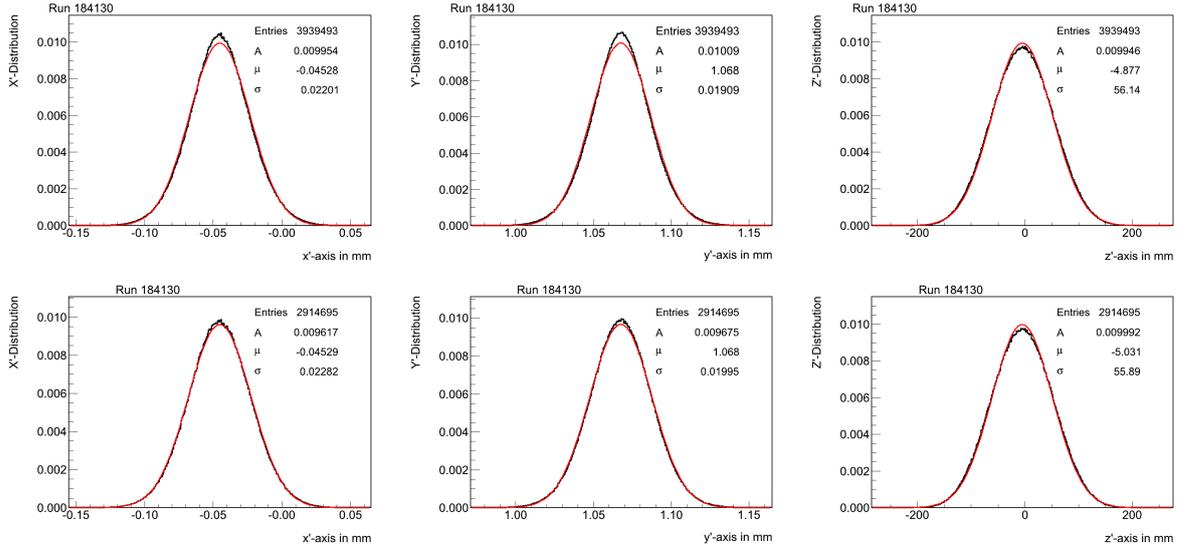


Figure 37: Distribution along the axes of the rotated coordinate system before and after rejecting low resolved vertices. The columns show the x-, y- and z-direction. The first row shows the whole data set, the observation differs from the model around the mean in the transverse directions. The second row shows the preselected data set, the difference is reduced.

looking at the Gaussian fits, it is clear that the cut rejecting low resolved vertices is preferable to the cut that rejects events by low or high track multiplicity.

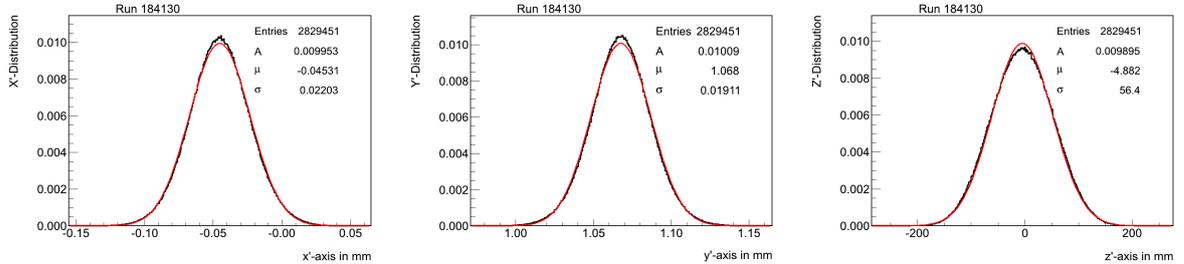


Figure 38: Distribution along the axes of the rotated coordinate system for only those vertices which have a track multiplicity within the RMS range of $23 \leq N_{trks} \leq 63$. This cut has less effect on the shape of the histogram than the limitation of the resolution.

The beamwidth was measured using the Gaussian fits for both selections throughout all examined runs. Figure 39 confirms the evaluation of the histograms in figures 37 and 38 as the beamwidth measured on the sample that was selected by resolution increased in both transversal directions. In the longitudinal direction the measured beamwidth has decreased due to the sample selection which indicates that the resolution is becoming worse with longitudinal distance from the interaction point.

The selection that limited the used events to the RMS range of the track multiplicity does not show a comparable result and actually even lowered the measured beamwidth in all three directions.

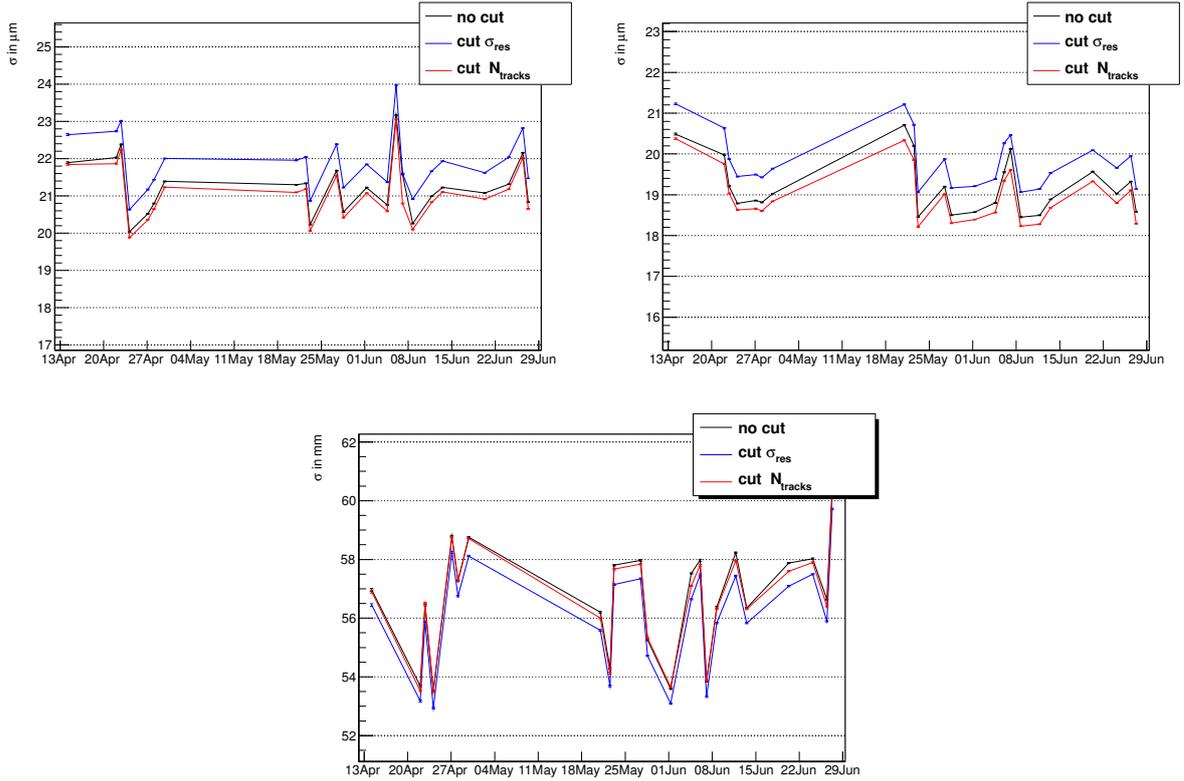


Figure 39: Found beamwidth for the whole dataset and the two examined selections. The blue line shows the cut that rejects events with a resolution that exceeds the average resolution plus one time the resolution RMS. The red line shows the selection that rejects events with a track multiplicity outside the RMS-range of the track multiplicity distribution .

In this context this demonstrated the likelihood of introducing a bias by using a sample design strategy. It further shows that the whole effect of the accumulation of low resolved vertices cannot be reversed by relying on the correlation between low resolution and a low number of tracks.

6.2. The mean estimation

The intuitive mathematical strategy to correct the beamwidth is based on equation 6, that can be rearranged to equation 32.

$$\bar{\sigma}_{real} = \sqrt{\bar{\sigma}_{obs}^2 - \bar{\sigma}_{res}^2} \quad (32)$$

The adaption of this equation is limited to sets of vertices. The idea of treating its distance to the mean, the residual, of the single vertex as its particular standard deviation fails to describe the physical background, as vertices that have the average interaction point lying within their confidence ellipsoid would cause equation 6.2 to map into the complex plane. Furthermore equa-

tion 32 was derived from the assumption that each sigma is a parameter of a normal distribution (see section 2.3.3), but for the single vertex this only applies to the average interaction point that is based on the whole sample, but not to the particular residual. The situation that vertices with lower resolution have a higher probability to be located close to the mean complicates the problem, since their σ_{res}^2 is high and their residual $(x - \bar{x})^2$ can be expected to be low with increased probability.

