Development and Performance Measurements of a Silicon Tracker Prototype for ATLAS and Study of Gluino Cascade Decays into Same Sign Dileptons using the ATLAS Inner Detector.

Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften der Abteilung Physik der Universität Dortmund

vorgelegt von Armin Reichold

Dortmund, 12. Juni 1996

Prologue

Curiosity is one of the fundamental properties of humans. Through following their curiosity humans as well as animals are able to acquire knowledge which will be beneficial for their lives. When living in a "society"[a] the knowledge obtained by a member will not only be spread but will also be preserved beyond the lifespan of the individual. In human societies the task of following curiosity in places where it is difficult to do so has been given to experts or specialists which form a fraction of society with special skills and functions. It is assumed that the expert, subsidised by its society will be able to guide and develop curiosity, express it in questions, find knowledge in the attempt to answer them and present the knowledge to the society together with interpretations and suggestions concerning its relevance. One "species" of specialist in modern societies are natural scientists of which high energy physicists are but a few. In the view of science as a function of society it is one of the essential responsibilities of the particle physics community to motivate and to a varying degree justify each scientific effort that is to be carried by society. This is in particular true for a project, the size of the LHC and its associated experiments. It would be foolhardy, and besides the point to attempt to justify the LHC in this thesis towards the many sides of society which could ask for justification. Instead of trying to justify ATLAS or the LHC, it will only be attempted to present the authors motivation for the effort invested into ATLAS, one of the proposed experiments at the LHC.

Common to a great fraction of specialists curiosity is the urge to find common reasons and principles in separate problems. This desire to unify has driven particle physicists to the creation of the standard model of particle physics. Questions related to the validity of this model and possibilities for its extension are the main fields of curiosity for particle physicists involved in the LHC and ATLAS. There are good reasons, some of which will be given later in this thesis, to believe that many of these questions could be answered by experiments like ATLAS. And we could hope to gain new insight into the structure and dynamics of fundamental particles, possibly bringing us one step further in the attempt to unify the understanding of the forces governing the behaviour of matter. Even though this understanding may be extremely complicated and require long studies and years of work in the field of particle physics, the interest into the underlying fundamental questions extends into society far beyond the community of particle physicists. To give up the attempt to condensate, explain and condense this knowledge until it becomes accessible for society, simply by saying that:" If even we can't understand it how could anyone else do?", as is often heard in particle physics, would deny one of the functions of scientists in our society. It may take a very long time - how long is to a great extent in our hands - to really understand the consequences of what is currently achieved on the frontiers of high energy physics but it is surely not impossible.

[a] Any form of living in a group that will be profitable for the individual

Contents

Chapter 1

Introduction

The standard model of particle physics still has many open questions, several of which, such as the search for the Higgs boson, could find an answer through the investigation of particle interactions at energies of $\mathcal{O}(1\ \mathrm{GeV})$. After the cancelation of the SSC[1], the LHC[2] offers the soonest possibility to perform these studies. Out of the two general purpose experiments approved for the study of LHC collisions, ATLAS[3] is one. Many of the subjects that can be studied with ATLAS require the precise reconstruction of charged particle trajectories. With silicon detectors it is possible to detect these trajectories with a very high intrinsic resolution. This is the reason why silicon strip detectors will constitute the majority of all subdetectors in the ATLAS inner tracking system. The requirements from the physics towards position stability and resolution, from the LHC environment towards the radiation hardness, and from engineering considerations towards reliability and mass production are setting the specifications for an ATLAS silicon tracker module. If one wants to make use of the excellent intrinsic properties of silicon as a detection material the detector modules have to be extremely carefully designed and verified before the modules are mass produced. This is why the design of a silicon strip module for ATLAS driven from the above specifications is presented in chapter 4 together with a description of the manufacturing process and conclusions concerning improvements of the module based on the experience of the construction of two modules with associated support, cooling and readout system for thermal and mechanical tests.

Due to the complexity of the mechanical and thermal problems involved in the design of a silicon tracker module and the in some cases very strong dependence of module properties on small variations in their production process as well as through the need for an extremely light and radiation transparent mechanical structure, leaving no room for mechanical safety in the design, it is crucial to test the modules after the design stage and verify their thermal and mechanical design properties under realistic operational conditions. In chapter 6, thermal and mechanical tests of the modules described in chapter 4 are shown. The two main techniques used for these tests, ESPI[4] and IR-thermography are described in chapter 5. The limitations and applicability of both techniques in the verification of the design of large area precision tracking detectors as well as the results of the measurements on several test structures are also discussed in chapter 6. The development of a calibration process for IR-thermography images using visible platinum resistance thermometers and the implementation of methods for the analysis of ESPI fringe data is

[1] Superconducting Super Collider
[2] **L**arge **H**adron **C**ollider
[3] **A** **T**oroidal **L**HC **A**paratu**S**
[4] Electronic Speckle Pattern Interferometry

also described.

Concerning the use of tracking detectors in ATLAS, a way of combining the precision measurements from silicon detectors with the information from the TRT, the continuous tracking detector of ATLAS, in form of a pattern recognition algorithm is shown in chapter 7. With a full ATLAS detector simulation the performance of the algorithm is intensively tested, in particular for its stability against high luminosity. The track parameter resolutions are measured as functions of detector position, track momentum and luminosity and the consistency of the results with analytical upper limits for the resolution is studied.

In chapter 8 the developed pattern recognition algorithm is used to study the decays of $\tilde{g}$-pairs in the ATLAS inner detector. Problems with signal intrinsic backgrounds, in particular from semi-leptonic decays of heavy quarks are discussed and the use of impact parameter cuts for their suppression is shown. Results for the track finding efficiency and charge misidentification probability as well as track parameter resolutions are shown and their consistency with the previously obtained results is demonstrated.

Chapter 2

Motivations

The motivations for building the LHC and ATLAS lie in the great potential for discoveries and precision studies that could possibly be made though pp-collisions at a center of mass energies of 14 TeV. Some of the questions currently considered to be answerable with the LHC and ATLAS will be given in this section.

2.1 The Standard Model

The standard model of particle physics (SM) has been described in many books [1] and will not be described here in length, but points in which the SM presents open questions will be discussed, focussing on those that could find an answer through experiments at the LHC.

One of the most unexplained and so far unmotivated properties of the SM is the separation of all fermions, spin 1/2 particles that make up matter, into three generations or families of identical internal symmetry but increasing mass scale. This property has recently been fully confirmed by the discovery of the top quark [2], the up to then last unobserved fermion of the SM. The generation phenomenon is particularly puzzling as only the first and lightest generation of fermions is required to describe stable matter. The properties of the top quark can be studied at the LHC in great detail as it will be produced with a large cross-section of approximately 1 nb [8].

The last particle of the standard model that has not yet been found is the Higgs boson. The scalar Higgs boson, originally introduced into the theory of the SM as a complex doublet with four degrees of freedom to solve the problem of the none gauge invariance of the massive vector boson fields represents the remaining degree of freedom after the spontaneous symmetry braking generated gauge invariant mass terms for the W and Z bosons. Spontaneous symmetry braking has been arbitrarily made possible by the choice of sign of one of the parameters in the Higgs potential and although the standard model allows for spontaneous symmetry braking, it does not explain its origins. There are several good reasons to assume that the Higgs boson, should it exist, would have a mass that makes it accessible for the LHC. In the SM the cross-sections for $W^+W^+ \rightarrow W^+W^+$ scattering would grow beyond the unitarity limit if the mass of the Higgs were to be much larger than $\mathcal{O}(1\text{ TeV})$. The world combined electroweak data is also setting an upper limit for the mass of the Higgs when trying to fit the value of $\sin(\theta_W)$ varying m_H. Assuming that the Higgs boson does not exist and interpreting the SM as an effective theory into which the gauge boson mass terms, previously supplied by the existence of the Higgs boson are

put in by hand one would have to specify an energy scale λ up to which this sub-standard model should be valid and beyond which new physics should become evident. In this case one can again use use electro weak data, this time looking at the dependence of global fits on λ and it is found that the limits on the Higgs boson mass are translated into equivalent limits for λ which again would indicate that some effects would be expected at $\mathcal{O}(1\text{ TeV})$. If the Higgs boson exists experiments at the LHC will have a good chance to find it and study it's properties.

2.2 Supersymmetry

Beyond the discovery of particles inside the SM there are arguments suggesting deviations from the SM at energy scales of 1 TeV. One is the fact that m_H is modified from the bare mass appearing in the Higgs potential, by loop corrections and is actually quadratically divergent in the cut-off energy scale λ_{m_H} up to which loop corrections are included into the mass calculation. So there is a natural limit to the validity of the SM. It is not obvious why this limit should be accessible with the LHC, but in the SM m_H or rather the vacuum expectation value of the Higgs potential which is proportional to m_H^2, is related to m_W and if λ_{m_H} grows, not only m_H but also m_W grows. This phenomenon is called the "gauge hierarchy" problem. It could in principle be avoided by choosing the bare mass of the Higgs which is one of the parameters of the Higgs potential and thus a fundamental input parameter to the SM to cancel the large loop corrections. If the validity range of the SM is extended up to the scale of grand unified theories $\lambda_{m_H} = M_{GUT} = \mathcal{O}(10^{15}\text{ TeV})$ the value of the bare mass has to be specified to a precision of 10^{-24} if the SM is supposed to stay consistent with current measurements. This problem is often referred to as the "fine tuning" problem.

There are many suggested extensions to the standard model which fall mainly into two classes. They are either predicting compositeness of the gauge bosons such as technicolour or they predict a new symmetry like GUTs or supersymmetry. Technicolour does predict the existence of new particles with masses below 1 TeV and these should be measurable at the LHC.
Supersymmetry[1], which is going to be discussed in some parts in chapter 8, is a symmetry between bosons and fermions, coarsely summarised by the statement that it demands the number of fermionic and bosonic degrees of freedom in each of its multiplets[2] to be equal. This leads to a large new particle spectrum in which each vector boson of the SM is complemented with a Majorana fermion as its super partner and for each helicity state of a fermion a new scalar particle (sfermion) is postulated. In addition to this the Higgs sector of a SUSY model consists of at least two Higgs doublets. The potential danger of the two fermionic Higgsino states developing triangular loop anomalies is canceled by the Higgsinos having hypercharge +1 and -1. Besides an abundance of new particles, which in itself is a rather unpleasant feature, the MSSM has several attractive properties. One of the problems of minimal grand unified theories is, that the evolution of the three coupling constants for electromagnetic, weak and strong interactions does not unify them at one

[1] Supersymmetry here always denotes the minimal supersymmetric model

[2] each particle of the SM is "promoted" into a multiplet being given a supersymmetric partner to balance the degrees of freedom.

energy. This problem does not appear in a supersymmetric GUT. Also the prediction of a too small proton lifetime from a minimal GUT is corrected by supersymmetry. In an unbroken supersymmetric theory the super partners would be degenerate with their partners from the standard model. This is obviously not the case, otherwise these particles would have been observed. Therefore SUSY has to be broken by some mechanism that maintains one of it's nicest properties, the cancelation of the quadratic divergences in the Higgs mass which were achieved by canceling each loop in the self energy diagrams by the appropriate loop of the super partner which would have equal and opposite sign contributions. The complete cancelation can not be maintained in broken SUSY but the remaining divergence is not quadratic but logarithmic in the cutoff scale and proportional to the mass difference between the particles of a SUSY multiplet. If it is assumed that fine tuning is at least an unaesthetic process the masses of the squarks, to be more precise this should be the mass of the stop, super partner of the top quark since this is the dominant contributer to Higgs self energy terms, should not significantly exceed 1 TeV and therefore be observable at the LHC. Coming back to the breaking mechanism for SUSY the condition of not degrading the behaviour of the Higgs self energy terms is a very weak one leaving many possible terms with many free parameters and very little predictive power. It is a plausible but more or less arbitrary choice which of these so called "soft breaking" terms are excluded from the model[3] and it is this choice which defines the **minimal** super symmetric model. One of the conditions restricting the number of breaking terms is the requirement that all sleptons and all squarks but the stop be degenerate. If R-Parity[4] is a good quantum number the lightest supersymmetric fermion, a neutralino, must be stable and should have a mass below 1 TeV, again being observable at the LHC.

All of the above arguments indicate that exploring the energy region around 1 TeV could lead to exiting new discoveries and may change the standard model of particle physics.

2.3 Technology

One less physically fundamental motivation which plays a major role in the long period of hardware development and tests is the the possibility to contribute to the huge amount of technological progress that needed and still needs to be made before the ATLAS detector can be built. This technological progress is much closer to being relevant or applicable for "the rest of the world". Some of the technological foci of ATLAS tracking detectors are on the design of detection elements, mechanical precision structures and readout electronics. The control and prediction of the mechanical and thermal properties of detector modules as well as the development of 0 CTE[5] composite materials and measurement techniques for the detection and understanding of small complex distortions are only a few tasks in this large project and have been described in this work. Another point of technological relevance is the development of the accelerator itself. Besides providing many other challenging problems the construction of the LHC certainly represents a project on the frontier of super conducting magnet design. The LHC is the only viable suggestion for an accelerator exploring the energy scale above 1 TeV in the near future, since linear e^+e^--colliders for this energy range today still represent unsolved technological problems.

[3] many are excluded already by phenomenology

[4] see chapter 8

[5] Coefficient of Thermal Expansion

Chapter 3

The Experiment

Looking at the LHC and ATLAS as parts of a project aimed at the investigation of pp-collisions it seems to be appropriate to consider both collider and detector as functional units of one experiment. Both components have to be closely matched to each other in order to exploit the full potential of physics subjects. As an example one could mention that the detector has to be able to cope with the radiation field caused by the collider and the collider has to be able to provide suitable beam parameters in the collision points introducing as few components as possible into the experimental hall.

3.1 LHC

The LHC is a proton proton collider with a center of mass energy of 14 TeV. It is going to be built in the existing LEP tunnel of 27 km circumference, using the LEP injection system and is going to have two collision points. It is supposed to reach a design luminosity of $\mathcal{L}_{design} = 1.0 \cdot 10^{34}\text{cm}^{-2}\ \text{s}^{-1}$ which is much larger than any luminosity achieved in a collider so far. At the design luminosity the interaction of two proton bunches will result in approximately 23 minimum bias interactions, each of which will produce 3-6 charged tracks per unit of rapidity, leading to a total of 15-30 tracks in the acceptance of the ATLAS inner detector. When compared to the TEVATRON, todays most powerful hadron collider or other hadron colliders the bunch length of 7.5 cm is very short, resulting in a total luminous region of approximately 40 cm length. The transverse beam size will be very small and averaged over its two axis the beam will be 16 μm in radius, allowing a good beam constraint in track finding. With 25 ns the time between two interactions is very short, presenting a big challenge to the detector electronics which will have to cope with enormous data rates. At the design luminosity, assuming 10^7s of operation per year the LHC will provide an integrated luminosity of $10^5\ \text{pb}^{-1}$ per year. Due to the very high interaction rate detectors for the LHC will be submitted to an intensive radiation field. In the ATLAS inner detector the yearly fluences of neutrons will be in the range of $2\text{-}8{\cdot}10^{13}\text{cm}^{-2}$ and the radiation dose will be 2-35 kGy per year, depending on the position in the inner detector cavity.

3.2 ATLAS

ATLAS is one of the two general purpose experiment approved for the LHC. ATLAS is described in great detail in its technical proposal [3] and only a coarse description of the major features relevant to this work is given here. ATLAS is now near to the end

of its planning phase and modifications of the design given in [3] are small. The major differences are:

- The baseline option for the forward discrete tracking in the inner detector are now silicon strip detectors.
- The layout for the inner detector has been slightly modified to accommodate the change of technology in the forward section and to achieve more flat distributions of the number of hits contributed by each subdetector
- The inner solenoid coil has been reduced slightly in length to remove a problem of excessive radiation length in front of the calorimeters in the barrel forward transition region.
- The super conduction toroid magnets will be reduced in size and field for cost reasons.

Figure 3.1 shows an isometric view of ATLAS with a quarter section cut out.
The most important dimensions are shown in table 3.1. ATLAS has a typical collider de-

component	Radius [m]	Length [m]	η-coverage
barrel muon detector	11	26	$\|\eta\|<1.4$
end-cap muon detector	11	2·8	$1.1<\|\eta\|<2.8$
barrel had-calorimeter	4.25	12.2	$\|\eta\|<1.0$
end-cap had-calorimeter	2.25	2·2.5	$1.5<\|\eta\|<3.2$
barrel em-calorimeter	2.25	6.42	$\|\eta\|<1.4$
end-cap em-calorimeter	2.25	0.63	$1.4<\|\eta\|<3.2$
forward em-, had-calorimeter	integrated in end-cap		$3.1<\|\eta\|<4.9$
inner detector	1.15	6.8	$\|\eta\|<2.4$

Table 3.1: Dimensions of the ATLAS subdetectors.

tector geometry. On the outside a muon detection system (figure 3.2) with approximately 1.3 million readout channels including trigger chambers is situated in a toroidal field of approximately 4T giving a momentum resolution for muons of 1TeV varying between 5% and 10%. Each muon track passes 3 measurement stations, each consisting of 2 separated half stations. Each half station consists of 3 or 4 layers of detection elements. The detection elements are pressurised, position monitored drift tubes in most of the covered area and cathode strip chambers in the areas with very high particle rates. Attached to the outside of each measurement station are two trigger chambers. The trigger chambers are either resistive plate chambers in the lower rate regions or thin gap chambers in the highest rate regions. Further inside the detector the hadronic calorimeters are found. Like all calorimeters they are separated into three subdetectors which are the barrel, end-cap and forward calorimeters. The forward calorimeters, both hadronic and electromagnetic, are liquid argon based and integrated into the cryostats of the end-cap calorimeters. The barrel and extended barrel region of the hadron calorimetry is an iron-scintillating tile calorimeter with 3 radial samplings, an active depth, including the em-calorimeter in front, of 9.5 λ_{abs} and an energy resolution of $\sigma(E)/E = 45\%/\sqrt{E/GeV} \oplus 1.5\%$. The granularity is going to be $\Delta\eta \times \Delta\phi$=0.1×0.1 and 0.1×0.2 in the last radial layer. In the end-cap region the hadronic calorimeter is based on liquid argon with a tungsten absorber because of the higher radiation levels. With a total active depth of 10 λ_{abs}, subdivided into four longitudinal samplings it is supposed to reach an energy resolution of $\sigma(E)/E = 50\%/\sqrt{E/GeV} \oplus 3\%$

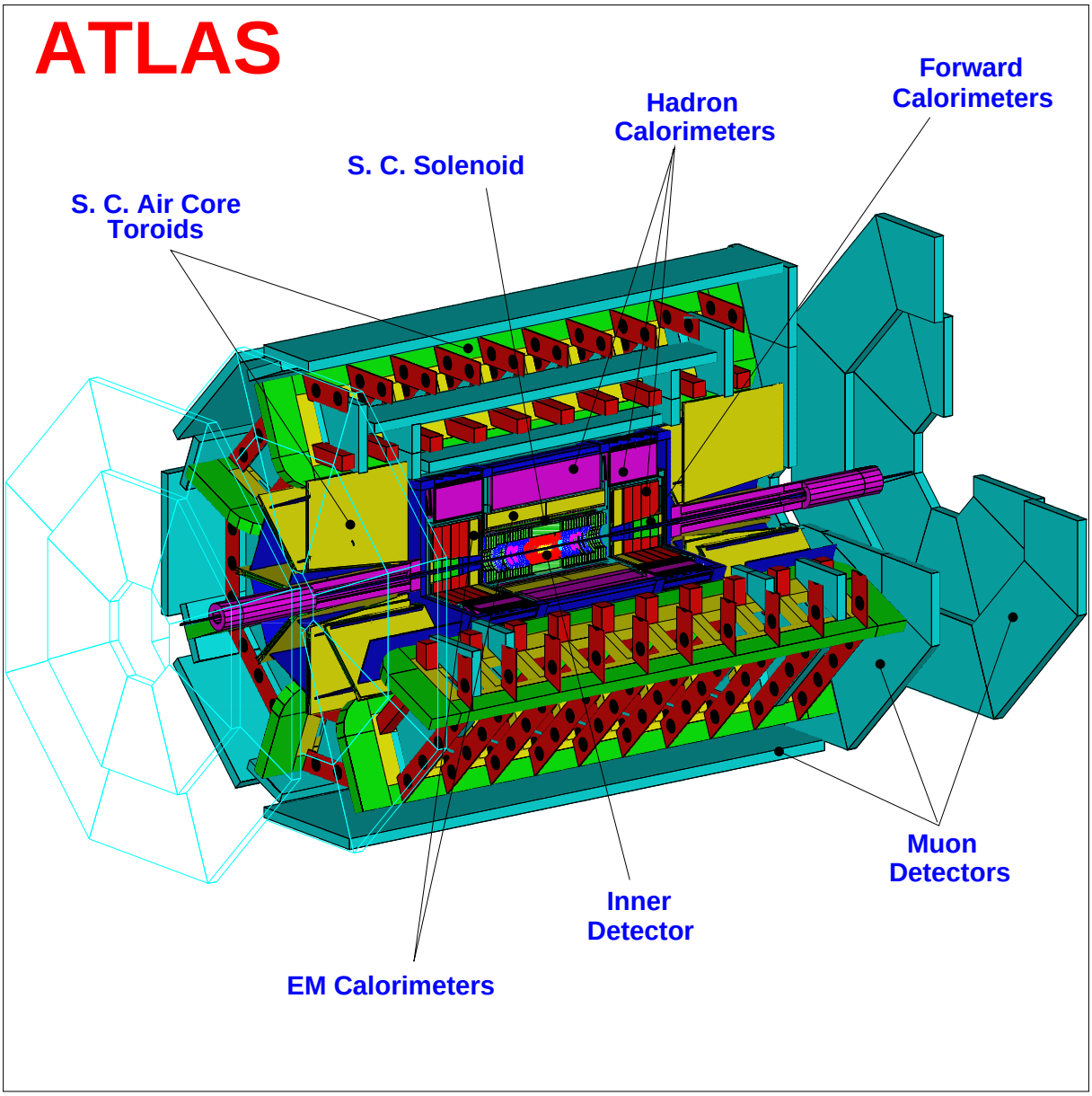


Figure 3.1: An isometric view of the ATLAS detector with a quarter section cut out [9]. (overall diameter 22 m.)

with a segmentation of $\Delta\eta \times \Delta\phi$=0.1x0.1. Going into the forward region, the granularity is reduced to 0.2×0.2 and the resolution is $\sigma(E)/E = 100\%/\sqrt{E/GeV} \oplus 10\%$.

All electromagnetic calorimetry is based on liquid argon. In the barrel and end-cap region an accordion type geometry with lead absorbers is foreseen. The total thickness of these calorimeters is 26.5 and $28X_o$ respectively. The granularity of the em-calorimeters varies between $\Delta\eta \times \Delta\phi$=0.003×0.1 and 0.05×0.025 depending on the longitudinal position. The barrel part is segmented into 3, the end-cap into 4 longitudinal samplings. The energy resolutions measured in test beams are:

$$\textbf{barrel}: \quad \frac{\sigma(E)}{E} = \frac{(9.99 \pm 0.29)\%}{\sqrt{E/GeV}} \oplus \frac{(0.288 \pm 0.017)}{E/GeV} \oplus (0.35 \pm 0.04)\% \quad (3.1)$$

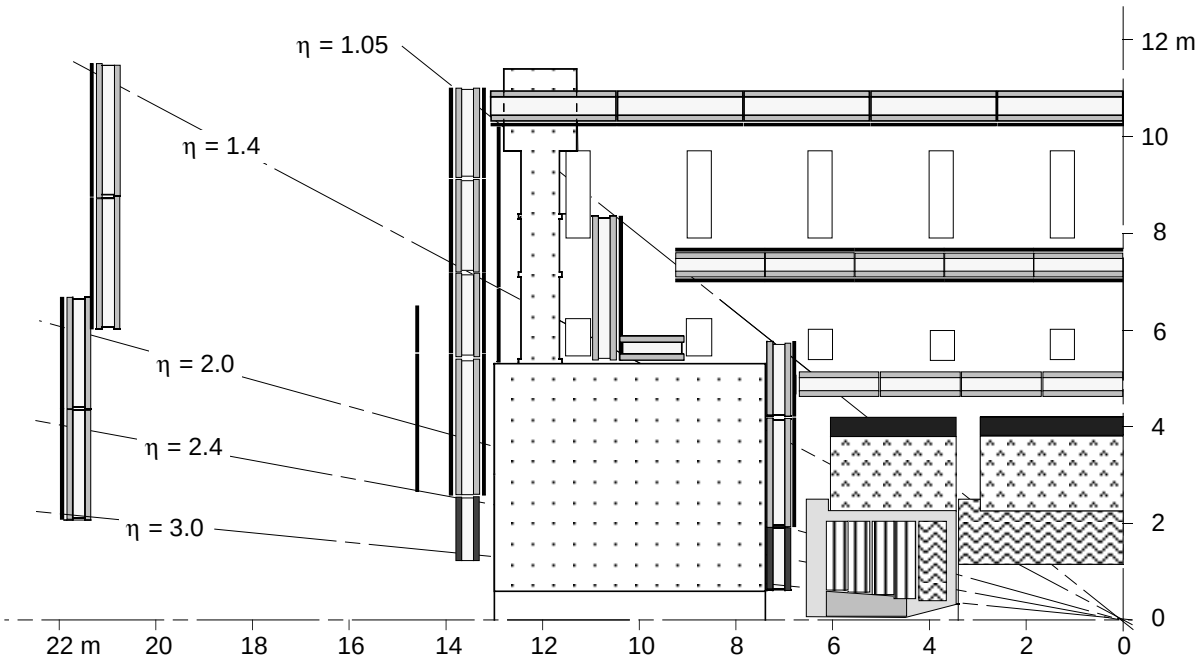


Figure 3.2: rz-quarter section of ATLAS. The muon precision chambers are shown in dark grey, the attached trigger chambers are indicated as solid black lines.

$$\mathbf{end-cap}: \quad \frac{\sigma(E)}{E} = \frac{(10.7 \pm 0.3)\%}{\sqrt{E/GeV}} \oplus \frac{(0.51 \pm 0.02)}{E/GeV} \oplus (0.30 \pm 0.04)\% \qquad (3.2)$$

For a forward region prototype using brass absorbers the resolution was found to be $\sigma(E)/E = 35\%/\sqrt{E/GeV} \oplus 4\%$.
Directly in front of the electromagnetic calorimeter a preshower detector, again using liquid argon will be placed. It will be integrated into the cryostat of the calorimeter and will have two samplings measuring ϕ and η with a granularity of 0.156 and 0.004 respectively.

3.2.1 Inner Detector

Inside the calorimetry the inner detector is situated and is described here in some more detail. Figure 3.3 shows an rz cross-section through the inner detector as shown in [3]. The same layout was used for the simulation work presented in chapters 7 and 8. The concept behind the tracker design is to combine the different properties of the continuous tracking detectors (TRT[1]) with those of the discrete detectors (silicon strip and pixel detectors, MSGC and GaAs strip detectors). The continuous detectors provide a large number of hits per track and therefore have good pattern recognition properties. The TRT performance suffers from high occupancies, in particular in the low radius layers because it is a projective detector giving only one dimensional measurements. The discrete detectors have only very few layers but high resolution and small granularity, giving two dimensional measurements. The physics and environmental requirements for the silicon detectors are described in chapter 4. The resolutions, granularities and total area of the subdetectors in the inner detector can be found in table 3.2 The precise measurements of the discrete detectors with their low occupancy contribute at least 5 hits to each track. The hits contributed from the TRT vary between 18 as a minimal value in the transition

[1]Transition Radiation Tracker

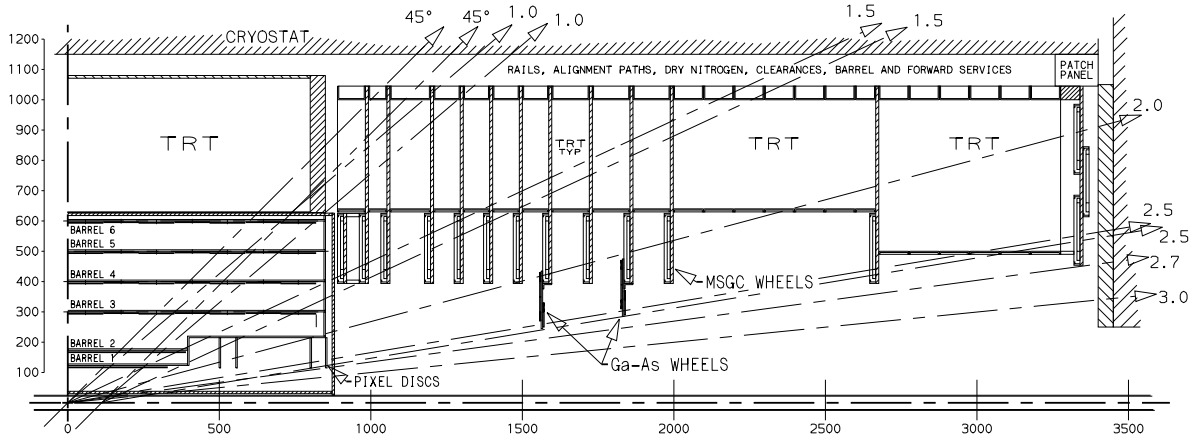


Figure 3.3: rz cross-section through the inner detector as used for the simulation work presented in chapter 7 and 8

System	Area (m^2)	Element size (μm)	Resolution (μm)	η Coverage
Pixels	1.38	50×300	$\sigma_{r\cdot\phi} = 14$ $\sigma_{z,R} = 87$	± 2.5
Silicon	41	75×12cm $\alpha_{stereo} = 40mrad$	$\sigma_{r\cdot\phi} = 15$ $\sigma_z = 770$	± 1.4
GaAs	3.3	50×7.6cm	$\sigma_{r\cdot\phi} = 10$ $\sigma_R = 1200$	2.0–2.5
MSGC	51	200×16cm	$\sigma_{r\cdot\phi} = 35$ $\sigma_R = 1800$	1.4–2.5
TRT		4 mm diameter straw tube	$\sigma_{r\cdot\phi} = 170$	± 2.5

Table 3.2: Main parameters of the subdetectors in the inner detector. The resolutions for the silicon detectors include charge sharing.

region between barrel and forward ($\eta = 0.9$) and the maximum of 60 at $\eta = 1.9$. The distribution of hits from the TRT can be seen in figure 7.1 on page 115. The silicon strip detectors and the MSGC use small stereo angles to achieve z or r resolution. The TRT is fully projective over the length of the drift tube and has only single hit capability[2]. The TRT in addition to its tracking function uses the good X-ray detection capabilities of its Xe-CO_2 based gas mixture to measure transition radiation photons produced in the radiator[3] separating the loosely packed drift tubes. Thus the TRT can be used to separate electrons from hadrons. For results from a prototype TRT see [10]. The GaAs detectors have been included into the design due to their radiation hardness against neutron irradiation. New measurements of charge collection efficiencies after charged particle irradiation [11] indicate a problem and the use of GaAs is under discussion inside ATLAS.

One of the central problems of this tracker is to find an optimal way of combining the information from the two different subdetector types, continuous and discrete tracking detectors, which has been addressed in chapter 7. The different subdetectors and their

[2]Only the drift-time of the first hit arriving in the tube inside a 30 ns window is measured

[3]the radiator will either be a ROHACELL foam or PE-foils

technology are described in great detail in [3]. For the initial low luminosity operation of LHC an additional b-physics layer at a radius of 4 cm is foreseen. The layer will probably be similar to the innermost pixel layer of figure 3.3 but will utilise charge sharing read-out to enhance the position resolution. Due to the quadratic increase of charged particle radiation with inverse radius, this layer will not survive the full operational period of ATLAS.

Chapter 4

Silicon Tracker Prototype

4.1 Specifications

The basic requirements for the ATLAS inner detector have to be translated into specifications for the tracker modules and support structures to serve as guidelines in the design. This is attempted here for the barrel SCT in the layout described in section 3.2.1 based on the work presented in [13] and [14]. Most of these specifications have to be applied to the combination of support structure and module since it is their combined functionality which defines the performance of the module. In the design described below the cooling system is physically a part of the support structure whereas the constraints of the operation temperature specified for this system are constraints on the temperature of the module. In this sense the module design is synonymous for the design of a functional unit of support and module. Aspects of the global mechanical and thermal system such as the alignment of a complete barrel or the temperature variations from barrel to barrel have not been addressed in this study. The measurement resolutions are derived from the requirements on momentum resolution and pointing accuracy.

- $r\phi$ -Resolution
  A heavy W' of $m_{W'}$ =2 TeV produces electrons of ${p_t}^e \approx 1$ TeV. The sign of charge of these has to be measured requiring a momentum resolution of $\Delta p_t/p_t < 30\%$. Using analytical calculations of the detector resolution in the proposed geometry a strip pitch of 75 μm with a resolution of $\sigma_{r\phi} = 21.5$ μm [1] is found to be suitable. It has to be noted that the strip pitch has been forced to be equal for all barrel silicon strip detectors to simplify the detector and module design and for the outermost silicon layers a larger pitch would have been possible.

- Z-Resolution
  The considerations determining the Z-resolutions are the mass resolution of Z', J/Ψ, B^o and K_s^o , the track shower matching and the matching of μ tracks from the inner detector to the muon system. These require a Θ-resolution of 2 mrad and thus a Z-resolution of $\sigma_z < 300$ μm .

To not degrade the above resolutions the total stability of the detector position during operation should be better than about 1/3 of the resolutions.
Other considerations concern the reliability and radiation damage of the modules and are discussed below.

[1] No charge division readout is assumed

- Operating Temperature
  The most stringent requirement on the operating temperature of silicon detectors comes from the shot noise of the high leakage current of a detector at the end of LHC operation. To reduce this noise the detector has to be operated cold. The second requirement arises form the increase in depletion voltage after irradiation through the high neutron fluences and charged particle doses in the inner detector cavity [3] and long term reverse annealing. This requires the detector to be constantly at low temperatures and not just during operation as can be seen by looking at the different scenarios described in the depletion voltage plot of figure 4.1. For the current the warm up does not make a significant difference. As figure 4.1 indicates, both leakage current and depletion voltage seem to not improve significantly at temperatures below -10 ^oC and temperatures below 0^oC seem to be possible ![12].

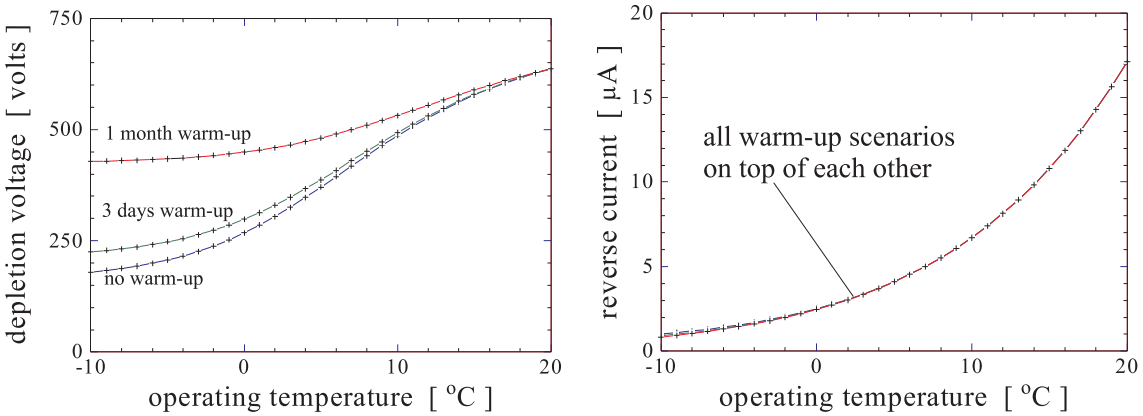


Figure 4.1: Depletion voltage and leakage current of a silicon detector of 300 μm thickness after irradiation expected for 10 years of LHC operation at a radius of 30 cm. Three warm up scenarios with no, 3 days and one month of warm up to romm temperature are shown. [12].

At these temperatures a dry gas atmosphere is required to avoid condensation.

- Temperature Stability with time
  The stability of the temperature over a run period has it's most stringent limit in the deflections resulting from the temperature change. These are very dependent on the detailed module geometry and can not easily be specified.

- Temperature Homogeneity
  Two factors limit the maximum temperature change across a module. First the resulting intrinsic mechanical stress and deflections have to be such, that the module is safe and geometrically stable.
  Secondly a temperature difference between diodes of the silicon detector leads to variations in the leakage current and thus the noise and power dissipation of the detector. Here in particular the effect of "thermal runaway", where the exponential increase in power dissipation of a silicon detector with temperature can lead to an instability, is of concern. Finite element simulations of this effect have shown that its critical parameters are the maximum distance between a point of the silicon detector and the cooling tubes, as well as the absolute temperature of the detector and its environment.
  A ΔT_{max} of 5 ^oC has been chosen as the maximum ΔT across a module. However

separate tests of the thermal and mechanical stability have to show that the specific module geometry is suitable for this ΔT.

- Power cycles
  A minimum number of 100 cycles between operating conditions and installation conditions[2] has to be tolerated by the module without any observable degradation. This point is supposed to cover fatigue of electrical and mechanical connections and sets a limit to the maximum stress induced by the changes in temperature and humidity.

- Worst Case Conditions
  The module is supposed to tolerate a temperature swing between 50 ^oC and 90 % relative humidity to operating conditions and stay within the specifications. This is supposed to represent accidental changes of humidity and temperature for example during transportation.

- Radiation Length
  A maximum thickness of the module of 1.7 % X_o averaged over one module at normal incidence has been chosen. This will determine the momentum resolution for low p_t tracks at $p_t < 30$ GeV and it will influence the performance of the electromagnetic calorimetry.