The bin-wise adaption of equation 32 in bins of track multiplicity is also observed to be troublesome by recalling figure 34 and looking at the range in the lower domain of the track multiplicity axis. The range that has to be excluded to avoid complex numbers reaching into the RMS-range of the track multiplicity distribution. In the range where the calculation starts to give real results, small changes of the beamwidth and resolution lead to big changes in the corrected beamwidth. In the region where the observed beamwidth forms a plateau the application of equation 6.2 looks much more reasonable, thus it is used here only as an estimator including all points by calculating the averages of the preselected sample.

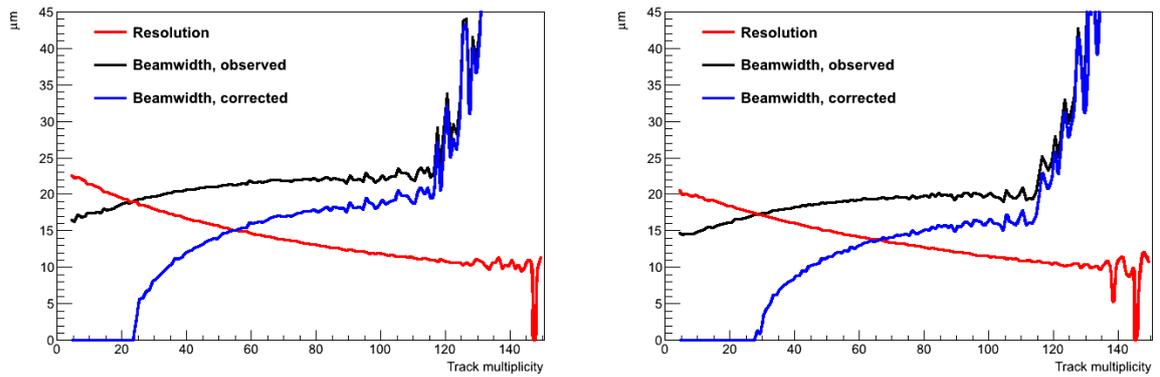


Figure 40: Bin-wise adaption of equation 32 along the track multiplicity axis. The means are calculated along the interval $35 \leq N_{trks} \leq 100$. The left plot shows the x-direction, the right plot the y-direction.

However, using equation 32 to estimate the real beamwidth with the results of the whole sample corresponds to the convolution of the resolution and the real vertex distribution and since it has been shown in the last section that the observed beamwidth using the cut rejecting low resolved vertices is normally distributed in very good approximation, this confirms this strategy. On the other hand, however it has been shown that the resolution is not standard distributed, which somewhat questions the applicability of the derivation in section 2.3.4. The evaluation of this method is given in figure 43, together with the evaluation of the other methods.

6.3. Deconvolution

The estimation of the real beamwidth as computed in the last section relied on the average standard deviations and the average resolution. Thus the gained information is valid for the

whole picture, but cannot be applied to the single events. The methods examined in this section can be seen as mappings which allow the following of each point through the process and the observation of the corrected position for each entry. The characteristic functions of the observed distributions are calculated, which are the Fourier transformed, since the distributions are discrete. Following the convolution theorem the convolution of the distributions of the real vertices and the resolution is equivalent to a multiplication in the frequency domain. Deconvolution by inverting this multiplication is referred to as inverse filtering.

This technique is applied by deblurring algorithms in image restoration which can in some sense be compared to the given problem [20]. A more sophisticated variant of this method is the Wiener filter which additionally considers the signal-noise-ratio in the frequency domain and systematically weighs by that calculation. Both algorithms are applied to the data sets here, but due to time limitations and issues with NumPy which memory leaks DFT coefficients, only the Wiener filter was examined for the whole run sequence.

6.3.1. Inverse filter

The inverse filter takes advantage of the convolution theorem and is probably the first intuitive mathematical approach to the inverse problem. The Fourier transformation can express the data set as a sum of periodic complex functions which equals the data at the measurement points and is complex otherwise. The discrete Fourier transformation finds the coefficients f_k for the periodic functions in the sum of

$$F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi i n \frac{k}{N}}$$

where $v_k = \frac{k}{N}$ are the frequencies corresponding to the dataset in the position-space. The inverse transformation is given by

$$f_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{2\pi i k \frac{n}{N}}$$

The filtering is realised here to the unbinned data which can be interpreted as a discrete set of pairs of the evaluations of the both distributions, the resolution and the observed vertex distribution. Thus the data represents $i = 1, \dots, N$ evaluations $(x_i, \sigma_{res, x_i}) = (f_{obs, i}, f_{res, i})$ of

$$f_{obs(x)} = f_{real}(x) * f_{res}(x)$$

Using the discrete Fourier transformation this gets a multiplication of the characteristic functions of the probability distributions.

$$F_{obs}(v) = F_{real}(v) \cdot F_{res}(v)$$

Dividing at this stage is not necessarily safe, since $F_{res}(v)$ can have frequencies with $F_{res}(v_i) \approx 0$, which would lead to problematic results obtaining very high absolute values for $F_{real}(v_i)$. Thus a

high pass on $F_{res}(v)$ is an important step to bear in mind, but in the given data sets $|F_{res}|$ has been tested not to exceed a range of $O(10^{-4} - 10^4)$ and was divided without either high or low pass. The inverse DFT gets

$$f_{real,est}(x_i) = iDFT \left(\frac{F_{obs}(v_i)}{F_{res}(v_i)} \right)$$

Of course this is limited by the fact, that the given values for σ_{res,x_i} are only estimators for the resolution.

The computation was done using NumPy.fft for the DFT and iDFT. The result is shown in figure 41, where the distribution before and after the inverse filtering can be compared. The beamwidth is reduced to $\sigma_{inv,x} = 13.23 \mu\text{m}$ and $\sigma_{inv,y} = 11.71 \mu\text{m}$, which corresponds to 63.5% of the observed beamwidth in x-direction and 62.1% of the measured beamwidth in y-direction.

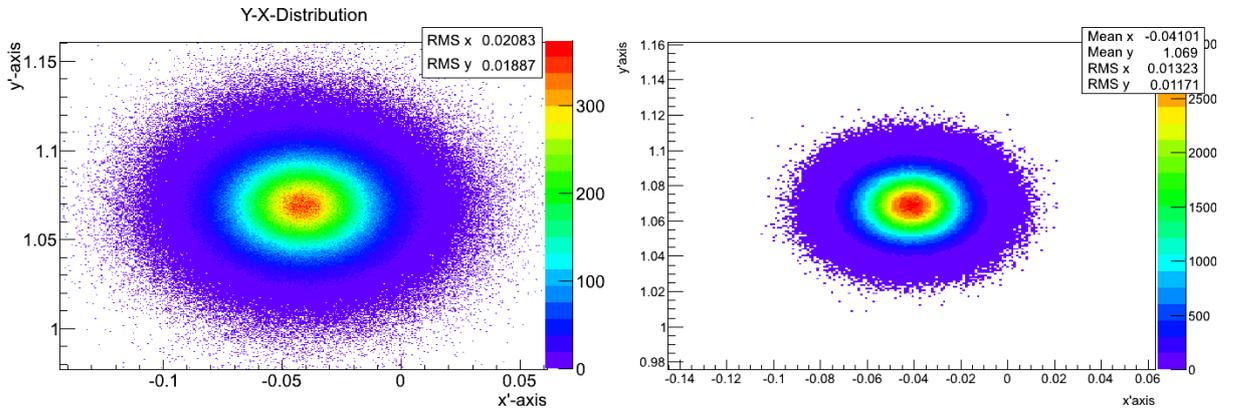


Figure 41: The deconvolution by the inverse filter produces a data set with a significantly contracted beam profile. Since it shifts points such that the result is centred at the origin, the point distribution is translated back into the average interaction point after the deconvolution.