- Radiation Hardness
  One of the most important features of the LHC environment are the extreme charged particle dose and the very large neutron fluences to which the detectors are exposed. Because there is a considerable uncertainty in the total neutron fluences and to allow for at least a factor two safety, all functional materials in the inner detector should not show any significant degradation in their relevant properties up to a 1 MeV equivalent neutron dose of $2 \cdot 10^{14}\,\mathrm{cm}^{-2}$ or the design should incorporate the degradation effect if it is well understood. The same should be true for a charged particle and photon dose of 600 KGy$\cdot\frac{11cm}{R^2}$ where R is the transverse distance to the beamline. The maximum radiation levels have been taken from [3] and a safety factor of two has been applied. These two requirements are of crucial importance and not yet fully investigated. For many materials such as electrically or thermally conductive glues, thermally conductive greases and structural glues the influence of radiation exposure is not clear. Intensive studies of all materials suggested for use in the inner detector cavity concerning their radiation hardness are ongoing.

Due to the large number of modules in the SCT ($\approx$ 1500 in the barrel SCT) the assembly yield and time, testability and installation procedure as well as the possibilities for repair and maintenance also play an important role in module design, but can not easily be translated into quantitative specifications. Points in the design addressing these problems will be pointed out as the module realised for this work is described. Other fundamental choices for module geometry and techniques used to build a module have been made and are listed here.

- The module should use back to back single sided detectors of 300 μm thickness.

- The stereo angle between the detectors should be 40 mrad.

[2] 20 ^oC and 60 % rel. humidity

- The cooling system is supposed to be operating at subatmospheric pressure for safety reasons (i.e. leakless cooling).

- The silicon detector dimensions are close to 50·60 mm^2.

- Two detectors should be concatenated into a pair to achieve a diode length of 120 mm in the z direction.

- A module should serve two pairs of concatenated detectors[3].

4.2 Module Design

In this section the design of an barrel silicon strip module and the corresponding support structure is presented. Work on the design started in 1993 and some specifications valid at this time differ from those described above. Also some of the above choices for preferred geometry and techniques were either different or not yet made. The main differences are described below.

- The maximal radial extent of the barrel SCT was 100 cm leading to a detector pitch of 200 μm .

- The default module configuration used one layer of single sided detectors but the requirements for diode length and number of read out channels per module were identical, doubling the area covered by one module.

- The cooling system was not restricted to subatmospheric pressure operation.

In particular the transition from single sided to double sided modules disfavours the application of some of the geometrical principles used in this design and the currently suggested strip pitch of 75 μm is too small for the detector concatenation method described here.

The design of the module focussed on four considerations, while trying to stay inside the detectors specifications.

1. maximum modularity[4]

2. 100% geometrical acceptance

3. minimal input capacitance to the front end electronics

4. minimal cost

Intrinsic to all detector components are of course the requirements of radiation hardness. Allthough this issue has not been studied in this work it should be said that a systematic study of the radiation hardness of glues and thermal compounds has recently been initiated by Dortmund university and the importance of this point is well understood withing the ATLAS SCT community. Figure 4.2 shows two modules mounted in $r\phi$ -overlap.

[3] this is equivalent to one double sided pair

[4] to give best reworkability, yield and repairability

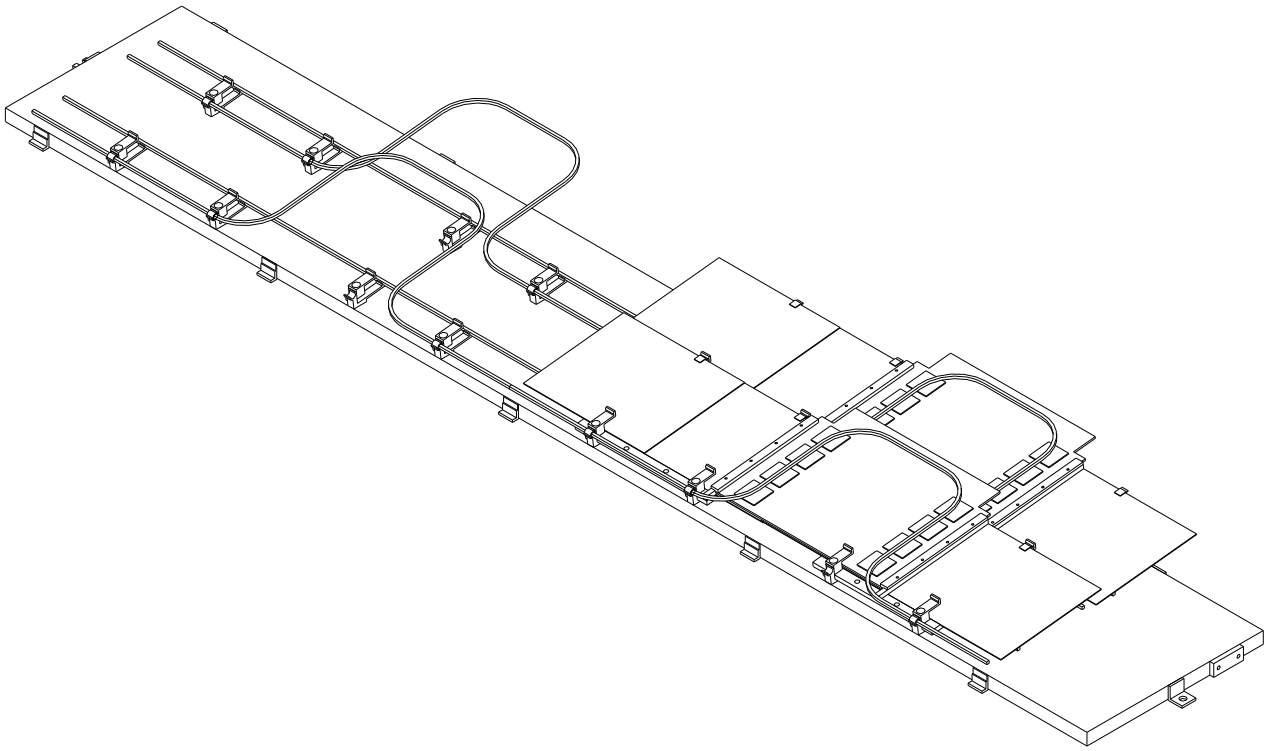


Figure 4.2: Two modules mounted in $r\phi$ -overlap on the support structure with both electronics cooling pipes in the "closed" position, but two module positions still unoccupied.

4.2.1 Detector-Detector Concatenation

In order to maintain a 100% geometrical efficiency of the modules the active areas of the two detectors forming an electrical pair have to overlap in the z direction. If the concatenation is done in the conventional way with wire bonds the strip side of one detector is glued to the backside of the other. In this geometry the large bias voltage of up to 400 V is applied across the thin glue layer connecting the detectors and the bond wires will have to pass close to this high field region. To avoid this problem the detectors can be concatenated with a bump bonding process. In this case the detectors are concatenated with their strip sides facing each other. The intrinsically higher costs of a standard solder bump bond compared to a wire bond can be avoided if a polymer bump bonding process such a PFC[5] process is used which is suitable for a pitch of 200 μm . This interconnection also has a smaller capacitance and is less fragile than the bond wires.

4.2.2 Detector-Electronics Connection

Special efforts were made to achieve a high degree of modularity for this module. This means that a maximum number of components are demountable and can be reused. In particular a demountable connection between the silicon detectors and the readout electronics using conductive rubber has been implemented and is sketched in figure 4.3 in an R-Z cross-section. The carbon fibre connection bars 1 and 2 are glued to the strip side of those silicon detectors of a pair facing upwards. An insulating epoxy resin suitable

[5]provided by EPOTEK, Billerica, USA

for stencil application is used to achieve even glue thickness. The process takes place in a jig ensuring that the thickness of the glue layer is stable during curing. In this way the problem of the conducting carbon fibre bar short cicuiting the detector strips can be avoided, even if the detector does not have a passivation layer. The height of connection bar 2 is dimensioned such that the compression of the conductive rubber reaches it's nominal value when the connection is fully closed. The connection is closed by screwing nuts onto the bolts extending from the connection bars through the connection lid. There are four screws in bar 1 and three in bar 2 as can be seen in figure 4.9. The conductive rubber between the two connection bars connects the pattern of the strips on the silicon detectors to the matching pattern on the underside of the electronic board. On the board the signals are routed through the board to the top surface.

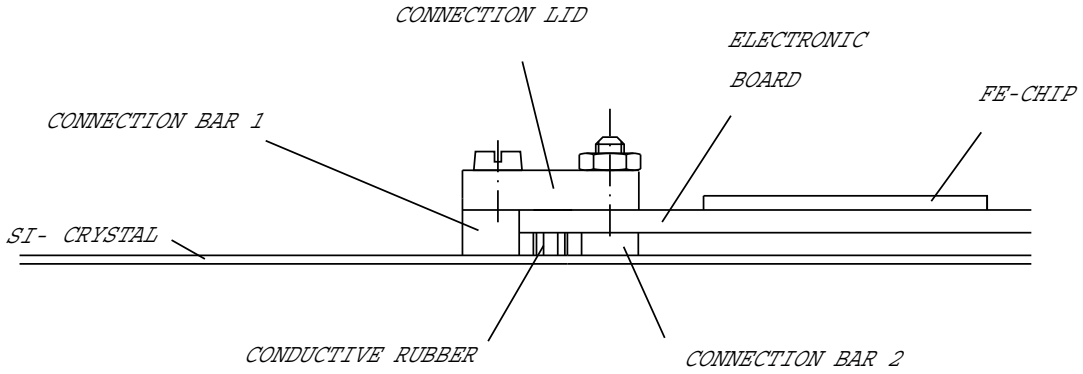


Figure 4.3: Sketch of conductive rubber contact.

4.2.3 FE-Chip Connection and Cooling

The signals arriving on the top of the electronic board are fanned into the area of their corresponding readout chips which are then connected to the signal lines with a chip-on-board flip chip process. This allows the reverse side of the FE-Chips to be directly cooled by a pipe running across the board in a u-shape. The pipe is pressed to the FE-Chips with miniature leaf springs that could finally be produced in an injection moulding process from a low radiation length polymer. The cooling pipe itself is a square cross-section aluminium pipe with 1 mm inner dimension and 0.1 mm wall thickness. It is pivoted in three bearings per module and can be rotated around its z-axis, as can be seen in figure 4.4. The pipe is rotated onto the board after the four modules of one row have been mounted onto the support structure. Then the springs are closed to fix the pipe to the front end chips. In order to obtain better contact between the pipe and the chips and to allow for the expansion of the aluminium pipe during cool-down a thermally conductive grease is used on the pipe-chip interface.
This cooling scheme has the advantage that it removes the heat from the front end chips directly and does not depend on good thermal conductivity of the electronic board material which allows the use of cheaper less conductive materials such as standard glass fibre and kevlar fibre materials. It also introduces a minimal heat flow through the electronic board and insulates the silicon detectors from the heat produced in the front end chips. The connection scheme offers a short signal path from the diode to the front end chip giving small input capacitance. Furthermore it avoids the use of fragile wire bonds replacing them with flip chip bonds and a conductive rubber contact.

4.2.4 Detector Cooling

The silicon detectors have to be operated at lower temperatures than the FE-electronics which could tolerate operation at room temperature. The detectors produce less than 20% of the heat generated by the module even at the end of the LHC running[6]. This is the reason for cooling the detectors with the incoming and the electronics with the return flow. The cooling pipe is of the same type as for the electronics cooling. It is equipped with Aluminium spacers which allow the straight pipe to follow the up and down pattern of the detectors resulting from their overlap. As for the FE-Chip cooling the interface between the pipe and the silicon is improved with a thermally conductive grease. The cooling pipe is supported by flexible arms in the lower adapter shown in figure 4.4 and is pressed with small tension against the silicon detectors. The flatness of the pipe and the correct height of the spacers are crucial for the performance of the system.

4.2.5 Positioning Scheme

Figure 4.4 shows an X-Y cross-section of two modules overlapping each other in $r\phi$ in the region where the modules are fixed to the support structure. The cooling pipe for the front end chips of the left module is rotated into the mounting position and the leaf springs holding down the pipes onto the front end chips are not shown.

Each detector pair is separately held to the support structure via a carbon fibre rack (see figure 4.5) which is glued to the long side of the pair in the region of the detector overlap. The racks extend along the side of the detector pairs from under the electronic board to the second[7] positioning pin. Because the electronic board should have minimal size and the conductive rubber contacts require it to overlap with an upward facing detector of each pair the racks for the two pairs of each module have different length. Each rack has two precision, miniature brass bearings into which tapered dowel pins from the support structure fit. The pins[8] have a central tapped hole and are held by an adapter (lower part) which is glued to the support structure. The tapped hole is used to screw down the upper part of the adapter onto the rack. The upper adapter has two functions. Firstly it presses the rack onto the pins and secondly it incorporates a flexible hook used to fix the free edge of the neighbouring module as can be seen in the enlargement of figure 4.4. Onto the free long edge of each detector a small plastic nose is glued which clicks into the flexible hook and thus restricts the upward movement of the silicon detectors. When the detectors click into these hooks they are slightly flexed keeping them pressed against the hooks.

4.2.6 Radiation Length

Table 4.1 shows a breakdown of the radiation length calculated for a module. In addition to the 0.82% X/X0 from the module, the support structure contributes 0.35% X/X0 giving a total of 1.17 % X/X0 for the complete solution.

---

[6] The power dissipated in the silicon detectors grows with increasing radiation damage as the required bias voltages and the leakage currents grow.

[7] upper in figure 4.5

[8] not directly visible in figure 4.4

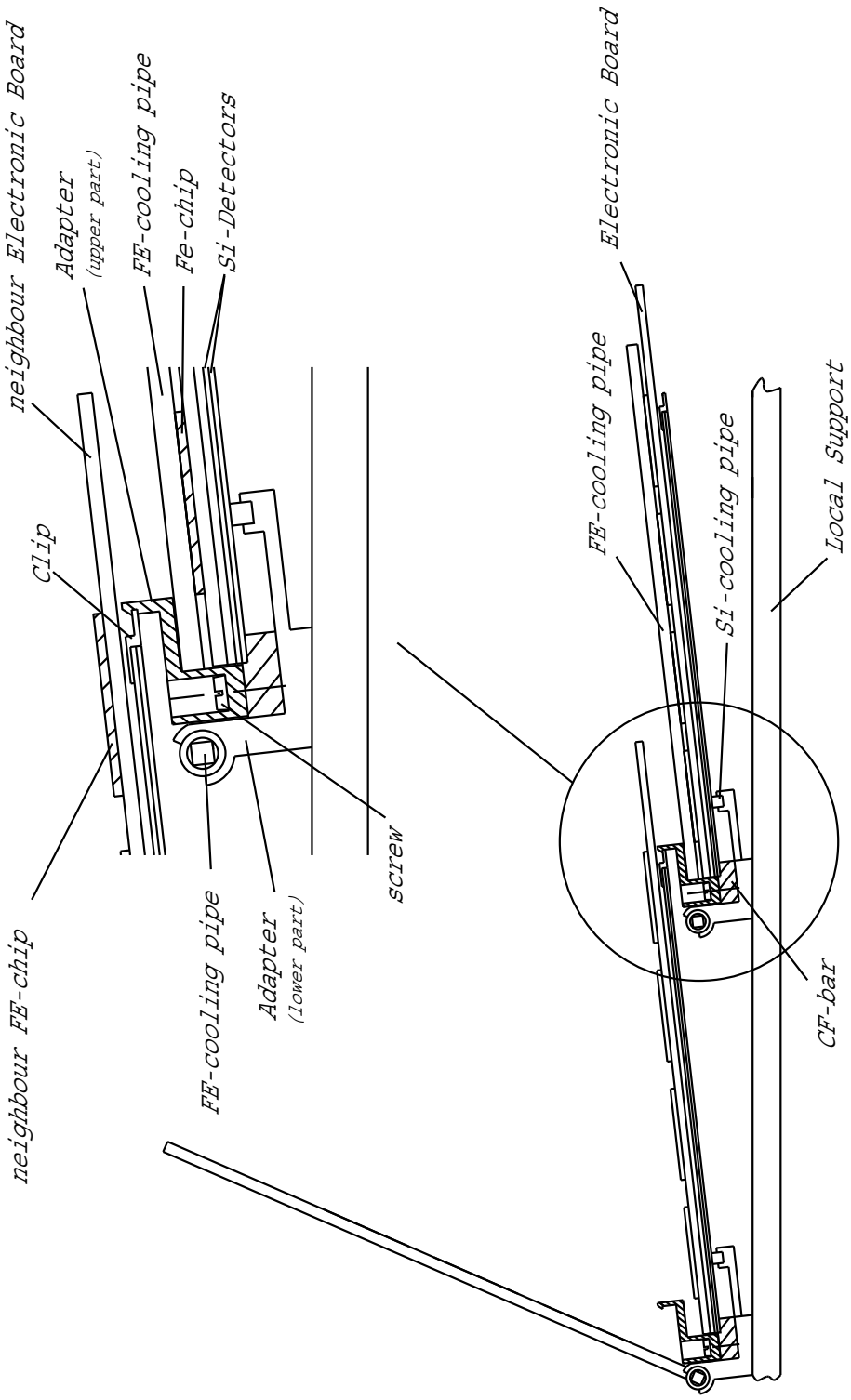


Figure 4.4: X-Y cross-section through two modules mounted in $r\phi$ overlap. The left electronics cooling pipe is in mounting position. The magnified region shows how the upper part of the adapter is used simultaneously to hold down a rack of its own detector pair and to fix the clip of the neighbouring pair. It is also shown how the flexible arm of the lower part of the adapter supports the silicon-cooling pipe and presses it against the detectors.

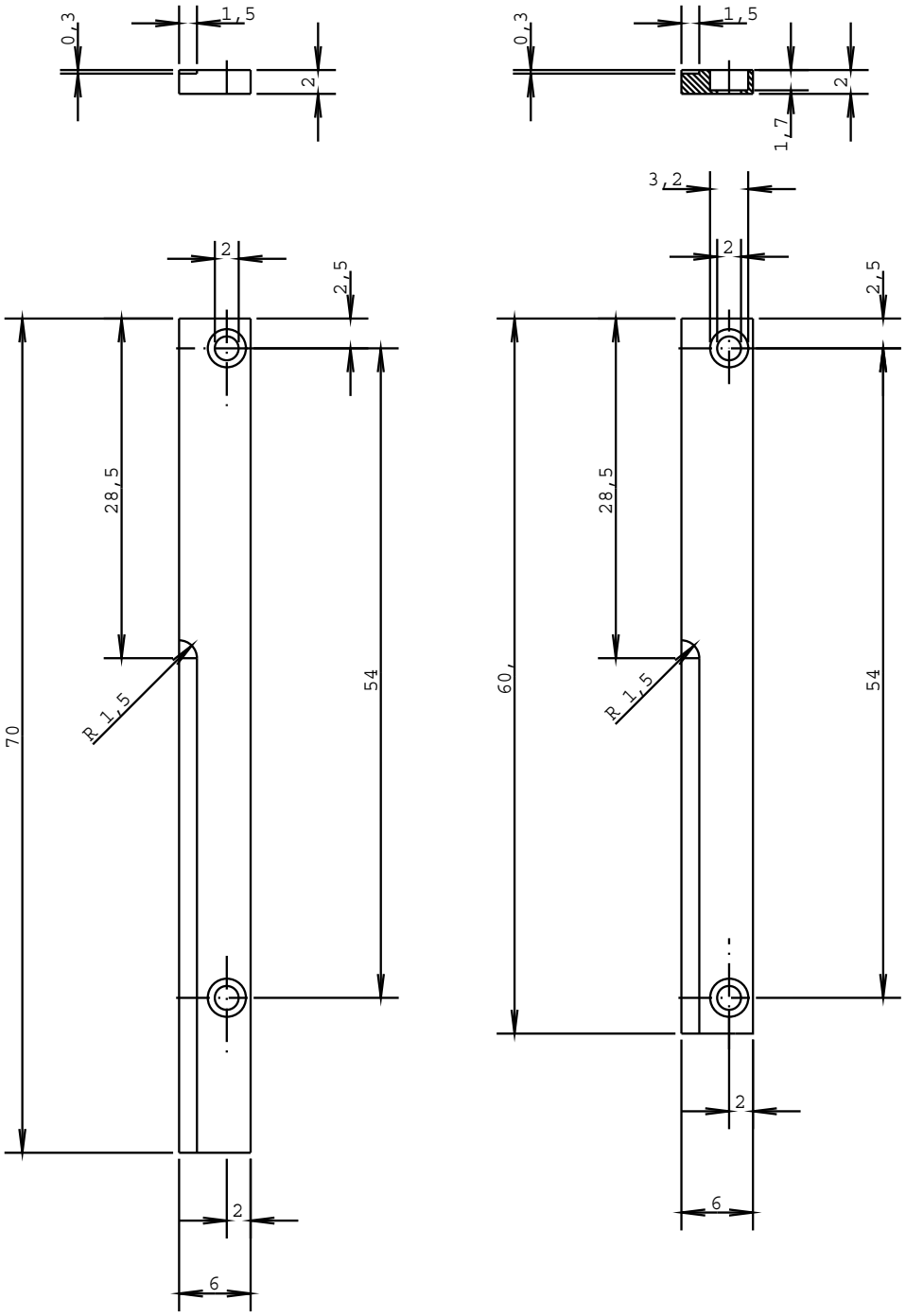


Figure 4.5: Workshop drawing of the two racks supporting the two silicon detector pairs of each module. The lower of the two detectors of one pair extends from 28.5 mm of the top of the rack downwards. The higher detector overlaps it by 2 mm. The racks have indentations of 0.3 mm for the lower detectors and oversized counter bored holes receiving the brass bearings for the support structures positioning pins.

component	material	X0 [mm]	X/X0 [%]
detectors	silicon	93.6	0.31
glue	epoxy	300	0.0007
racks	CFC	250	0.052
board (lines)	copper	14.3	0.185
board (bulk)	kevlar	287	0.13
contact bars	CFC	250	0.014
contact lid	CFC	250	0.027
conductive rubber (pins)	brass	14.3	0.003
conductive rubber (matrix)	silicone rubber	200	0.05
screws	brass	14.3	0.014
FE-chips	silicon	93.6	0.023
RO-controller	silicon	93.6	0.0013
adapter	CFC	250	0.018
clip	CFC	250	0.0005
cooling pipes	Al	89	0.024
coolant	water	361	0.0135
pipe bearing	GFC	194	0.0013
Total:			0.82

Table 4.1: Radiation length breakdown of a module. CFC=**C**arbon **F**ibre **C**omposite, GFC=**G**lass **F**ibre **C**omposite

4.3 Module Assembly

The process of assembly refers to all process steps necessary to create an operational module from it's elementary components. Some of the elementary components are prefabricated carbon fibre parts (8 parts of 5 different types). The other components are the individual silicon detectors, the electronic board which is assumed to be fully functional and tested and the conductive rubber strips which are precut and optically inspected. The upper part of the adapter mentioned in subsection 4.2.5 is also considered to be a part of the module.

4.3.1 Detector Pairs

The assembly process starts with the concatenation of detectors into pairs. The alignment of the detectors of one pair with respect to each other is in principle an issue that is under control of the bump bonding company and is usually done by a traveling microscopic optical system which requires either fiducial marks on the otherwise unprocessed backside of the detector or a bidirectional visual alignment viewing two sides of the alignment edge. In this case an alignment using the edges of the detectors is possible and helps to drastically improve the reliability of the process and reduces the cost. This is possible because the positioning accuracy of the detector pairs inside the module with respect to the mounting points is also achieved by mechanical alignment along the contour of the silicon detectors. The required accuracy of the cutting edge with respect to the detectors active elements has been achieved by two manufacturers. These manufacturers[9] have delivered detectors with a cutting accuracy of better than 5 microns. Although the required cutting accuracy

[9] Hamamatsu and Micron

increases the detector cost, the bump bonding of the detector pairs and the assembly of the module is less labor intensive and faster. Therefore it is also considered to be cheaper.

4.3.2 Mounting Racks

After the pairs have been manufactured and tested for their electrical properties the mounting racks are glued to their sides. This process is described below. In order to keep the demands on the manufacturing tolerances of the carbon fibre mounting racks small, the holes for their brass bearing have been oversized and all sides facing the detectors for later glue joints are dimensioned to stay away from the silicon by at least 0.05 mm. The process of gluing the racks to the detector pairs takes place in a high precision jig. Into the jig which is shown in figure 4.6 duplicates of the jig positioning pins (13a,b in fig. 4.6) from the support structure are glued. The holes for these pins and the two sets of alignment pins for the silicon pairs (1-4 and 9a,b in fig. 4.6) were machined using the same CNC machine that manufactured the jigs for the application of the positioning pins onto the support structure. This ensures that the pins have exactly the same relative positions. The gluing process has the following steps.

1. Into the bottom part of the rack the detector pair is positioned, where it is pressed against $r\phi$ positioning pins (1-4) with leaf springs (5) and held down on a teflon coated vacuum chuck.

2. The rack is positioned onto the rack positioning pins (13a,b) in the lid.

3. The brass bearings are inserted over the pins 13a,b into the rack and the gap is filled with epoxy resin. The rack and the bearings are held in position by two nuts screwed onto pins 13a,b. Only the position of the bearings is precisely defined.

4. Onto the surfaces of the rack which are to be glued to the silicon, epoxy resin is dispensed.

5. The lid is slid onto the bottom part and aligned. The alignment blocks (15a,b) are pressed against pins 1, 4 with the lid leaf spring 10 and against 9b by the operator.

6. The vacuum in the lid is switched on to pull down the lid with the rack onto the silicon and fix the lids position.

7. Now the glue is cured at slightly elevated temperatures around 25^oC. Hot curing is avoided to minimize the expansion of the jig during curing.

4.3.3 Connection Bars

The connection bars are glued to the silicon pair after the racks. The pair is held via vacuum in the lid of the assembly jig. A special tool (not shown) holds the connection bars to which epoxy glue is already applied. The tool (bar holder) is positioned against pins 14a,b onto the detector pair and the glue is again cured at slightly elevated room temperature.

Figure 4.6: Assembly jig in 3 views.

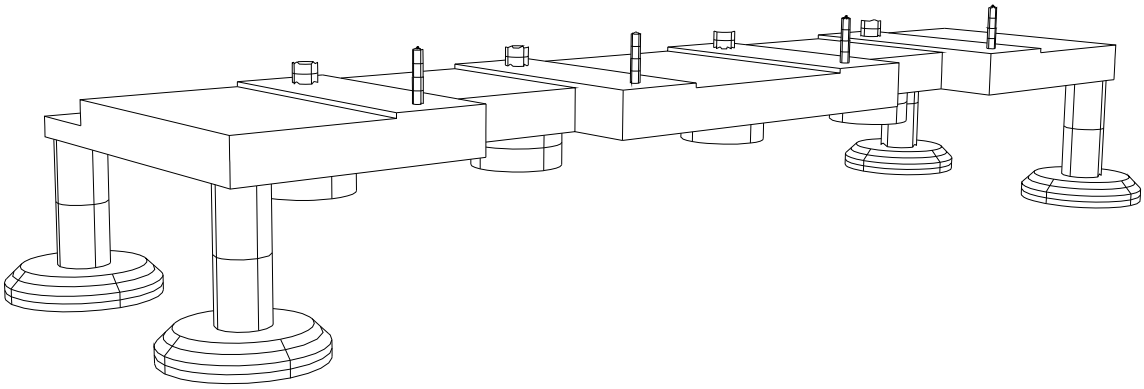


Figure 4.7: Perspective drawing of the storage-jig. The front row of small pins are the rack mounting pins. In the back the row of thicker pins are Teflon screws which can be raised until they touch the silicon detectors from below and stop them from vibrating.

4.3.4 Electronic Boards

When both pairs are completed they are mounted onto a transport and storage jig shown in figure 4.7. Into the gap between the connection bars of both pairs strips of conductive rubber are laid and the electronic board is positioned onto the two pairs, aligning it on two of the bolts extending from the connection bars. Now the connection lids are screwed onto the conductive rubber clamp ensuring that the pressure is evenly built up by tightening each of the seven nuts of each clamp through half a turn at a time. Now the module is complete and can undergo all tests before it is mounted onto the support.

4.4 Module Mounting

One of the aims of the module design was to provide the possibility to mount the module avoiding any complicated robotic process and the need for position feedback systems. The module positions are supposed to be determined by form closure[10] between the module and the positioning pins on the support. This can be done fully manually which is considerably cheaper. All connections between the module and the support structure are demountable and use screws as fixation elements. The mounting process is described by the following sequence:

1. Thermal grease is applied to the part of the silicon cooling pipe onto which the module is mounted.

2. The fully assembled module is taken from the storage jig with a special transfer tool (not shown) that maintains the relative position of the two individual racks inside the module. If the two pairs of the module were only held together by the board they would not maintain their positional stability.

3. The module is placed onto the rack alignment pins and the upper adaptors (see figure 4.4) are screwed onto them

[10] Defining the position by the matching shapes of the positioning elements on the support and on the module, i.e. a pin in a bearing.

4. The small clip on the far end of the silicon detectors is clipped into the neighbouring rows upper adapter hooks.

5. When all four modules in one row along Z, using the same cooling pipes are mounted thermal grease is applied to the front end chips, the electronics cooling pipe is rotated onto the chips and the spring clips holding the pipe are closed.

Now the module (row of modules) is fully mounted and the cooling system tests as well as connectivity and readout tests can start.

4.5 Prototypes

This subsection describes the prototypes of two silicon strip detector modules and support structure built for mechanical and thermal tests at the university of Dortmund. The emphasis lies on the modifications to the design described in section 4.2 that were necessary to do the tests described in chapter 6. Figure 4.8 shows a photograph of two modules mounted on the support structure. It can be seen that only one out of two rows of module positions has been filled.

4.5.1 Detector Pairs

Each detector is 50*60 mm^2 in area and 300 μm thick. They have been cut to size by the manufacturers to approximately 5 microns tolerance. Two detectors are concatenated to a pair with a non conductive glue. The gluing takes place as an integral part of the module assembly in the assembly jig described in section 4.3. The electrical detector-detector interconnects described in chapter 4 are not functional in these pairs. Because the detetectors were not operated under bias voltage they were heated via their front face metallisation. To do this a gold plated strip of Al-foil strips is conductively glued to each short end of all detectors. Where possible the strip has been connected to the power supplies on both ends to reduce inhomogeneities in the power dissipation caused by a voltage drop along the strip. The strips supply the detectors with a current of up to 5 A that passes along the strips in the detector metallisation. To minimize the power dissipated in the supply lines and to avoid mechanical stress transmission from the supply, highly flexible, large cross-section cables (4 mm^2) were used. Because no extra heater element had to be applied to the silicon detectors their mechanical and thermal properties were minimally disturbed and the heat was generated evenly across the complete surface of the detectors as would be the case during real detector operation.
In order to measure temperatures on the detectors a total of ten 1mm^2 Ni-100 resistors were glued to each module. The positions and the numbers of the resistors are shown in figure 4.10 as small black squares. Each sensor was connected with 4, 0.003 mm^2 enamel coated copper wires to solder pads on the electronic board.

4.5.2 Electronic Boards

The electronic boards do not carry any front end chips but are equipped with high accuracy (less than 0.1 % variation) 100Ω surface mount resistors. Three resistors in parallel represent one front end chip or read out controller. The board also has solder input pads for 4 wire interconnects to 10 temperature sensors located on the silicon detectors. The

Figure 4.8: Two modules mounted on the support structure.

Figure 4.9: Enlargement of the left module in figure 4.8 in the region of the electronic board.

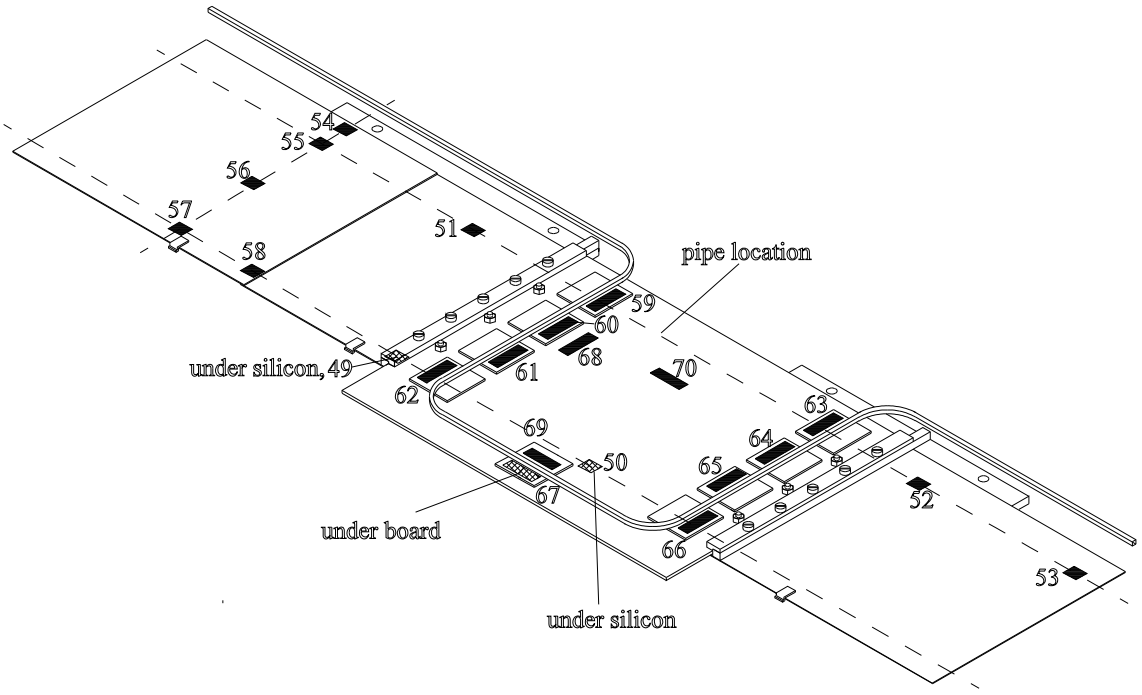


Figure 4.10: Positions and numbers of the Ni and Pt-100 Sensors on module 1. The sensor numbers shown here have an offset of +1 with respect to those in the readout as used on temperature plots. The numbers for board 2 are from 72 to 93 in the same sequence.

board itself is equipped with further 12 temperature sensors (Pt-100), the position of which is shown in figure 4.10 as black rectangles. Each board has two 50 pin double density ribbon cable connectors through which all sensors are connected to the DAQ and the resistors on the board are supplied with power. In total four boards have been made; two using standard glass fibre material and two using a kevlar paper material[11] which is supposed to have a CTE of less than 6 ppm.

4.5.3 Support Structure and Local Cooling System

The support structure and cooling system manufactured for this prototype are designed to receive two rows of two modules each. They have the full functionality of the design discussed in section 4.2. In addition a total of 16 Pt-100 resistors have been distributed over the cooling system and the support plate. The positions of these sensors are shown in figure 4.11. The cooling tube is made from a square Al-Cu-Si alloy extrusion with 0.1 mm wall thickness and 1 mm inner dimension. It has been shaped to follow the path of the electronics cooling pipe by cold bending. All pipe joints were glued with epoxy glues. The connectors on the cooling pipes are held by two clamps on each end of the support to isolate mechanical stress introduced by movements of the outer hose system. This can bee seen in figure 4.8. During operation the cooling pipe can slide in one of the clamps to allow it to expand and contract.

[11] the material was DURAMID PCE by ISOLA, Düren

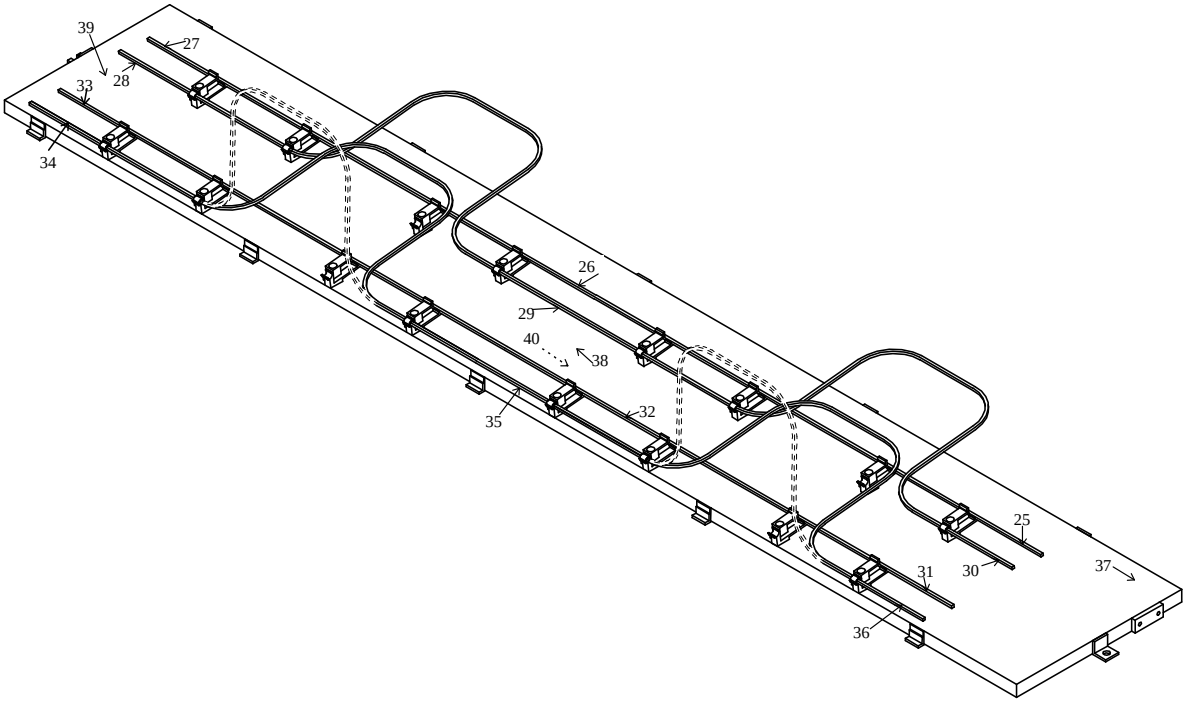


Figure 4.11: Position and numbers of Pt-100 sensors in the support and cooling system. Sensor number 40 is positioned on the reverse side of the support plate. The sensor numbers have an offset of +1 with respect to those in the readout as used on temperature plots.