6.3.2. Wiener filter

The Wiener filter applies basically the same concept, but introduces an additional approach to weigh the frequencies by their signal-to-noise-ratio SNR . It depends on uncorrelated signal and noise, here the discrete sequence of the measured points and the disturbance by the resolution limitation. As has been shown there are some correlations between the spatial distribution and the measurement precision, but these are small enough to apply the Wiener filter. The Pearson coefficient for the vertex coordinates and the resolution evaluates to $|corr(x, \sigma_{res,x}^2)| < 0.01$ in all three directions. Provided that $x_{obs,i} = x_{real,i} + \sigma_{res,i}$, where the addition is compensated for by the symmetry of the luminosity region, in the frequency domain an estimator is searched to minimize the least square error.

$$\sum_i (F_{real,i} - H_i \cdot (F_{real,i} + F_{res,i}))^2 \rightarrow Min \quad (33)$$

Minimizing this expression and using that the correlation between F_{real} and F_{res} is approximately zero gets equation 34, where Φ_{xx} is the autocorrelation function of the coordinate sequence and Φ_{rr} is the autocorrelation function of the resolution sequence [20].

$$F_{real,est} = H \cdot F_{obs} = \frac{(|F_{obs}|)^2 \Phi_{xx}}{(|F_{obs}|)^2 \Phi_{xx} + \Phi_{rr}} \quad (34)$$

Dividing by Φ_{xx} gets the term Φ_{rr}/Φ_{xx} added to the denominator, which becomes large for a bad NSR and thus lowers the frequency, and becomes small for a good NSR letting the filter behave neutral since the fraction approximates 1. The inverse DFT gets the filtered sequence.

The Wiener filter is used as a function provided by SciPy.signal.signaltools here and performed on the coordinate sets of the analysed data.

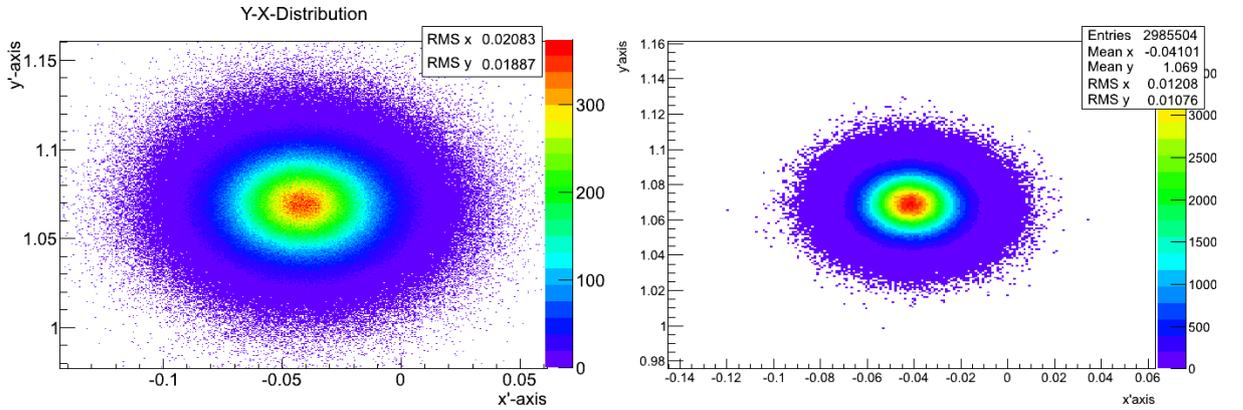


Figure 42: The deconvolution by the Wiener filter also significantly contracts the beam profile. The Wiener filter does not shift the points like the inverse filter, the means in both directions are the same after the deconvolution.

The beamwidth is contracted by the deconvolution using the Wiener filter to 57.0% in x -direction and 58% in y -direction. The results for all runs are shown in figure 43, in order to compare the estimation by the quadratical subtraction and the deconvolution using the Wiener filter. The Wiener filter gives lower results in x -direction, in y -directions both methods show less difference. The average beamwidth found by the Wiener filter is $12.3 \mu\text{m}$ in x -direction and $11.1 \mu\text{m}$ in y -direction. In y -direction the estimation evaluates to the same average beamwidth for the whole run sequence. In x -direction the estimation using equation 32 gets $13.7 \mu\text{m}$. The runs evaluated using the inverse filter have values that are of the same order. The Wiener filter applies a more sophisticated strategy to the problem that takes all events into account. There are ways to apply an error estimation to this process, but this cannot be examined here due to the limited time.

Since both methods evaluate to comparable values, both decreasing the beamsize significantly, the expected confirmation can be given, that the real size of the luminosity zone significantly lower than the observed beamsize.

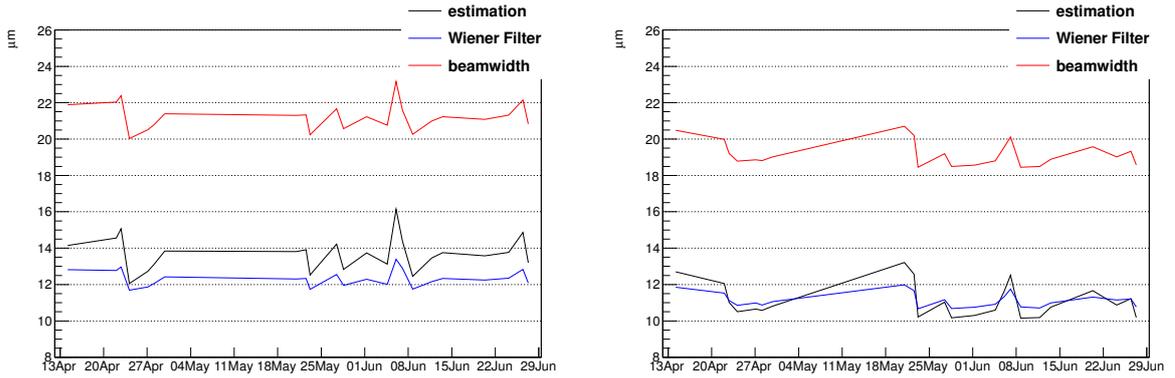


Figure 43: Comparison of the corrected beamwidth using the deconvolution with the Wiener filter and the estimation by quadratic subtraction. The left plot shows the x-direction; the right plot shows the y-direction.

7. Distribution of the pile-up events

The average amount of interactions per bunchcrossing was 5-6 at the time of the examined run sequence. The analysed data sets have a corresponding number of additional vertices per event. An intuitive idea is at first hand to include these vertices into the distributions in order to get a better statistic. A brief discription shall be given here to explain that the pile-up events cannot be used to measure the size of the beamwidth. Even applying a very strict selection is not likely to get any agreement of the distribution of the primary vertices and the pileup events. Measuring the beamwidth by only taking the pileup events into account gives a very different picture as figure 44 illustrates. Both transverse beamwidths measured in the distributions of the pileup vertices are smaller than $10 \mu\text{m}$. But in the longitudinal direction the beamwidth of the pileup distribution differs less than 3% from the one of the primary vertices.

This observation can be explained by the results of section 5. Following their lower average resolution, the transverse location of the pileup vertices is traced with significant less precision, which lets them appear as being located closer to the mean in the observation. Since the analysis showed as well that vertices with fewer tracks do in average have a lower resolution, the comparison of the track multiplicity complies with the analysis as well.

The RMS range for the track multiplicity of the pileup vertices is $N_{trks} = 14 \pm 13$. The correlation of this domain on the track multiplicity axis and a low resolution has been shown in the dicussion of figure 34. Thus a further analysis of the spatial distribution of the transverse vertex coordinates of the pileup events is dispensable, since it is not likely to have a good agreement between the observation and the real distribution of the pileup vertices. One has to keep in mind, that the tracking process has microseconds to analyse the bunch crossing, thus the distinction of the pileup vertices from the primary event by buiding clusters with respect to their z-coordinates is efficient enough and the pileup vertices don't need to be better resolved. However, they cannot count as an indication of the size of the luminosity region.