4.5.4 Base Plate

In order to mount the support structure to a pillar a base plate was made. The plate was supposed to show maximum dimensional stability under temperature and humidity changes. The plate is a sandwich of two skins, each made from 6 plies of 26 g/cm^2 YLA-XN50A/RS-3 ultra high modulus carbon fibers in a 0/-60/60/60/-60/0[12] arrangement and a 1cm Nomex ECA 3.2-48 core attached with a YLA-RS3-C adhesion layer. The fibre arrangement was designed to give an isotropic 0 CTE behavior of the plate. The resin used was an YLA-RS-3C Cyanate Ester which is supposed to give smaller water absorption than standard epoxy type resins. In order to avoid exposure of the core to humidity the edges of the plate were sealed with the same fibre arrangement. Into the back of the plate threaded steel inserts were glued so that the plate could be attached to a steel finger in the environmental chamber. Four Pt-100 sensors were glued to the base plate. Their positions and numbers are shown in figure 4.12

4.5.5 Environmental Chamber

In order to minimize the influence of environmental conditions on the test setup a chamber was built to enclose it. The chamber had to be transparent in the visible to allow optical inspection and ESPI-measurements[13]. It also has to be transparent in the infra-red for wavelength between 8 and 15 μm to facilitate IR temperature measurements. After measuring IR transmission spectra of several materials, the only material found to be suitable at both wavelength for large area windows was mylar. Figure 4.13 shows a transmission

[12] the numbers given are the angles (in degrees) between the fibre and the long axis of the plate

[13] see section 5.2 in chapter 5.

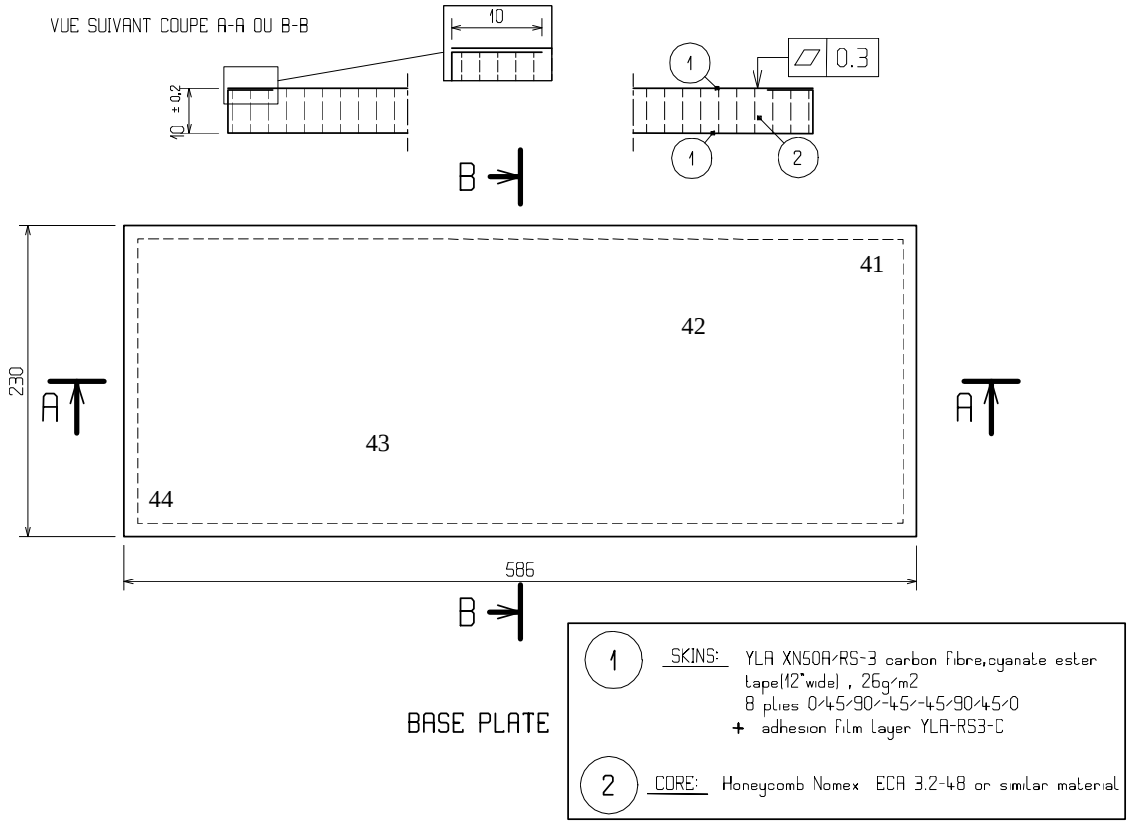


Figure 4.12: Description of the manufacturing details and the position and numbers of the sensors on the CF-base plate. The ply angle sequence in the box is incorrect. The correct sequence is 0/-60/60/60/-60/0.

spectrum of the 3 micron thick mylar film used for the window.

The chamber has to interface the internal to the external cooling system and it has to provide connections for the readout of the resistance thermometry system. Furthermore it has to have feedthroughs for the silicon detector heaters power supply lines. Additionally the chamber has to maintain the setup in a dry nitrogen atmosphere with minimal flow so that it's influence on the detectors thermal properties is small. Figure 4.14 shows the complete setup in the chamber. The white colour of the detector setup is explained in chapter 5, section 5.3. The top and bottom side of the box each show 18 feedthroughs for the silicon heater power supplies. On the left side four 25 way flat cables are connected to the box to read out the thermometry system. Also on the left are the input and output hoses of the cooling system with their insulation sleeves removed. On the right side the return loops of the cooling system extend from the chamber. On top of the box an IR-calibration Peltier element can be seen. The 2 pairs of 50 way flat cables from the two electronic boards leave the box through the lid seals. The front of the box carries the mylar window in it's frame. The mylar window is 3 μm thick and it is stretched across the frame which keeps it under tension. The frame allows for a double window which was considered necessary to run at very low temperatures to avoid condensation on the window.

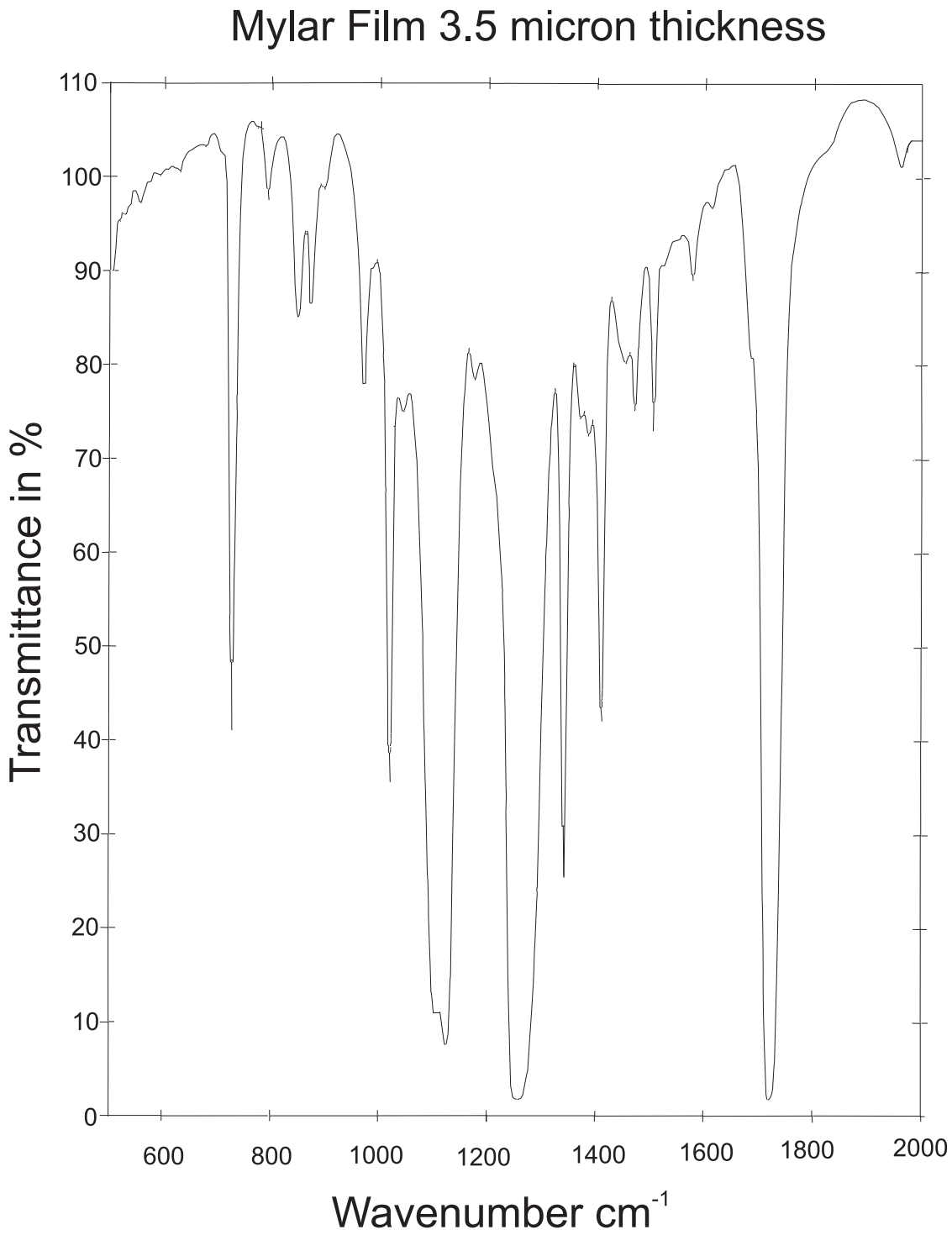


Figure 4.13: Transmission spectrum of the mylar foil used to construct the window in the environmental chamber. The relevant wavelength band of 8 to 15 μm corresponds to wavenumbers between 666 and 1250 cm^{-1}. The spectrum does occasionally exceed the 100% level due to inaccuracies in the calculation of the vertical axis scale which is done in the spectrometer. The central broad dip is a thin film interference minimum.

Figure 4.14: Two detector modules mounted in Z-overlap inside the environmental chamber. The reasons for the white colour are explained in chapter 5, section 5.3

4.5.6 Cooling system

The cooling system uses a mixture of water, ethylene glycol and Al-corrosion inhibitors to protect the thin wall Al-pipes from corrosion through ion exchange with copper and brass fittings elsewhere in the system. A schematic drawing of the cooling system is shown in figure 4.15. The central problem of the system lies in the very high pressure required to achieve turbulent flow in the thin pipes of the primary cooling circuit[14]. Calculations of the pressure drop in the pipes were based on the following formula [20]. The head or pressure lost in a pipe of hydraulic equivalent diameter d with length L at a mean flow velocity v is given by:

$$\Delta H = f \cdot \frac{L}{d} \cdot \frac{v^2 \cdot \alpha}{2 \cdot g} \tag{4.1}$$

where α is an empirical correction factor which is approximately 1 (2) for turbulent (laminar) flow and g is the gravitational acceleration. For non circular pipes d is given as the ratio of the cross-sectional area over the wetted perimeter of the pipes cross-section. The expression $\frac{v^2 \cdot \alpha}{2 \cdot g}$ is often called velocity head. f is called the friction factor and was numerically calculated via a Newton-Raphson method from the following equation.

$$\frac{1}{\sqrt{f}} = -2 \cdot \ln\left(\frac{\epsilon}{3.7 \cdot d} + \frac{2.51}{R_E \cdot \sqrt{f}}\right) \tag{4.2}$$

In this equation ϵ is the surface roughness of the pipes internal wall and R_E is the Reynolds number for the flow which is given by

$$R_E = \frac{v \cdot d}{\nu} \tag{4.3}$$

Here ν is the kinematic viscosity of the flowing medium. Assuming pure water at room temperature as the cooling liquid it was found that at a flow velocity of 4 m/s and a Reynolds number of 4065 the pressure drop along 1 m of straight pipe varies from 3.2 to 3.6 bar as the surface roughness increases from 0.1 μm to 10 μm [15]. Losses at a 90^o elbow are of the order of 0.07 bars and a sudden enlargements of the pipe diameter by a factor of 3 as in the hose to pipe transitions of the system contribute a pressure drop of 0.1 bar. The total system inside the environmental chamber was estimated to have a pressure drop of 6.7 bars at a flow velocity of 5 m/s and a Reynolds number of 5068 which is definitely turbulent. The total flow in each pipe would be 0.3 l/min.

The high pressure and low flow requirements severely limit the choice of pumps available for this application. An intensive search found only one possible pump for the primary cooling circuit. A piston metering pump was chosen. It allows control of the total flow by adjusting the mark to space ratio and can supply as little as 0.1 l/min at a maximum pressure of 13 bar. The primary circuit is cooled by two paths. The bulk of the heat is extracted from the primary circuit by cooling the pumps cylinder head with a peltier assisted heat exchanger transferring the heat into a pump head cooler circuit which dumps it via a heat exchanger to the laboratories chilled water system. This cooling circuit is not actively controlled, but the heat transition in the pump head cooler can be adjusted via the peltier power supply. The second cooling path for the primary circuit is via a smaller peltier assisted heat exchanger to the secondary circuit which dumps the received heat into the ambient air via a small chiller unit. This heat exchanger is connected to a PID

[14] inside the environmental chamber

[15] The manufacturer estimated the roughness of the pipes to be of the order of 1-2 μm

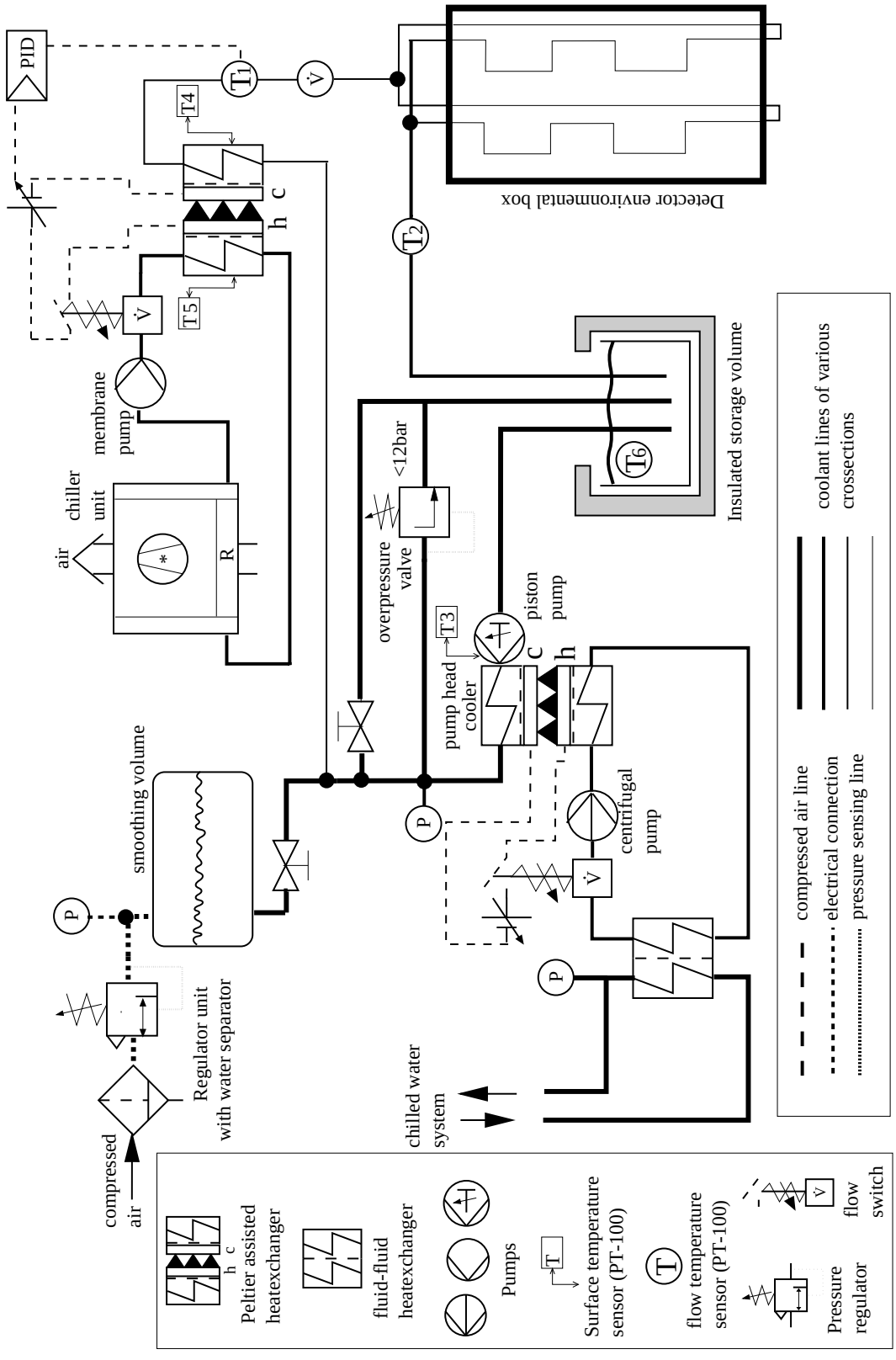


Figure 4.15: Schematic drawing of the cooling system

controller which regulates the heat transfer by pulsing the small peltier power supply. A temperature stability of 0.2 ^oC has been routinely achieved. Besides the problem of cooling the detector the extreme pressure variations and vibrations introduced into the cooling system had to be eliminated because they would have moved the detector setup and ESPI measurements would not have been possible. The main damping element was an expansion volume of approximately 3.5 liter volume directly behind the pump outlet. The air in the volume was compressed by the incoming coolant until the pressure of the gas reached the operation pressure in the primary circuit. Because a full piston volume was only 0.04 liters and at 10 bar the remaining expansion gas volume was 0.35 liters the static pressure fluctuations could be reduced by a factor 9. To further enhance the smoothing the cavity could be prepressurised with compressed air to 7 bars giving a usable smoothing volume of 2.5 liters at 10 bar and a suppression factor of 60.
Great care has been taken to ensure that failure of any coolant circuit does not result in damage to the system. This is why the power supplies for the peltier heat pumps have been connected to flow switches in their cooling circuits which will switch off all power to the peltier element if the coolant flow drops below a safe level. An overpressure valve has been included into the primary circuit to protect the piston pump and the detector cooling system from destruction through excessive pressure.

4.5.7 DAQ

A schematic drawing of the data acquisition system used to read out the thermometers inside the detector and the cooling system is shown in figure 4.16. The system is based on four-terminal resistance measurements of both platinum and nickel resistance thermometers. It multiplexes the inputs from the thermometer to a Keithley DMM 196 for sequential measurements via a set of relay scanner cards developed in the university of Dortmund. The system is controlled by a PC running Lab View via a GPIB bus. The maximum scanning frequency is 2 Hz. Data is stored as resistance and optionally as temperature values. Online monitoring of all temperatures via a graphical interface is possible. The input from the thermometers is standardized and comes via 5 pin DIN-connectors for individual sensors in the cooling system and via sets of four 25 pole flat cables for all other sensors. Inputs from the sensors on the detector modules themselves come via double density flat cables and are translated in an adaptor box which also fans the power supply lines for the electronic boards into the double density cable. The system has been tested on a thermally stabilized reference resistor and an rms spread of 0.7 mΩ in the resistance values was found. Because the measuring current of the Keithley 196 of approximately 1.4 mA introduces a power of typically 0.2 mW into each sensor, the sensor temperature would change slighly when the scanning system measures it. To keep all sensors under constant conditions, a fixed voltage of 145 mV was applied across each sensor whenever the sensor was not measured. This was achieved by connecting all sensors in parallel to a constant voltage source through the reset connections of the relay cards.

4.5.8 Design Improvements

This subsection deals with detailed improvements of the specific design discussed in section 4.2 from the point of view of assembly and mounting. For more general conclusions to be drawn in particular from the tests of the module see section 6.5.1 in chapter 6. The assembly scheme suggested in this chapter can be improved in several ways.

- The process of gluing the connection bars onto the silicon detectors requires a more easy to operate bar holding jig.
- The alignment of the electronic boards with two of the bolts from the connection bars requires metallized holes which have been drilled and reamed after metallisation.
- The precision with which the bolts were glued into the connection bars was insufficient for mass production. This could be improved by using a drilling and insertion jig.

In order to establish a similar detector module in a serial production several of the small size components such as the lower and upper adaptors, the pipe holding springs and the silicon clips have to be manufactured by injection moulding rather than precision milling as has been the case for the prototype. Furthermore better tools for the handling of the small size nuts on the connection clamps and the small leaf springs attaching the cooling pipes to the front end chips have to be built. It would also be in the interest of improved production yield and jig efficiency to introduce several fixation glue spots using UV-curing epoxy resins on the long term curing joints.
The cooling system can be improved by increasing the flexibility of the silicon pipe holder arms. It would also be necessary to establish means of bending longer sections of pipe and to miniaturize the pipes external connections.

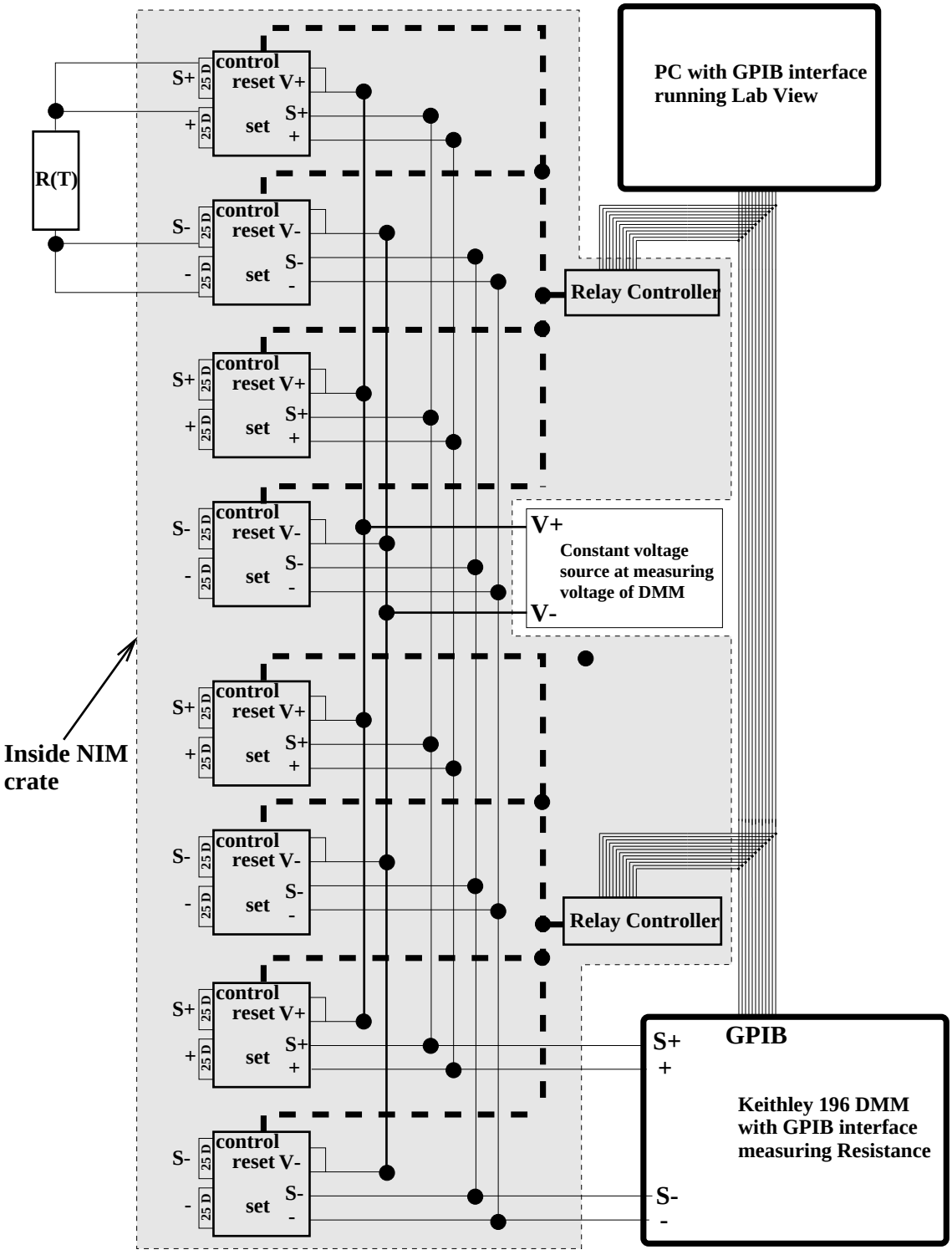


Figure 4.16: Data acquisition system for the thermometry of the detectors and the cooling system.

Chapter 5

Measurement Techniques

5.1 Introduction

This chapter will explain in detail the measurement techniques used to evaluate the prototype described in chapter 4 section 4.5. The experimental setup for thermal and mechanical measurements are described along with the fundamental principles of the techniques and the analysis methods used to evaluate the measurements. The resistance thermometry system has already been described in several places in chapter 4, focussing on the read out system in subsection 4.5.7. Results from measurements using the techniques are given in chapter 6. D

5.2 ESPI

a This section will first motivate why ESPI[1] was used as the method for deformation measurements. The fundamental principles of ESPI will be explained using the example of deformation measurements and the experimental setups used will be described. Finally the procedure of making an ESPI measurement starting with the adjustment of the system will be given.

5.2.1 Why ESPI

The problem of measuring the position stability of a silicon detector module during operation requires a technique that, in an ideal case, does not significantly disturb the modules mechanical and thermal properties. Due to the complexity of the expected movement and the intention to test and tune a finite element model of the movement and thermal properties it would be advantageous to not only measure displacements in single points but over the whole surface of the module. ESPI allows these measurements with an accuracy of $\mathcal{O}(1\mu\mathrm{m})$ over an area up to $\mathcal{O}(1m^2)$. It also measures distortion rather than shape which is advantageous for the comparison with the finite element model. It can be complemented with infra red thermographic measurements and thus supply a complete set of input parameters for the finite element analysis.

5.2.2 Fundamentals

ESPI is based on the manipulation of digitally stored images of optically rough surfaces which are illuminated by coherent light, usually from a laser. The term optically rough

[1] **E**lectronic **S**peckel **P**attern **I**nterferometry

refers to the illuminated surface varying in height by $\mathcal{O}$(wavelength) inside an area smaller than the resolution limit of the imaging system.

Subjective Speckle Pattern

When such a surface is coherently illuminated and imaged, the intensity in the image plane varies randomly. This granular appearance is called subjective speckle. It should be distinguished from objective speckle which refers to the granularity in the intensity distribution of the scattered light and not of that in the image. For the ESPI measurements the subjective speckle pattern is the relevant one. The granular appearance of the image can be understood as follows:
In the simplest case of an imaging system as shown figure 5.1 the image of a single point is a diffraction pattern, the light intensity distribution of which has the form[22]:

$$I(\theta) = I_0 \cdot \left(\frac{2 \cdot J_1 \left(\pi \cdot a \cdot \sin(\frac{\theta}{\lambda}) \right)}{\pi \cdot a \cdot \sin(\frac{\theta}{\lambda})} \right)^2 \tag{5.1}$$

$I(\theta)$ is called the point spread function of the optical system, J_1 is the first order Bessel function, λ the wavelength of the illuminating light, a the diameter of the aperture and θ the viewing angle as shown in fig 5.1. In the following f will denote the focal length of the lens, u the distance between the lens and the object and ν the distance between the lens and the image. Note that it will be assumed that the image of the speckle pattern is in focus and therefore ν =f. Because of the roughness of the surface the light coming from each point, i.e. P_1 in figure 5.1 has a random phase. P_2 is a point separated from P_1 by a distance that will make Q_1 be in the first minimum of its diffraction pattern from P_2 and vice versa. The corresponding distance $\overline{Q_1Q_2}$ in the image plane can be found from equation 5.1 to be $\frac{1.22 \cdot \lambda \cdot \nu}{a}$ which translates to $\overline{P_1P_2} = \frac{1.22 \cdot \lambda \cdot u}{a}$. If we neglect the light amplitude beyond the first minimum we can say that light from any point P, separated from P_1 by less than $\overline{P_1P_2}$ will interfere with the light from P_1 and the light from P will have a random phase with respect to that of P_1. The resultant intensity will be random and the image of an element of surface with diameter smaller than $2 \cdot \overline{P_1P_2}$ can be said to have one resultant amplitude and phase. This image is called a subjective speckle. The speckle diameter is proportional to the numerical aperture[2] $F = \frac{f}{a} = \frac{\nu}{a}$ and the wavelength λ.
The light complex amplitude distribution of a subjective speckle pattern defined to be in the x-y plane is usually written as:

$$U = u(x,y) \cdot e^{i \cdot \psi(x,y)} \tag{5.2}$$

where $u(x,y)$ and $\psi(x,y)$ are the randomly varying amplitude and phase of the speckle.

Correlation Interferometry

In the following the intensity distributions of electronically subtracted interferograms obtained with the two interferometers described in subsection 5.2.4 will be explained. It is assumed that the object and image are in the x-y plane. A lot of details about the formation and use of speckle pattern interferometry can be found for example in [22]. An image obtained with the out of plane sensitive interferometer of figure 5.3 is an interference between a speckle pattern of intensity distribution I_1^{spek} produced by the object and

[2] if the pattern is in focus

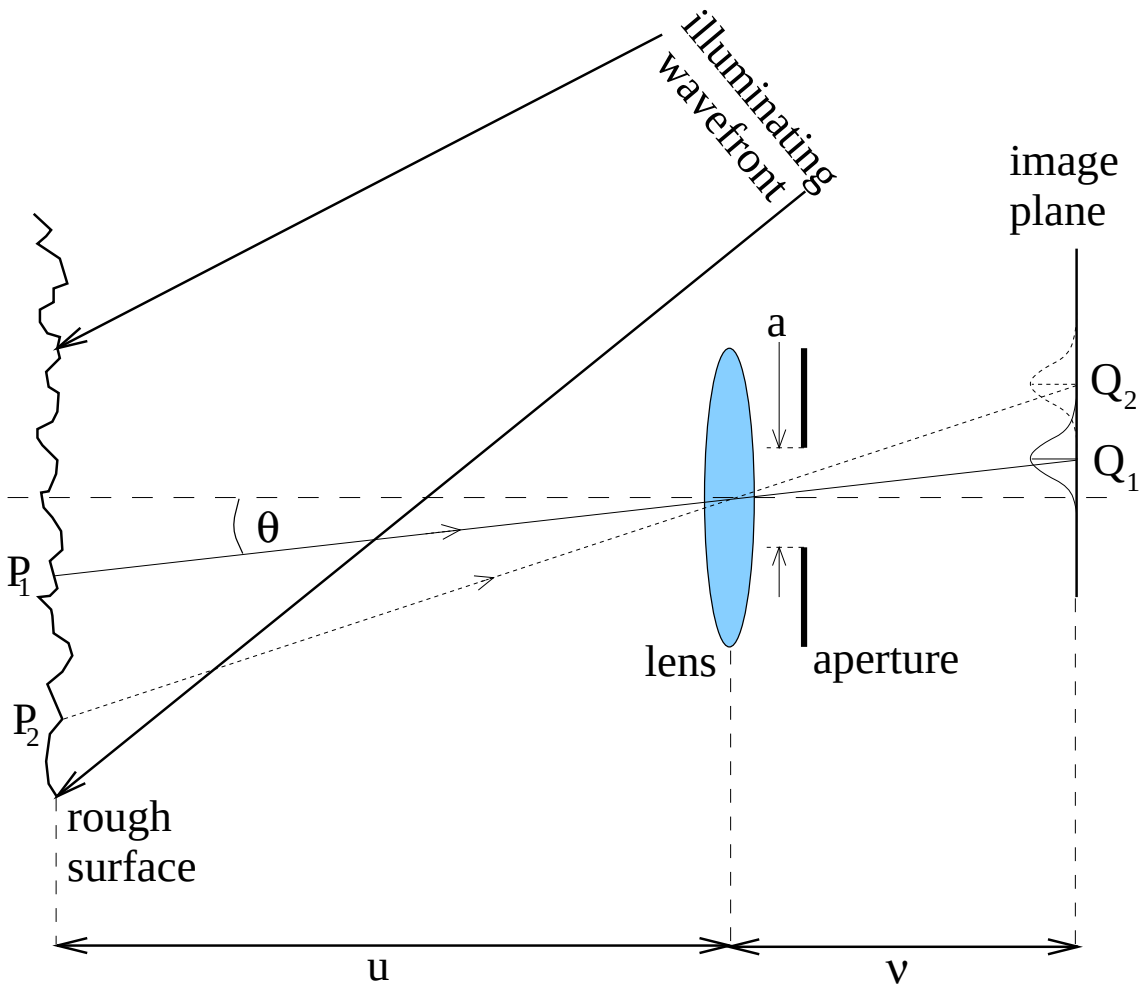


Figure 5.1: Formation of a subjective speckle in the image plane. The two distributions in the image plane are the point spread functions of P_1 and P_2 up to their first minimum. If the object is in focus ν=f, where f is the focal length of the lens.

the plane wave of the reference beam with intensity distribution I_1^{ref}. The amplitude distribution $U_1^{img}(x,y)$ of the interference is similar to that of the speckle pattern and the resulting intensity distribution which is recorded by the video camera can be written as:

$$\mathcal{I}_1 = I_1^{spek} + I_1^{ref} + 2 \cdot \sqrt{I_1^{spek} \cdot I_1^{ref}} \cdot \cos \psi_1^{spek-ref} \tag{5.3}$$

In the above equation all variables are functions of the coordinates in the image plane, (x,y) and the dependence has been suppressed in the formula. A second image, obtained after the object has undergone a displacement $d_Z(x,y)$ out of the object plane has the intensity distribution $I_2(x,y)$. The displacement can be expressed as a change of relative phase $\Delta\phi(x,y)$ between the light from the object and the reference beam. Simple geometry shows that:

$$\Delta\phi(x,y) = 2 \cdot \pi (\cos \theta_{illum} + 1) \cdot \frac{d_Z(x,y)}{\lambda} \tag{5.4}$$

The intensity distribution of neither the speckle pattern nor the reference beam should change with $d_Z(x,y)$ and therefore $I_2^{spek} = I_1^{spek} = I^{spek}$ and $I_2^{ref} = I_1^{ref} = Iref$. Then the second image has the intensity distribution (again dependence on (x,y) suppressed):

$$\mathcal{I}_2 = I_1^{spek} + I_1^{ref} + 2 \cdot \sqrt{I_1^{spek} \cdot I_1^{ref}} \cdot \cos (\psi_1^{spek-ref} + \Delta\phi) \tag{5.5}$$

Assuming that $\mathcal{I}_1$, $\mathcal{I}_2$ and $d_Z(x,y)$ are statistically independent and $\langle\mathcal{I}_2\rangle = r \cdot \langle\mathcal{I}_1\rangle$ it[3] is straight forward to calculate the correlation between the two images $\rho(\mathcal{I}_1,\mathcal{I}_2)$ defined as:

$$\rho(\mathcal{I}_1,\mathcal{I}_2) = \frac{\langle\mathcal{I}_1 \cdot \mathcal{I}_2\rangle - \langle\mathcal{I}_1\rangle \cdot \langle\mathcal{I}_2\rangle}{\sqrt{(\langle\mathcal{I}_1^2\rangle - \langle\mathcal{I}_1\rangle^2) \cdot (\langle\mathcal{I}_2^2\rangle - \langle\mathcal{I}_2\rangle^2)}} \tag{5.6}$$

one finds that

$$\rho(\mathcal{I}_1,\mathcal{I}_2) = \frac{1 + r^2 + 2 \cdot r \cdot \cos(\Delta\phi)}{(1+r)^2} \tag{5.7}$$

which simplifies to $\rho(\mathcal{I}_1,\mathcal{I}_2) = 0.5 \cdot (1 + \cos(\Delta\phi))$ in the case of $r = 1$. The correlation has a minimal value of $[(1-r)/(1+r)]^2$ for $\Delta\phi = (2 \cdot n + 1) \cdot \pi$ and always reaches a maximum value of 1 when $\Delta\phi = 2 \cdot n \cdot \pi$. If the intensity of any beam changes during the displacement the correlation will not reach 0. From equations 5.3 and 5.5 it can be seen that for the case of maximum correlation $\mathcal{I}_1$ and $\mathcal{I}_2$ are equal and thus a subtraction of the images will result in a zero intensity region for displacements of $d_Z = n \cdot \lambda/(\cos\theta_{illum} + 1)$ which for perpendicular illumination becomes $d_Z = (n/2) \cdot \lambda$. In general (assuming r=1) the process of taking the absolute value of two subtracted images will result in an intensity distribution of the form:

$$\begin{aligned} \mathcal{I}_{diff} &= \sqrt{(\mathcal{I}_1 - \mathcal{I}_2)^2} \\ &= 4 \cdot [I^{spek} \cdot I^{ref} \cdot \sin^2(\psi^{spec-ref} + \frac{\Delta = \phi}{2}) \sin^2(\Delta\phi)]^{1/2} \end{aligned} \tag{5.8}$$

In the above equation $\psi(x,y)$ is a rapidly varying random phase, whereas $\phi(x,y)$ depends linearly on the displacement to be measured. If equation 5.8 is averaged along lines of constant ϕ, $\mathcal{I}_{diff}$ will show minima and maxima depending on the value of ϕ. The first term $2\sqrt{\cdot I^{spek} \cdot I^{ref} \cdot \sin^2(\psi + \frac{\Delta\phi}{2})}$ which varies with ψ will determine the amplitude of the intensity variations. The maxima also correspond to the locations of minimal correlation and show values of $2 \cdot \sqrt{I_1 \cdot I_2}$ which is not homogeneously white but is a speckle pattern with an average intensity of the geometric mean of the two images. The displacement necessary to induce one complete fringe is $\lambda/(\cos\theta_{illum} + 1)$.
In the case of the in plane sensitive interferometer shown in figure 5.4 the relation between the in-plane movement and the phase difference is slightly different and reads:

$$\Delta\phi = \frac{4 \cdot \pi}{\lambda} \cdot d_x \cdot \sin\theta_{1,2} \tag{5.9}$$

This corresponds to a movement of the object between two fringes of $\lambda/(2\sin\theta_{1,2})$.