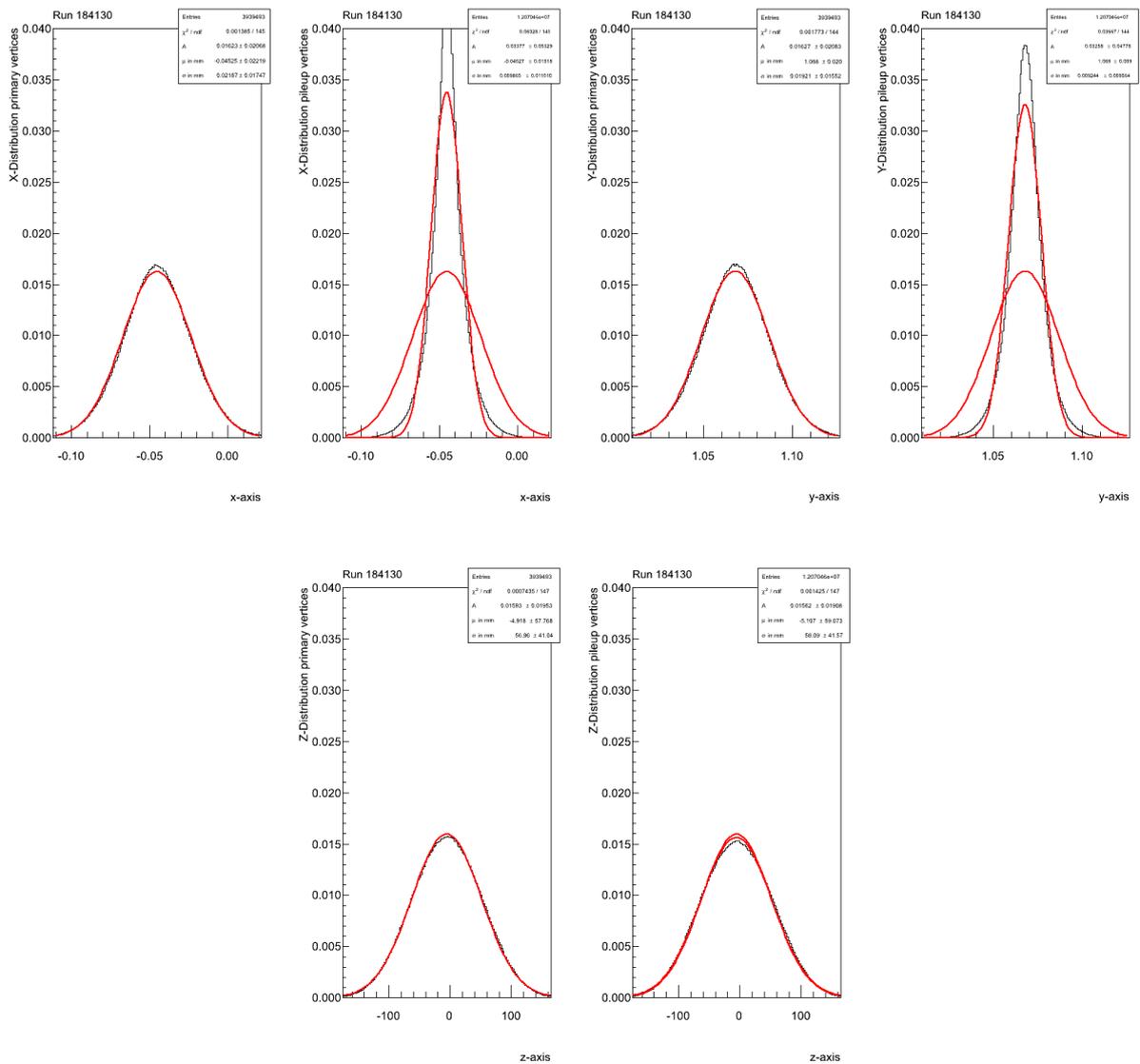


Figure 44: Distribution of the primary events (left) and the pileup events (right). The fit function for the primary events is always as well shown on the right side, to provide a better comparison.

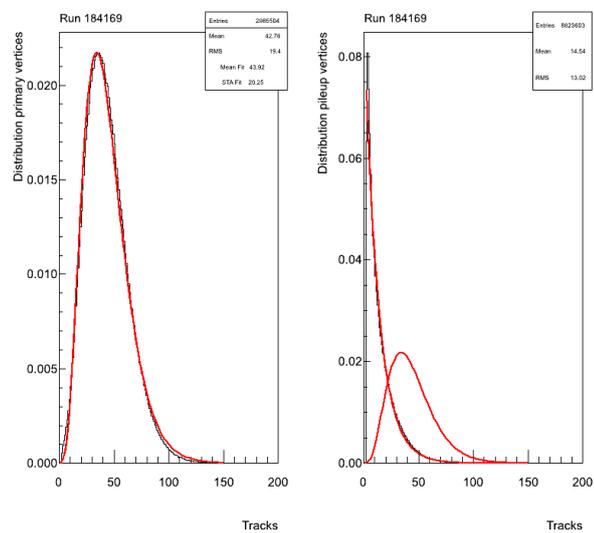


Figure 45: Track multiplicity for run 184169 for primary and pileup vertices. The curve of the fit of the primary vertices is added to the distribution of the pileup vertices, in order to provide comparability.

8. Conclusions

In this section a summary of the results is given as well as a brief discussion of the advantages and disadvantages of the analytic procedure presented in this thesis. The limitations of this thesis are addressed and an outline of topics for further work on the subjects “size of the luminosity region” and “vertex resolution” is given.

8.1. Summary

This thesis provides a detailed analysis of the structure of the luminosity region of the ATLAS detector. This includes a discussion of the measurement precision, which deepens the understanding of the observed vertex distribution. The findings were validated using a sequence of 23 runs which gave a good opportunity to find correlations between different observations and to confirm the central topics.

It has been found that the resolution of the vertex reconstruction correlates with the beam width such that a run with smaller beamwidth can be expected to produce a data set with a better average resolution. This has been found in both of the transverse directions and it has been shown that not only the beamwidth was smaller in the vertical direction than in the horizontal direction, but the average standard deviation assigned to each event due to its resolution correlates with this and was also smaller in the vertical direction in all examined runs.

The spatial distribution of the average local resolution in the luminosity region has been analysed in detail and it is understood that the accumulation of vertices which is observed in the one-dimensional transverse distributions is related to the vertex resolution. It has been shown that a selection rule that rejects low resolved vertices can reverse parts of this effect and can be expected to lead to the choosing of a subset that has a better agreement with the assumption of a normal distribution of the events in the luminosity region. This has been compared with the distribution along the track multiplicity axis and it is understood, why the distribution of the bin-wise observed beamwidth decreases with low track multiplicity as the average standard deviation due to the resolution increases in those subsets.

Ways to correct the observation from the resolution limitation have been examined and an estimation of the real vertex distribution could be derived from the analysis.

8.2. Discussion

The determination of the beamwidth was realised by applying Gaussian fits to the data, which means additional work for not necessarily better results. One could also deploy the computational formula for the variance that is implemented in most working environments. However, in that case it is recommended not to compute on the whole dataset without any test for outliers, since the underflow and overflow buffers of the histograms were never unfilled.

The efficiency of the coordinate system rotation is low and the simplified variant of the used

matrix should be sufficient as long as there is no significantly higher tilt. The computation takes CPU time and the results have to be stored separately in order to avoid the frequent repetition of the process. It may be of further interest to measure the beamwidth with a parallel orientation to the beamline, although an analysis of the measurement precision of the detector might in turn be better carried out in the detector's coordinate system.

A main issue discussed in this thesis is the observed shifting of points with low resolution towards the beamline by the HLT fit process. The observation was based on the computation of the resolution from the perigees by calculating their covariance. An improvement to this method is presented by [16], using the split vertex method. This method aims to minimize the confidence interval that the resolution represents by dividing the group of tracks for a given vertex iteratively into two half-sets and finding the position of these half-vertices. This method is likely to be capable of reducing the suggested effect of a backlink of the computed beamwidth into the resolution values.

The suggested and analysed sample design strategies were only roughly evaluated by looking at the histograms. When testing sample designs and selection criteria, the χ^2 -test should be used to express a comparable parameter. This was just observed in the computing shell window, but not used as a criterion to evaluate the effects of the sample selections for the runsequence. However, more reliable results could be achieved by examining this more thoroughly.

The resolution correction by deconvolution goes the path of the inverse problem which is probably not the best strategy provided reliable prior information about the real vertex distribution. A Bayesian analysis which goes forward through the convolution and minimizes the loglikelihood function to find the best estimator for the real vertex distribution is likely to provide good results on estimating the real size of the luminosity region.

8.3. Future work

It is of great interest to examine the distribution of the vertex resolution as it is correlated with the vertex distribution. As discussed in section 2.3.2 there are additional offline vertex finding algorithms in place. A study which applies the analysis given in section 5 to those algorithms may obtain new conclusions.

It has been shown that sample design strategies should be well thought through and there are a lot of interesting evaluations to carry out by applying selection rules and analysing the effect on the observation.

Very interesting would be an attempt to do a Bayesian analysis in order to find the corrected beamwidth. Its distribution can be easily estimated with good approximation, since the real standard deviation of the assumed normal distribution of the events in the luminosity region is the only parameter that is unknown. There is a huge set of sophisticated methods about statistical estimation as well as about strategies solving the inverse problem, that could be deployed to the

measuring of the luminosity region. In this thesis only the most common approaches to this topic have been examined.