5.2.3 Fringe Pattern Analysis

The process of ESPI fringe pattern analysis is divided into three main tasks. First to improve the quality of the very noisy fringe patterns, second to extract the change of phase that has caused the fringe pattern and third to unwrap the phase change into a physical distortion of the object.

Noise Reduction

The two methods used here to improve the fringe contrast and reduce the noise were initially a simple contrast enhancement and a Fourier filtering method. To enhance the

[3] there could be a variation of average intensity between the two images but usually $r = 1$

fringe contrast, all intensities above a given threshold were set to a maximum value and intensities below a second threshold were set to 0. It was found that for most illumination conditions the thresholds could be calculated from the image as:

$$I_{thresh}^{max} = \langle I_{image} \rangle + K_{max}\sigma(I_{image}) \tag{5.10}$$

$$I_{thresh}^{max} = \langle I_{image} \rangle - K_{min}\sigma(I_{image}) \tag{5.11}$$

with good results for $0.5 < K_{max} < 1.2$ and $-0.3 < K_{min} < 0.8$ where $\sigma(I)$ is the root of the variance of the intensity distribution.
The Fourier filtering method was used for images with few and rather simple fringes. With SEMPER[4] a fast Fourier transformation of the image was made and the power spectrum of the image was displayed. The fringe pattern signal was easily visible in the area around the origin (at low spatial frequencies). Then a graphical cut was drawn around the signal. The Fourier spectrum outside the marked area was set to its average value. The edge of the cut was usually softened over 5 to 10 pixels. If the cut is made symmetrically around the origin of the Fourier plane[5] the back transformation directly results in images with most of the high frequency speckle noise removed, (low pass filtering) but the image will appear blurred and fringe edges will be very soft. If the cut is drawn asymmetrically as in figure 5.2 the back transformation will be a complex image. Taking the inverse tangent of the quotient consisting of the imaginary over the real part of the image will represent the phase of the image. Figure 5.2 gives an impression about the methods performance and limits. It shows an unfiltered fringe pattern, a contrast enhanced fringe pattern, the power spectrum (logarithmic) of the Fourier transform of the fringe pattern with the signal area marked and the resulting back transformation (a phase distribution modulo 2π) after the asymmetric filter and arctangent calculations were applied. It should be noted, that for fringe pattern which are dominated by a set of parallel fringes this method will directly yield a correct phase map. This can be achieved by superimposing a known tilt onto the objects distortion, usually by moving it with a piezo stack. These superimposed fringes are usually referred to as carrier fringes because the distortion of interest will be visible as a modulation of these carrier fringes. In figure 5.2 the method shows its limitations in the places where the fringes change direction. Here the ambiguity of the sign of the phase can not be solved uniquely for all fringes and the image represents no longer a modulated carrier pattern. For fringes in the lower half of the image, the process causes the phase (brightness) in the left part of a fringe to increase in the opposite direction as in the right part of the same fringe and the change from one direction to another leads to a confused image in the region where the fringes change direction. For a review of asymmetric Fourier filters and their application to ESPI see [23].

Phase Stepping

The fringe patterns described above only give information about the absolute value of the phase shift that induced them, but not about the sign. This results in an ambiguity in the distortion which will also have an undetermined sign. In cases were the distortion changes sign, i.e. if in an out of plane distortion measurement a flat object distorts into a corrugation like shape, this can lead to serious misinterpretation of the fringe pattern. In order to resolve these ambiguities a so called phase stepping method was implemented. The method used here has first been described by Carré [24] and is discussed with other

[4]SEMPER is the image analysis software used throughout this work. SEMPER is a product of Synoptics Limited

[5]The Fourier transformation of a real image is always symmetric around the origin of the Fourier plane

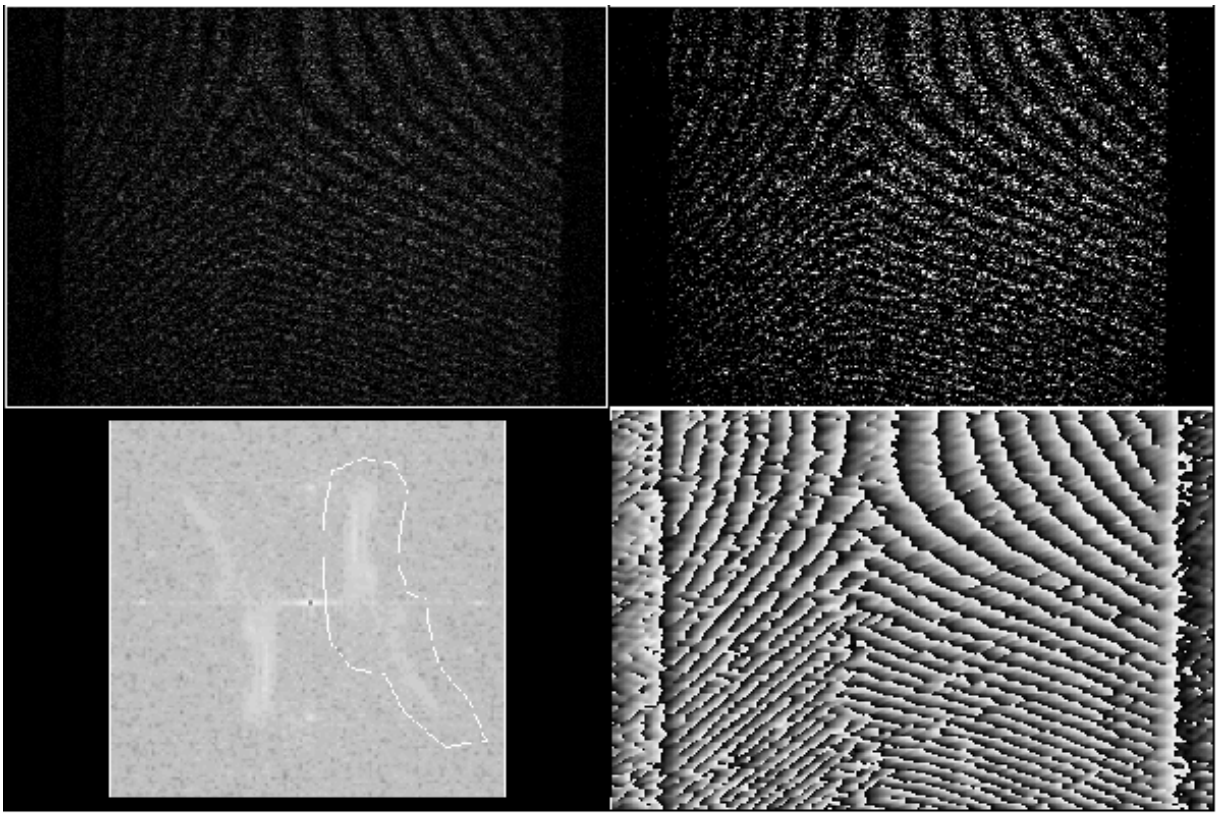


Figure 5.2: A raw fringe pattern (top left) after contrast enhancement (top right) and asymmetric Fourier filtering and phase calculation (bottom right). The central part of the logarithmic power spectrum of the fringe pattern is shown, together with the filter acceptance area in the bottom left part. The origin of the power spectrum is indicated by the black dot in the center[7]. The horizontal and vertical axis correspond to the real and imaginary frequency components.

algorithms for phase stepping in [25]. It is assumed that 4 images can be taken of the object without any change in object geometry. Between the 4 images the phase of one of the beams (the illumination beam common to the out of plane and in-plane setup) is shifted via a piezo driven mirror[8] through an unknown phase angle α. The four intensity distributions are then given by:

$$I_1(x,y) = I_0(x,y)\{1+\gamma\cos[\psi(x,y)-\frac{3}{2}\alpha]\} \quad (5.12)$$

$$I_2(x,y) = I_0(x,y)\{1+\gamma\cos[\psi(x,y)-\frac{1}{2}\alpha]\} \quad (5.13)$$

$$I_3(x,y) = I_0(x,y)\{1+\gamma\cos[\psi(x,y)+\frac{1}{2}\alpha]\} \quad (5.14)$$

$$I_4(x,y) = I_0(x,y)\{1+\gamma\cos[\psi(x,y)+\frac{3}{2}\alpha]\} \quad (5.15)$$

The advantage of the Carré method is, that it does not require a calibrated phase step, even though it will work best if α is close to 90 degrees. In the above equations γ is called the intensity modulation distribution. It can be calculated assuming that $\alpha \approx \pi/2$.

$$\gamma(x,y) = \frac{1}{2I_0}\sqrt{\frac{[(I_2-I_3)+(I_1-I_4)]^2+[(I_2+I_3)-(I_1+I_4)]^2}{2}} \quad (5.16)$$

[8] see subsection 5.2.4 for details of implementation

The error on γ induced by α deviating from $\pi/2$ is proportional to the difference between $\sin(\alpha-\pi/2)$ and $\alpha-\pi/2$. In the above equations I_0 is the average intensity of all 4 images. With equation 5.16 the tangent of the phase can be calculated by solving the system of 4 equations given above:

$$\tan\psi = \frac{\sqrt{[3(I_2-I_3)-(I_1-I_4)]\cdot[(I_2-I_3)+(I_1-I_4)]}}{(I_2+I_3)-(I_1+I_4)} \tag{5.17}$$

The arc tangent of this equation will yield the phase modulo π which has to be expanded into it's values modulo 2π. This can be done by examining the sign of quantities proportional to $\sin(\psi)$ and $\cos(\psi)$. One set of such quantities for the Carré method is:

$$A = (I_2-I_3) = [2\cdot I_0\cdot\gamma\sin\alpha]\sin\psi \tag{5.18}$$

$$B = (I_2+I_3)-(I_1+I_4) = [2\cdot I_0\cdot\gamma\cos\alpha\sin^2\alpha]\cos\psi \tag{5.19}$$

Given the sign of A and B the phase modulo 2π can be calculated from table 5.1.

A	B	ψ_{new}	Range of ψ_{new}
pos.	pos.	ψ	0 to $\frac{\pi}{2}$
pos.	neg.	$\pi-\psi$	$\frac{\pi}{2}$ to π
neg.	neg.	$\pi+\psi$	π to $\frac{3\pi}{2}$
neg.	pos.	$2\pi-\psi$	$\frac{3\pi}{2}$ to 2π
0.	any	π	π
pos.	0	$\pi/2$	$\pi/2$
neg.	0	$3\pi/2$	$3\pi/2$

Table 5.1: Expansion of the phase modulo π into its values modulo 2π

Phase Unwrapping

In order to obtain a field of distortion of the object the phase which is still modulo 2π has to be unwrapped. There are many algorithms suggested for this problem but none of them is applicable to the fully automated unwrapping of interference fringe images of a complete detector module. This is due to the very complex surface structure which has many discontinuities and complicated shapes. The effort involved in the computer based unwrapping was considered too big in view of the fact that in the majority of cases quantitative data with maximum accuracy is not needed. Most interference fringe images are used to give a qualitative impression of the detectors behavior.

5.2.4 Experimental Setup

Two ESPI setups were constructed, serving different purposes. The principle innovation in both setups is the complete decoupling of the laser from the measuring setup with optical fibers.

He-Ne-Laser setup

The first and smaller arrangement was based on a 5 mW He-Ne Laser operating at 633 nm. The setup was only capable of measuring distortions out of the plane of the object in an area of approximately 200 cm^2. No phase stepping was implemented and the primary aim of the setup was to demonstrate the feasibility of the technique and gain experience

necessary for the installation of a larger system. The laser was mounted on a small optical table, separate from the measurement table. Approximately 1% of the beam was split off and sent through a variable attenuator behind which it was focussed into a polarisation preserving mono-mode reference fibre. The remaining beam was focussed into a similar fibre used to illuminate the object. Both fibers had widening optics directly attached to the output end. The camera, illumination fibre and object were placed on a second optical table and surrounded by a light tight box with black internal walls. The surface of the optical table was also blackened as far as possible to minimize the influence of stray light. The reference fibre output was attached to the video camera and the beam was coupled into the camera with a beam splitter behind the camera lens. In this way the reference beam was not imaged but evenly illuminated the CCD of the camera. The object fibre output illuminated the object with an angle of 5 to 15 degrees between the surface normal of the object and the illumination direction. 0 degree illumination was not possible because of specular components of the reflection. This setup was used to study the behavior of a single crystal of the module with the connection bars glued to its back.

Ar^+-Ion-Laser setup

The arrangement for the two interferometers realised in this setup can be seen in figures 5.4 and 5.3. The second setup involved a much more powerful Ar^+-ion laser operated at 514 nm. The laser was a Coherent Innova 90-4 [21] with 4 W output power in multi-mode operation and 870 mW in single mode operation at 514 nm. Single mode operation was achieved by introducing an etalon and a grating in front of the mirrors. The laser input power of 15 kW was cooled by a water cooling system connected to the in house chilled water supply. The stability of the laser was much better than that of the He-Ne laser used before and the minimal frequency drift velocity was 19 MHz/h. The laser was again positioned on a separate optical table and the length of all fibres transporting the light to the measurement table was 25 m.

Fibres

Long fibres were chosen to be able to perform measurements with the complete measurement table in an environmental chamber and the laser being operated outside the chamber. The long fibre length introduces a problem for phase stepping operation as the length and refractive index of the fibre changes with temperature by $5.5 \cdot 10^{-7}$/K and $9.9 \cdot 10^{-6}$/K respectively. If the two fibres of the interferometer are exposed to different temperatures the relative phase between them will drift. It can be seen that the phase will drift by 2π if the temperature of one fibre, integrated over it's total length changes with respect to the other fibre by 1 mK. This requires a very stable temperature environment if phase stepping methods are to be performed. A non phase stepped image is only affected by showing an overall phase shift of the fringe pattern which can at maximum introduce an error of one fringe when counting fringes over several images. The fibres were mounted on 9×9 cm^2 cross section Al extrusion pillars on both long edges of the table. The height and the two illumination angles could be varied continuously. The position along the length of the table can be varied in steps of 15 cm. For phase stepped measurements and in cases of vibration sensitive setups the fibres were mounted on a combination of two cross connected pillars[9] stabilised in the direction perpendicular to the tables long axis by cross bracings.

[9] the pillars were separated by 15 cm along the viewing direction

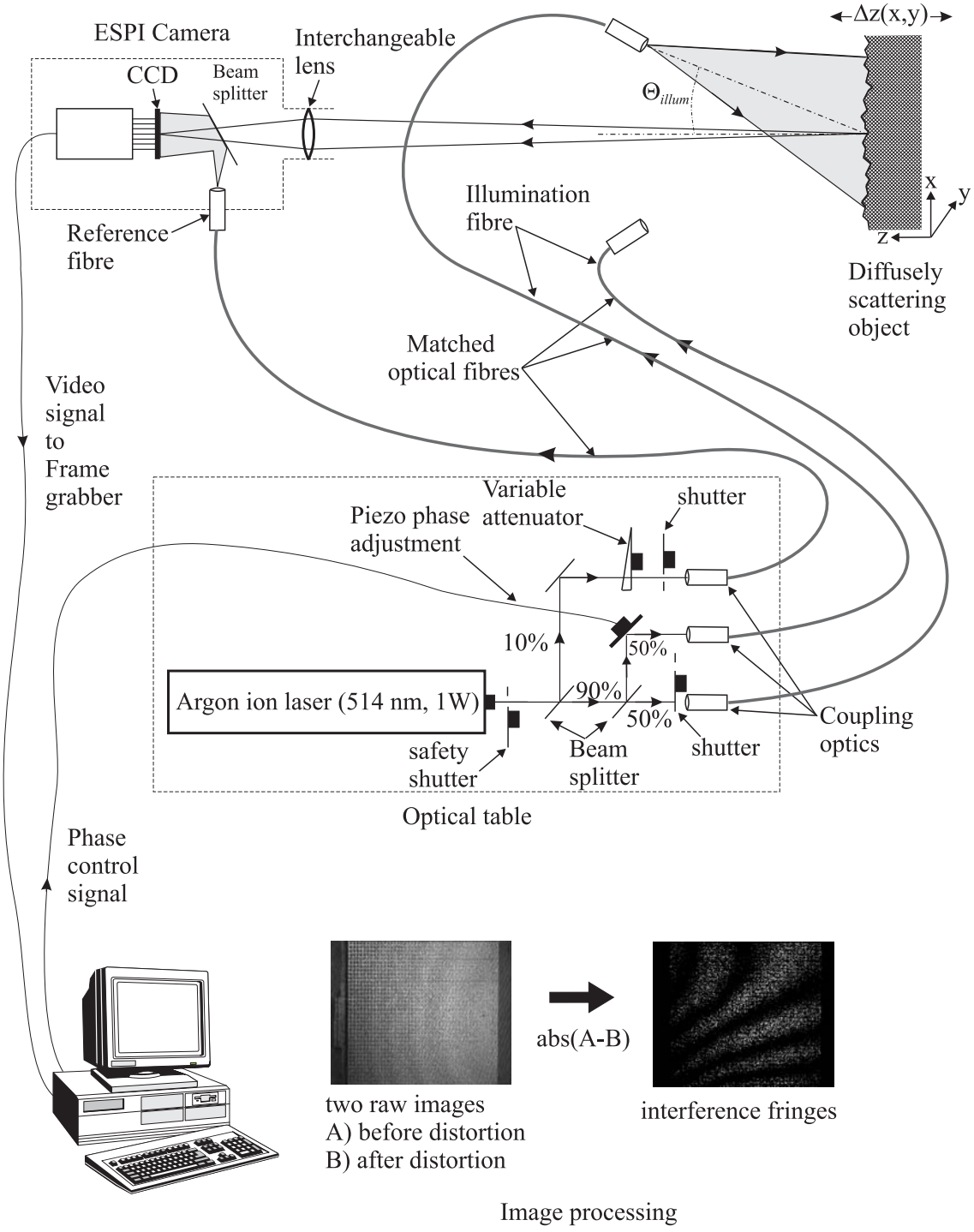


Figure 5.3: Setup of the out of plane sensitive interferometer

Tables

The laser table was a standard optical sandwich table of 2x1 m size. The measurement table was constructed at Oxford NAPL. Its outer dimensions are 1.5x5 m^2. It was an aluminium sandwich structure with 2 cm thick skins and 3 I-beams of 40 cm height running along the long dimension of the table. The table was positioned on 4 air springs with a resonance frequency of 6 Hz. The laser table was positioned on top of the measurement table at the short end, opposite the observed object.