A. Appendix

A.1. Collections of plots for all 23 runs

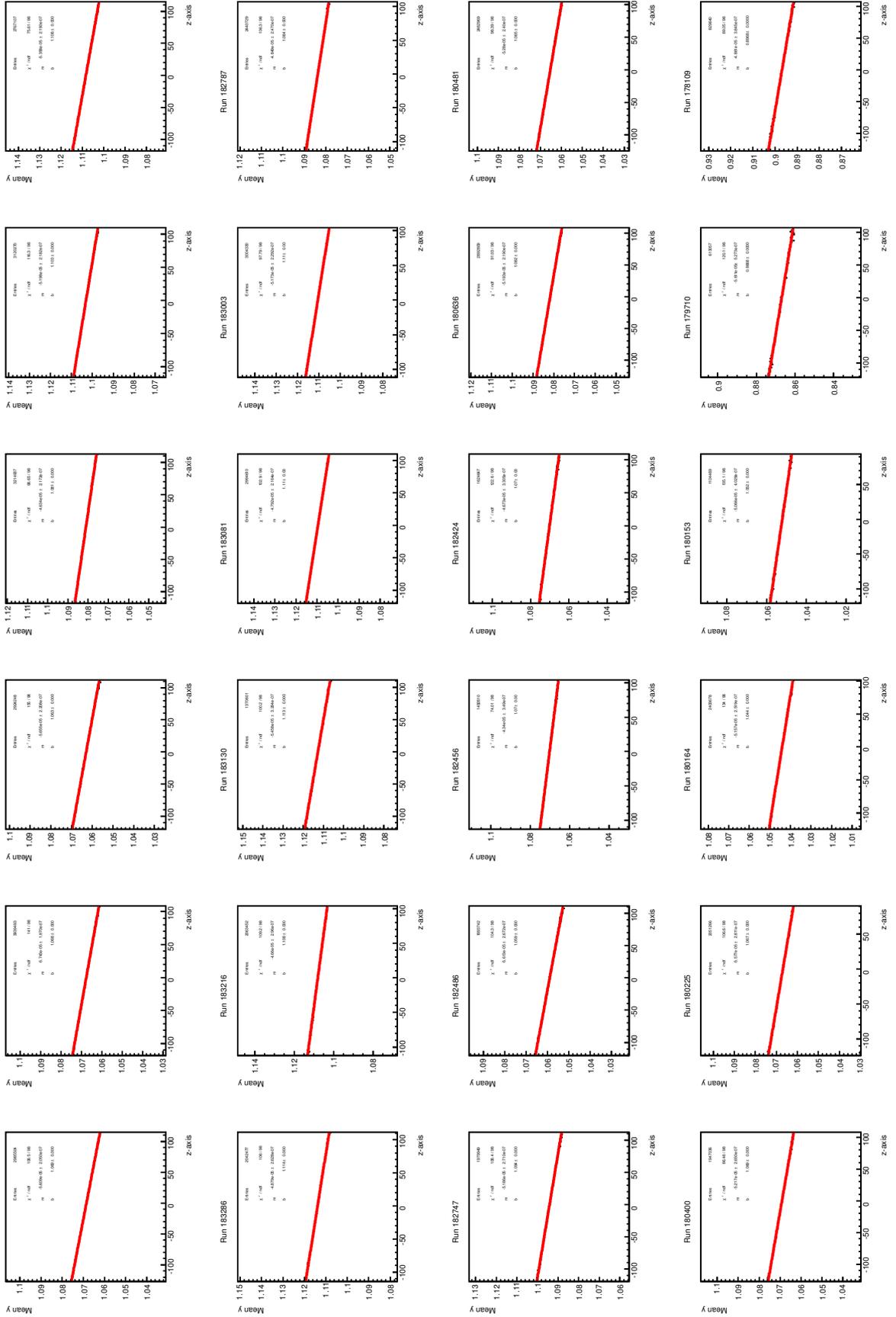


Figure 46: Beamspot mean in y-direction taken in 100 intervals along the z-axis. The slope of the fitted linear function represents the y-tilt of the beamline. The drawn intervals show a range of $\pm 2\sigma$ around the mean on both axes.

A.1 Collections of plots for all 23 runs

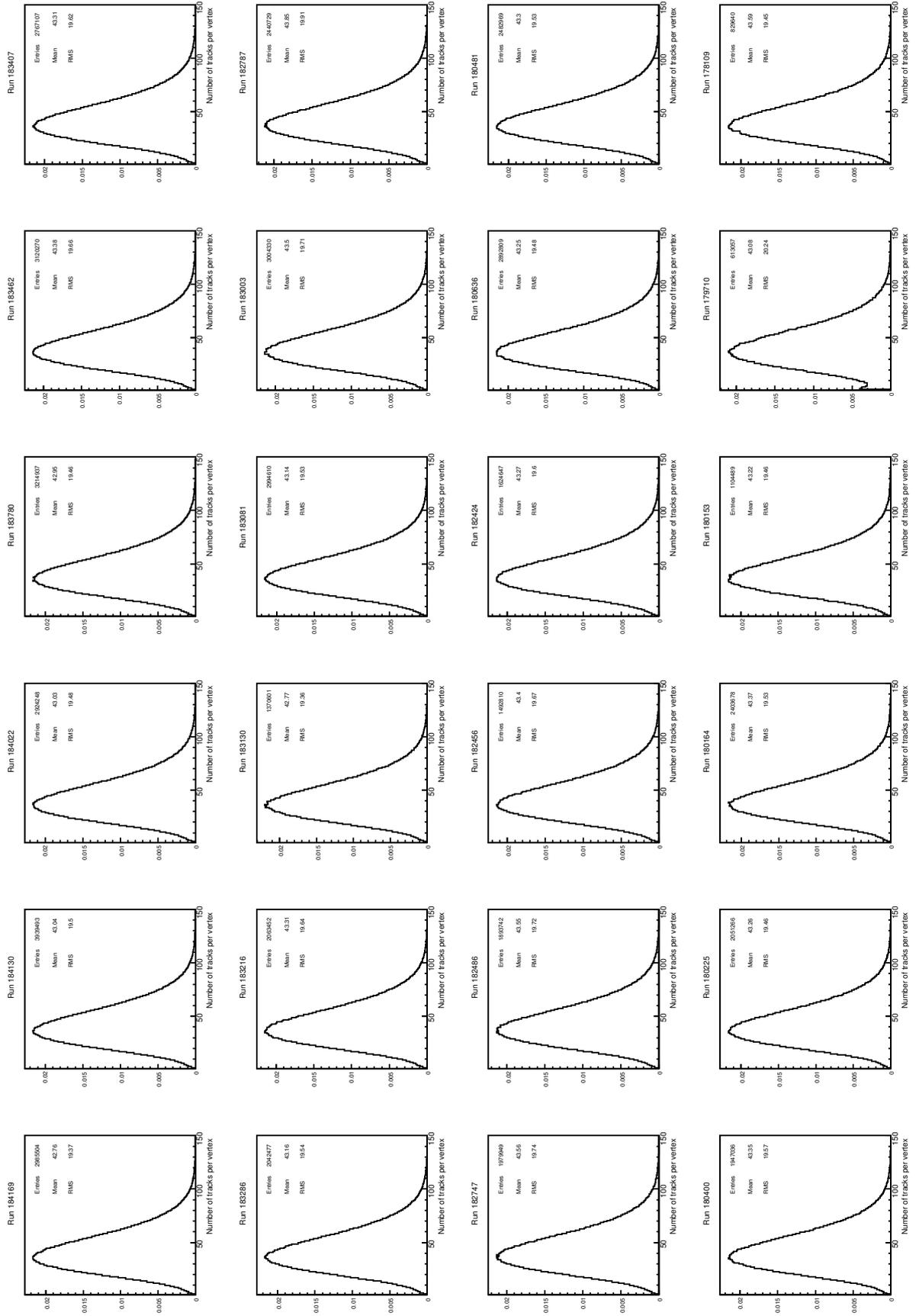


Figure 47: Distribution of number of tracks per vertex for all runs

A.1 Collections of plots for all 23 runs

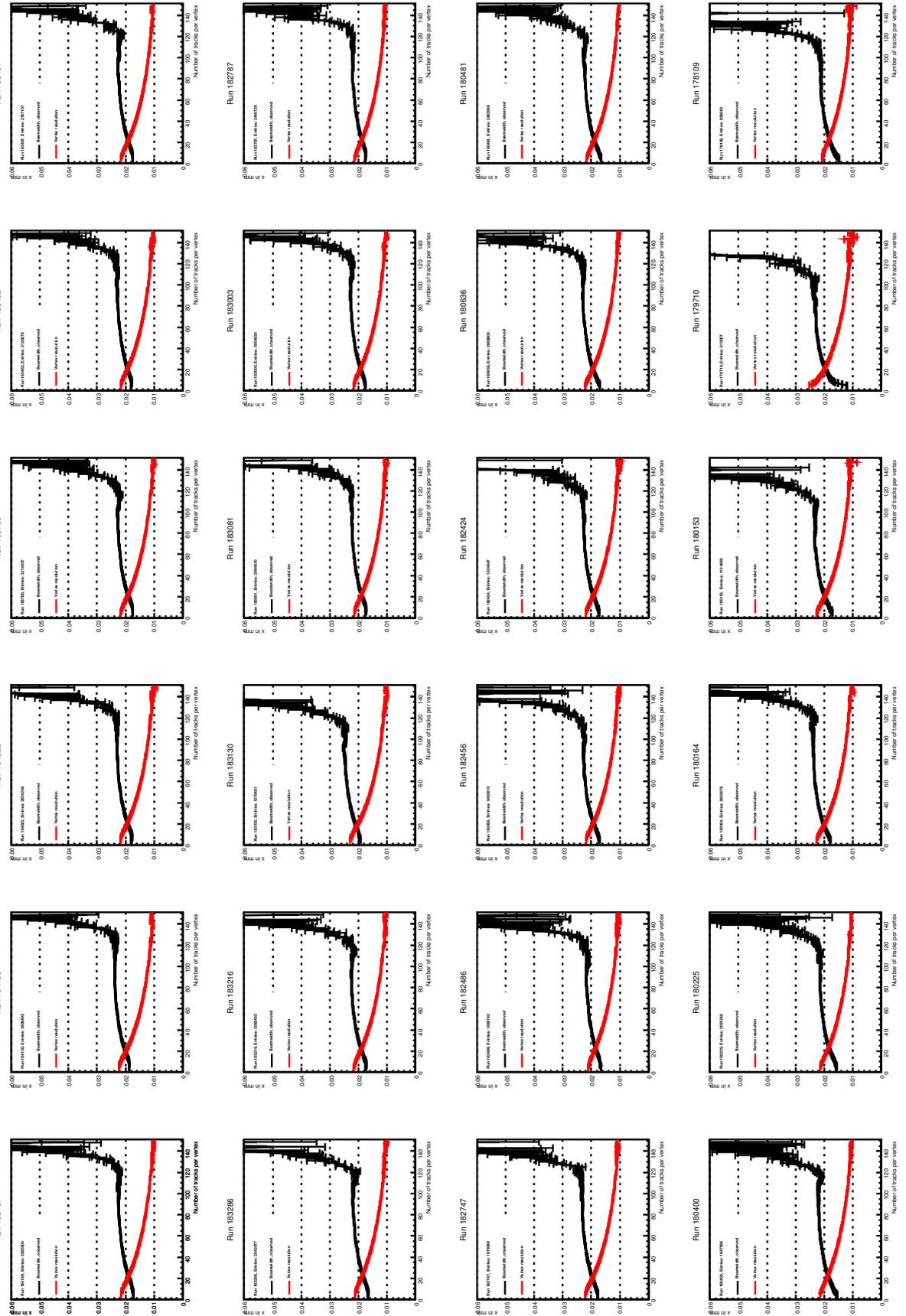


Figure 48: Observed beamwidth and resolution in x-direction vs. number of tracks. Collection for all runs.

A.1 Collections of plots for all 23 runs

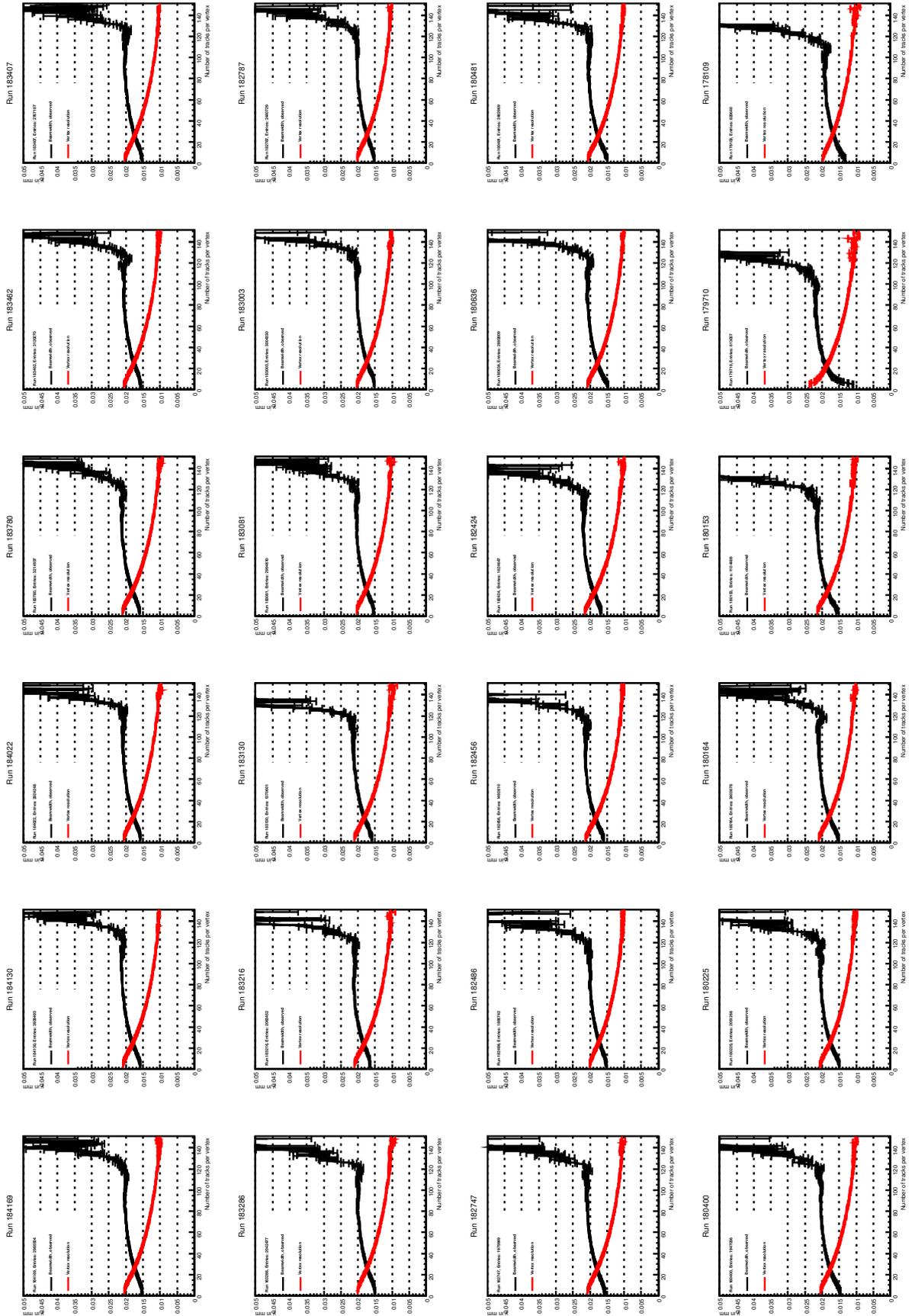


Figure 49: Observed beamwidth and resolution in y-direction vs. number of tracks. Collection for all runs.

A.1 Collections of plots for all 23 runs

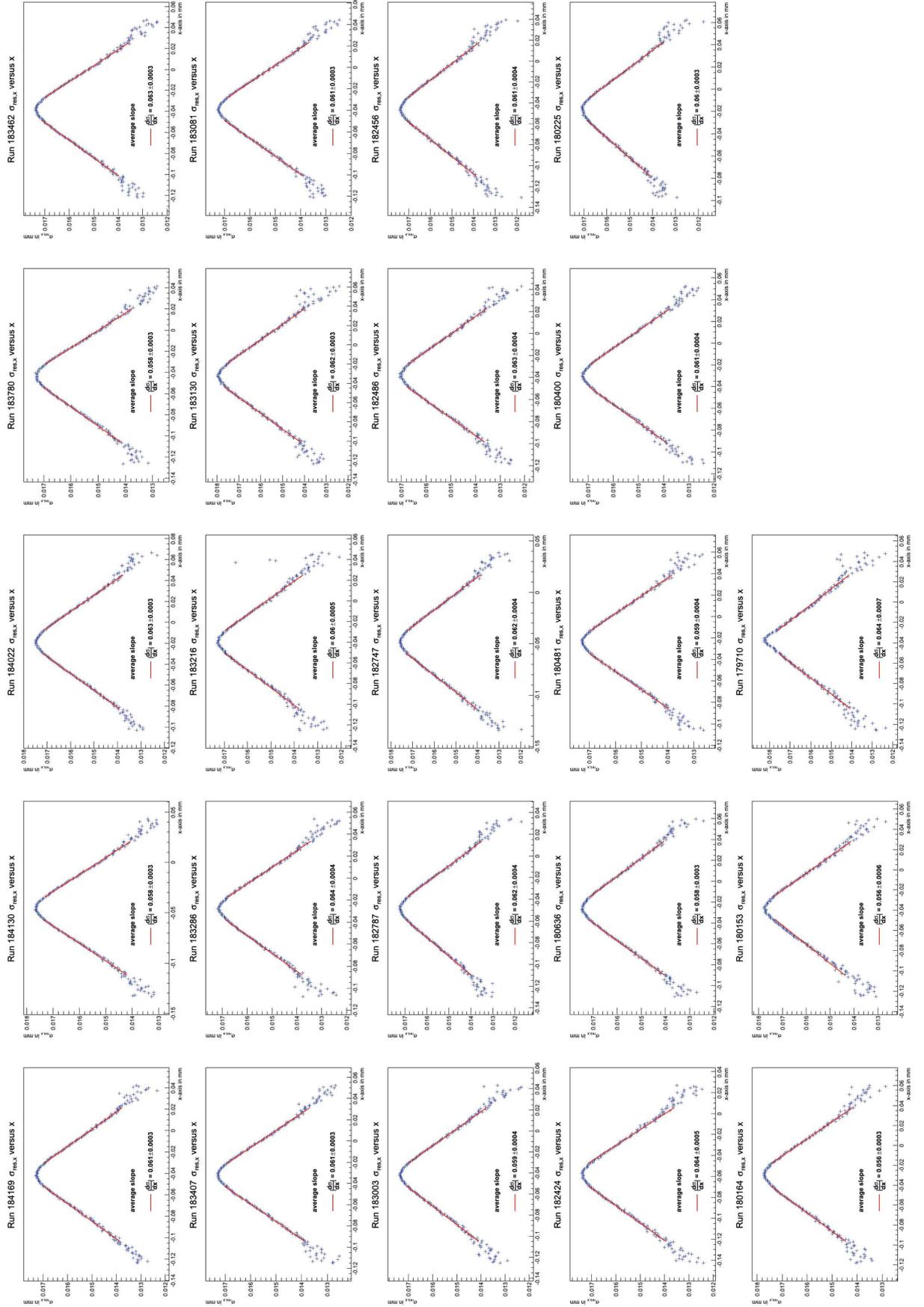


Figure 50: Local transverse resolution $\sigma_{res,x}$ along the x-axis.

A.1 Collections of plots for all 23 runs

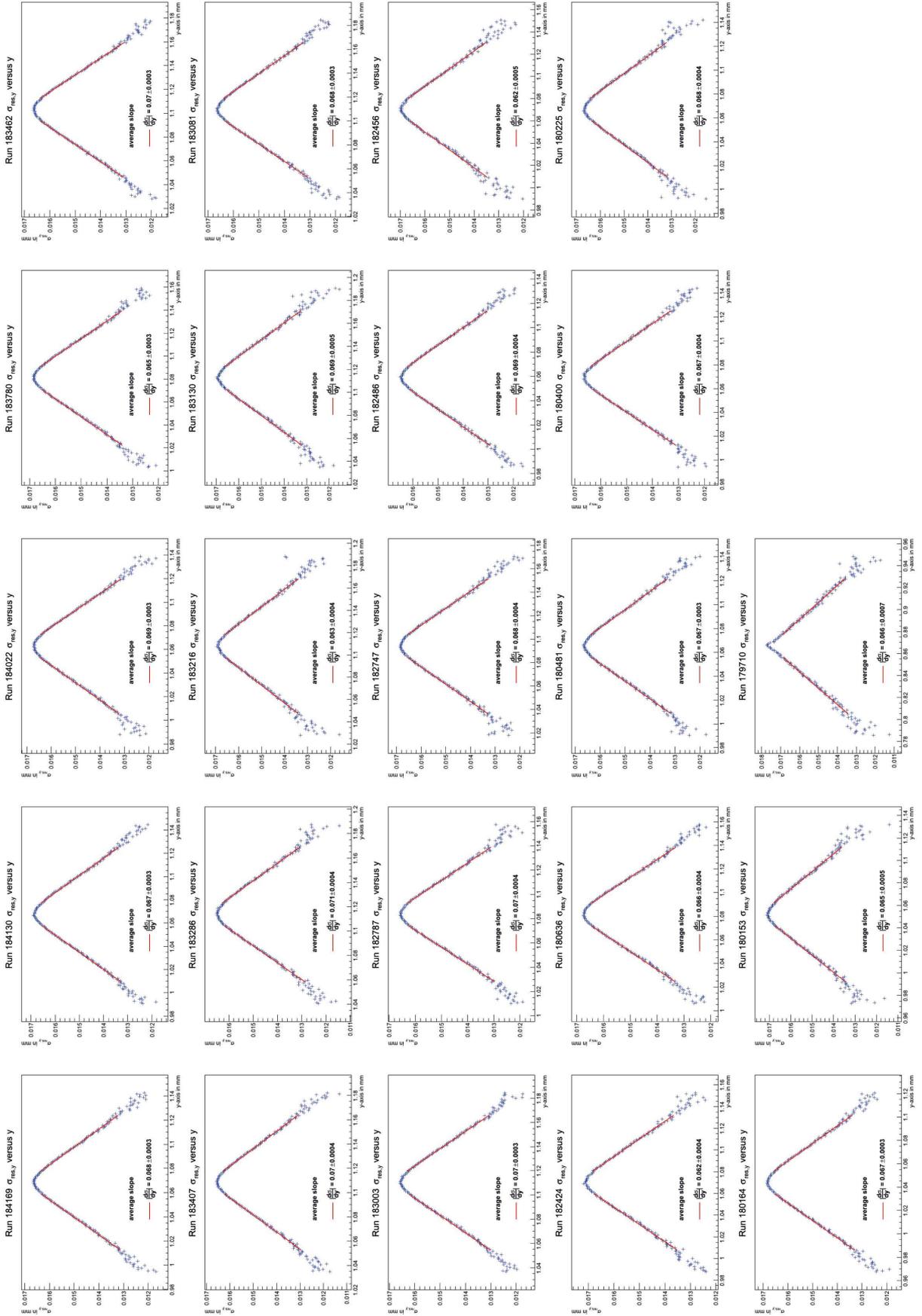


Figure 51: Local transverse resolution along the y-axis.

A.1 Collections of plots for all 23 runs

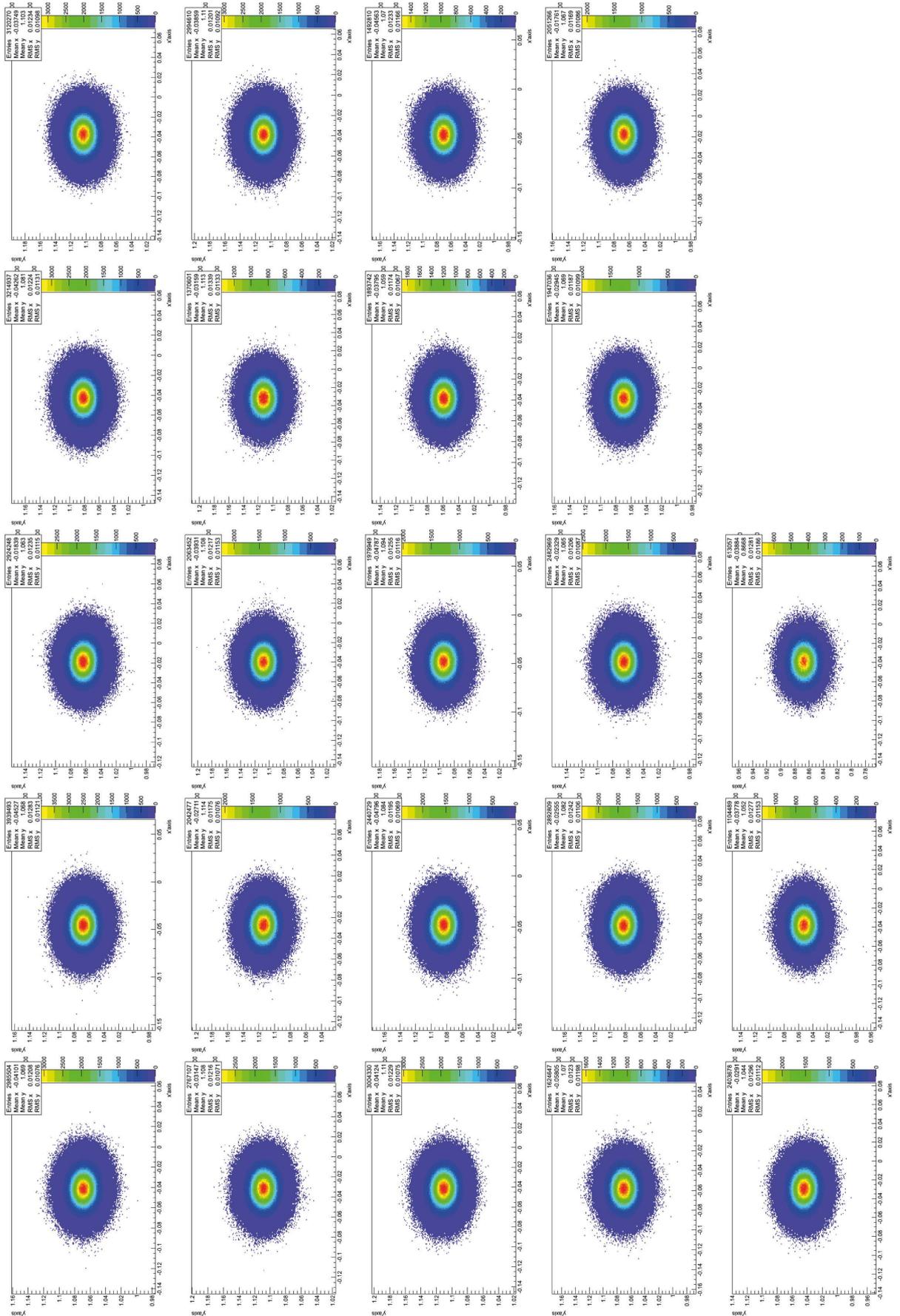


Figure 54: Distribution in the x-y-plane after the Wiener filter.

References

- [1] CERN. Atlas and cms experiments present higgs search status. *Cern Press Release*, 2011.
- [2] Rene Brun and Fons Rademakers. *ROOT - An Object Oriented Data Analysis Framework*. <http://root.cern.ch>, 1997.
- [3] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.
- [4] Travis E. Oliphant. Guide to numpy. numpy.scipy.org, 2006.
- [5] J Bleck-Neuhaus. *Elementare Teilchen, Moderne Physik von den Atomen bis zum Standard Model*. Springer-Verlag, 2010.
- [6] K. Nakamura et al. Pdglive particle summary. *Particle Data Group*, 2011.
- [7] LHC design report. lhc.web.cern.ch/lhc/LHC-DesignReport.html.
- [8] www.weltmaschine.de/cern-und-lhc.
- [9] S. Baird. Accelerators for pedestrians. <http://cdsweb.cern.ch/record/1017689>, 2007.
- [10] Klaus Wille. *Physik der Teilchenbeschleuniger und Synchrotronstrahlungsquellen*. Teubner Studienbücher, 1996.
- [11] Lhc machine outreach. <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/>.
- [12] The ATLAS collaboration. *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS-Detector at the LHC*, 2010. ATLAS-CONF-2010-027.
- [13] The Atlas Collaboration. Measurement of the inelastic proton-proton cross-section at 7 TeV with the atlas detector, 2011.
- [14] atlas-runquery.cern.ch.
- [15] <http://lhc-statistics.web.cern.ch/LHC-Statistics/>.
- [16] R. Bartoldus. Online determination of the lhc luminous region with the atlas high-level trigger. Technical report, 2011.
- [17] Primary vertex reconstruction in the atlas experiment at lhc. Technical report.
- [18] Hongbo Zhu. Atlas status and results. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic>, 2011.

- [19] Eric W. Weissstein. *Convolution*. MathWorld – A Wolfram Web Recource, mathworld.wolfram.com/Convolution.html, 2011.
- [20] F Wahl. *Digitale Bildsignalverarbeitung, Grundlagen, Verfahren, Beispiele*. Springer-Verlag, 1988.

List of Figures

1.	The W^- boson as exchange particle in the beta decay	6
2.	The Large Hadron Collider and the four detectors	8
3.	Bunch filling scheme of run 184169	9
4.	Phase-space ellipse of the particle motion in the $x-x'$ -space	12
5.	Relative beamsizes around the interaction point in the ATLAS Detector	13
6.	Cut-away view of the ATLAS Detector	14
7.	Online luminosity and beam intensity	16
8.	Luminosity production 2011 during the proton-proton physics program [15]	17
9.	Spatial distribution for all primary vertices of run 184169	24
10.	Spatial distribution using all primary vertices. Overview for all 23 runs	25
11.	Variation of the observed beamwidth for all 23 runs	26
12.	Observed beamline tilt of run 184169 for both transverse directions	27
13.	Tilt in mm^{-1} for the run sequence	28
14.	Tilt in percent of beamwidth σ for the run sequence	28
15.	2-dimensional distributions in the x - z -plane and the y - z -plane	29
16.	2-dimensional distributions in the x' - z' -plane and the y' - z' -plane	31
17.	Observed vertex distribution in the adjusted coordinate system	32
18.	Change of the measured beamwidth after rotation in percent	33
19.	Area of the transverse beam projection	34
20.	Absolute value of the vertex resolution in the transverse directions	36
21.	Resolution distribution for all spatial directions for run 184169	37
22.	Ratio of the transverse resolutions	38
23.	Absolute value of the vertex resolution in transverse vs. longitudinal direction	39
24.	Trend of the transverse resolution in dependence of the longitudinal resolution	39
25.	Change of the measured beamwidth after rotation in percent	41
26.	Transverse beamwidth ratio $\sigma_{obs,y}/\sigma_{obs,x}$ vs. median of $\sigma_{res,y}/\sigma_{res,x}$	42
27.	Average local resolution in the transverse plane	43
28.	Average local transverse resolution in the transverse plane	44
29.	Average local resolution in the x - z and y - z -plane	44
30.	Average local resolution $\bar{\sigma}_{res,x}$ along the x -axis and $\bar{\sigma}_{res,y}$ along the y -axis	45
31.	Range of the local transverse resolution for the examined run sequence	45
32.	Local longitudinal resolution variation along the z -axis	46
33.	Distribution of the track multiplicity	47
34.	Observed beamwidth and vertex resolution vs. number of tracks per vertex	48
35.	Beamspot position vs. number of tracks	49
36.	Histogram of the points with a track multiplicity of $N_{trks} \geq 100$	49
37.	Vertex selection by resolution	51

38.	Vertex selection by track multiplicity	51
39.	Comparison of the vertex selections for the run sequence	52
40.	Beamwidth correction by bin-wise quadratic subtraction	53
41.	Deconvolution by the inverse filter	55
42.	Deconvolution by the Wiener filter	56
43.	Corrected beamwidth for the run sequence. Wiener filter and quadratic subtraction	57
44.	Distribution of the primary events and the pileup events	58
45.	Track multiplicity for primary and pileup vertices	59
46.	Beamspot mean in y-direction	III
47.	Distribution of number of tracks per vertex for all runs	IV
48.	Observed beamwidth and resolution in x-direction vs. number of tracks	V
49.	Observed beamwidth and resolution in y-direction vs. number of tracks	VI
50.	Local transverse resolution along the x-axis	VII
51.	Local transverse resolution $\sigma_{res,y}$ along the y-axis	VIII
52.	Distribution in x-direction after rejecting vertices with low resolution	IX
53.	Distribution in y-direction after rejecting vertices with low resolution	X
54.	Distribution in the x-y-plane after the Wiener filter	XI

List of Tables

1.	Table of fermions	4
2.	Table of bosons	5
3.	Angular range covered by the subdetectors of the ATLAS Detector	15
4.	ATLAS trigger configuration	18
5.	Selected runs and their fill parameters	22
6.	σ for primary vertices, average over all 23 runs	25
7.	Beamwidth in the rotated system and in the ATLAS coordinate system	34
8.	Average resolution versus observed beamwidth	40

Acknowledgements

I want to express my appreciation of Prof. Dr. W. Wagner's support, he has splendidly advised me through the time of the thesis writing. I thank Priv.-Doz. Dr. D. Wicke for agreeing to be the co-examiner for this thesis.

Dr. Klaus Hamacher and Jun. Prof. Dr. Thorsten Dickhaus, Georg Sartisohn and Christian Riegel helped me reflecting upon topics and getting started with ROOT.

JC French did a great job on spell-checking my English.

This is the right place to state my appreciation for all who work on the open source projects Ubuntu, Python, Scipy, NumPy, ROOT and Lyx, you provided the technical background to do this.

My greatest THANKYOU goes to my familiy and friends who supported me at their very best.

Eidesstattliche Erklärung

Hiermit erkläre ich, Michael Schuh, dass ich diese Arbeit zum Thema
A STUDY OF THE LUMINOSITY ZONE OF THE ATLAS DETECTOR AT THE LHC
selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate
kenntlich gemacht habe.

Wermelskirchen, den 19.01.2012