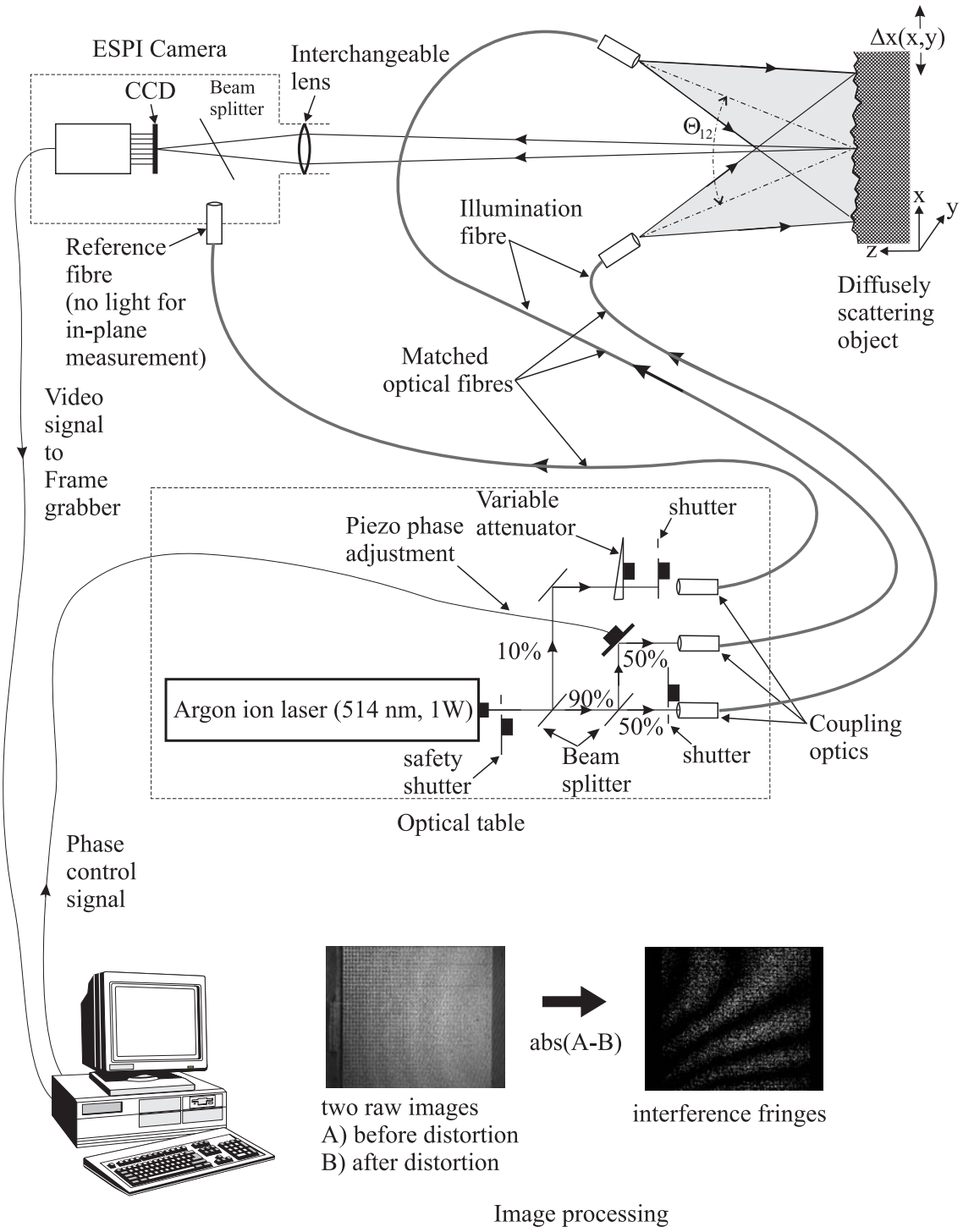


Figure 5.4: Setup of the in-plane sensitive interferometer

Safety

The laser was enclosed on the laser table by a black aluminium box with a semitransparent perspex lid which was interlocked to the laser shutter. The table was surrounded by a ring of black curtains which were also interlocked with the laser shutter. A further curtain box enclosed the volume on top of the measurement table excluding the laser table. The laboratory door was also interlocked to the laser shutter. The laser power supply was interlocked to a flow switch in the laser cooling system as well as to a temperature sensor in the laser cooling water return flow. Laser operation was only possible for a limited

number of trained persons registered as users of this laser. Laser start up required two keys for the Laser remote control and the laser shutter controls. A fixed attenuator could be remotely set into the main laser beam from the shutter control box. When working inside the outer curtain area with the laser in operation safety goggles had to be worn. Both output fibres were enclosed in a metal tube. This way it was impossible to approach the fibre with the eye closer than the minimum safety distance of 30 cm. At this distance the light from the fibres, which had widening optics with angles between 10 and 20 degrees was considered safe.

Camera

ESPI images were recorded with a Pulnix TM-460 surveillance camera. The camera had a CCD with 573 lines times 417 pixels with a physical pixel size of 8.3 μm vertically and 15.2 μm horizontally. The analog signal from the camera was standard CCIR and the camera was operated in non interlaced mode[10]. Each field of a frame had 287 lines, each line being the sum of two adjacent lines. This so called square sampling results in an effective pixel size of 16.6 by 15.2 μm . The lenses on the camera ranged from focal length of 28mm to 75mm and numerical apertures between 2.8 and 16 were available. The camera was built into a new frame to incorporate a beam splitter behind the lens. Unfortunately in out of plane operation only half the pixels of each line were usable due to an interference effect between the incident reference beam and the CCD face plate. The camera was mounted on a 15×15 cm^2 cross-section aluminium extrusion pillar and could be positioned continuously across the width of the table. The camera height was also continuously adjustable. The position along the length of the table could be varied in steps of 15 cm. The camera viewing axis was always parallel to the long axis of the table.

DAQ

Images from the camera were either displayed on a control monitor or stored in a digital framestore inside a PC running SEMPER 6. The framestore sampled the cameras analog signal with an 8 bit ADC into frames of 512 lines and 728 samples per line. This PC also had control over the primary piston pump of the cooling system and the phase stepper. It was possible to gate the pump for phase stepping applications so that the pump would not operate and cause vibrations as long as images were acquired. The PC was also able to trigger the IR scanner. A VISUAL BASIC interface had been written for SEMPER 6 which allowed the acquisition of sets of phase stepped images. The program had control over the piezo power supply and the pump via a DAC card in the PC. The card controled the output voltage of the piezo power supply and switched a solid state relay controlling the mains power of the piston pump.

Phase Stepper

Into one of the illumination beams a piezo driven mirror was positioned. The power supply of the piezo was under control of the ESPI data acquisition and could be driven in 256 steps sizes up to a maximum shift of 2.5 μm . The illumination fibre was phase stepped in both in plane and out of plane configuration. In out of plane configuration usually the low power reference beam is phase stepped, but this would have meant that the piezo had to be moved when switching from out of plane to in-plane operation.

[10] two identical fields per frame

5.2.5 Procedure

This subsection will describe how sets of phase stepped out of plane images are acquired and how the setup is adjusted beforehand. After approximately 30-60 min the laser has stabilised and the alignment of the fibre inputs can be established at very low laser powers to not damage the fibres with a badly aligned beam. Since the fibres only preserve the polarisation of the light when it is input in the correct direction the fibre rotation angle also has to be tuned. The optimal fibre rotation is found if the output polarisation does not change when the free parts of the fibre are moved. When the above two steps have been performed for both fibres the output polarisation of the fibres has to be aligned. After this the intensities of the fibres are tuned to be equal to give best fringe contrast. This is done either by using a variable attenuator, or in cases of very small differences by detuning the input of the brighter fibre taking care to not damage it by detuning it too far. The distance of the fibres from the object is set such, that a good compromise between the absolute brightness and the homogeneity of the illumination[11] is achieved. The angle with respect to the viewing axis has to be chosen such that the sensitivity of the spectrometer is sufficient and that specular reflections do not enter the image. The camera is adjusted to give the biggest possible numerical aperture[12] to maximise speckle size whilst keeping the intensity of the object centered on the sensitive range of the camera and framestore. The illumination and imaging geometry is noted in the logbook and the system is checked for stability by taking a series of images without changing any object parameters. Calculating the difference of these images with respect to the first image and determining the average brightness gives a good indication of the stability and noise. If the average brightness is stable it represents the amount of noise in the complete system. If it is unstable the object is moving or the beam intensities are varying.

5.3 IR-Imaging

5.3.1 Why IR-Imaging

The method of measuring the temperature of an object by sampling its infrared radiation field differs from conventional thermometry in three major respects:

- The radiation field carries information of the complete temperature field on the objects surface whereas a thermometer only measures the average temperature of a finite subsection of the surface.
- The infrared measurement disturbs the test object neither mechanically nor thermally. A thermometer always adds mass to the object and therefore disturbs the objects temperature, heat capacity and mechanical properties.
- However, the infrared radiation field depends neither uniquely nor simply on properties of the test objects temperature distribution. It is influenced by many parameters such as ambient temperature, air temperature and humidity, object emissivity and scanner sensitivity to name only a few.

Since the temperature measurements in this case had to be done simultaneously with high resolution mechanical measurements the influence on the mechanical properties had to be

[11] the fibres output light with a Gaussian intensity distribution

[12] smallest opening

minimised.
The complexity of the detector modules and the need to obtain complete temperature fields as input variables for finite element simulations left no other choice but infrared scanning.
It was however necessary to combine the infrared measurements with a reduced number resistance thermometers to be able to calibrate the temperature fields. There is a usefull complementarity between the two methods of temperature measurement. Where IR-images can easily be used to visualise the geometrical connectedness of the temperature distributions, sequences of discrete temperature measurements are more suitable to show the evolution of temperatures with time. This chapter will explain in detail how the relationship between object temperature and measured infrared intensity was modeled. It will also explain how this model was calibrated using platinum- and nickel-resistors. The equipment used to measure the infrared intensities as well as the image processing and analysis procedures are also described.

5.3.2 Thermography

This subsection is meant to give a short overview of the theory of infrared thermography emphasizing the dependence on those parameters of interest for this work. A black body represents the simplest case for the relationship between the infra-red radiation and temperature of an object. The radiation field of a black body is determined solely by its temperature and is given by Planck's radiation law:

$$W_{\lambda_b}(\lambda,T) = \frac{2\cdot\pi hc^2}{\lambda^5}\cdot\frac{1}{\exp\frac{hc}{\lambda k_B T}-1}\cdot 10^{-6} \quad \left[\frac{W}{m^2\mu\mathrm{m}}\right] \tag{5.20}$$

With the following definitions:

- $W_{\lambda_b}(\lambda,T)$ is the spectral power surface density emitted by a black body at Temperature T, it is sometimes also called the spectral radiant emittance.
- $k_B = 1.4\cdot 10^{-23}\mathrm{J\cdot K^{-1}}$ is the Boltzmann constant
- λ is the wavelength
- h= $6.63\cdot 10^{-34}$J·s is Planck's constant
- T is the absolute temperature in K
- c is the velocity of light

Figure 5.5 shows Planckian distributions for various temperatures indicating the locations of the maxima as a function of temperature as predicted by Wien's displacement law[15]. This simple form is not directly applicable to typical objects studied with IR images. An object property called spectral emissivity $\epsilon(\lambda)$ modifies the IR spectra from their Planckian shape. $\epsilon(\lambda)$ is defined as the ratio of $W_{\lambda_b}(\lambda,T)$ over $W_{\lambda_b}^{blackbody}(\lambda,T)$. If $\epsilon(\lambda)$ is constant with λ and T the object is called a grey body. This is assumed for all objects measured in this work. However this does not imply that $\epsilon(\lambda)$ be constant across the complete surface of the object, even though an attempt was made to homogenise the emissivity for some of the objects measured in this study. A further assumption about the measured radiation fields has to be made. It concerns the angular distribution of the emitted radiation. It is

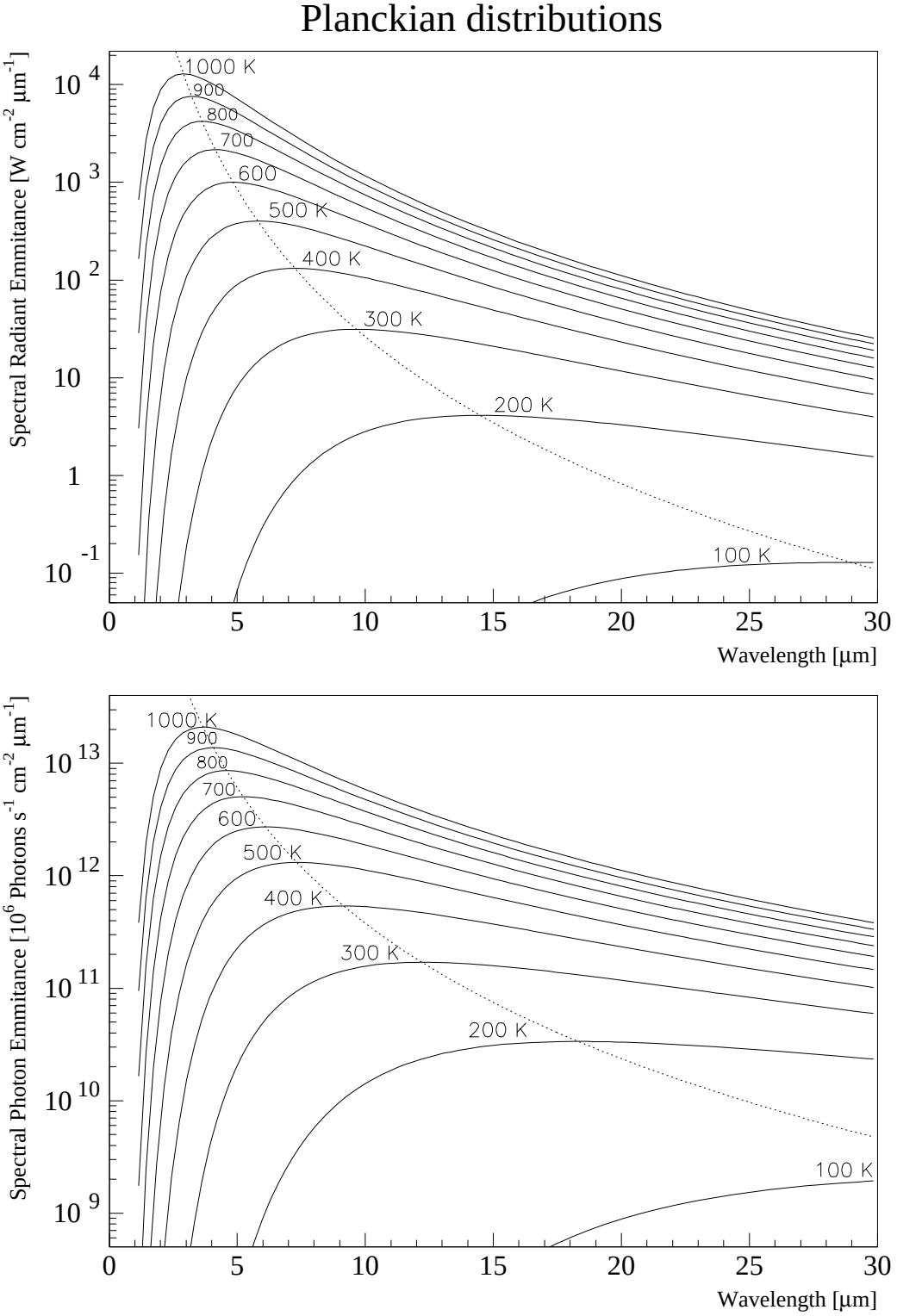


Figure 5.5: Logarithmic plot of Planckian distributions with locations of maxima as predicted by Wien's displacement law. The top curves show the distributions of the emitted spectral power whereas the bottom curves show the distributions for the rate of emitted photons.

assumed that the objects have Lambertian surfaces. This means that the $W_{\lambda_b}(\lambda, T)$ is not a function of the angle α between the surface and the observation line. This is usually true for perpendicular observation but starts to be wrong as α approaches 70^o where $W_{\lambda_b}(\lambda, T)$ goes through a maximum and then falls to 0 at $\alpha = 90^o$. This is understandable as the emittance of an object according to Kirchhoff's law is proportional to its absorbtance[13] which is solely given by the objects refractive index and the angle under which its surface is observed. Since most objects in this study have been measured perpendicular to their surfaces the assumption of Lambertian surfaces holds.
The spectrum emitted by the object does not fully describe the dependence between the measured detector signal and the object temperature, because the spectral response of the detector is neither constant nor unity.
Because the detector elements used for IR scanning type applications are usually IR-photon sensitive semiconductors which have a defined efficiency of detecting a single photon of any given wavelength, the relevant quantities in describing the input to the detector response is not the spectral radiant power but the spectral photon emittance which is defined as the number of photons per wavelength interval and unit time emitted by one unit of surface of the object. It is obtained by dividing Planck's law by the energy of a photon hc/λ and is also shown in figure 5.5. The curves of the spectral photon emittance are similar to those of the spectral radiant emittance but have a less abrupt fall after the maximum which is shifted towards longer wavelength. This means that for a black body at T=300 K the maximum of the spectral photon emittance is at 12.2 μm whereas the corresponding power density would have a maximum at 9.7 μm. The photon count version of the Stefan-Boltzmann law only shows an increase of the number of photons emitted from a black body with the third, rather than the fourth power of T.

5.3.3 The Scanner

The process of scanning an image falls into two main tasks. First to subdivide an image of the object into a field of equal sized samples and to focus the radiation from each sample onto a detector for a fixed time. This is usually done with a moving mirror system. Secondly the photons from each sample will be integrated in the detector over the exposure time for each sample. In the detection process the incoming photon spectrum will be weighted with the detectors spectral response and integrated over wavelength. Thus an electrical signal proportional to the weighted energy deposition, usually in form of a voltage is generated for each sample.
The IR-scanning equipment used was an AGEMA Thermovision 900 LW system. The system was controlled by a two processor VME unit with internal disk and two dedicated 16 bit IR-data busses for the scanner readout. In addition a VGA monitor, a colour video printer, a video recorder and two germanium lenses were part of the system. The system was connected to the central Oxford Unix cluster via ftp on ethernet for data backup. The scanning was done by two rotating mirrors, a single sided mirror for the vertical and a polygon mirror for the horizontal scanning. The detector element used in the scanner is a CdHgTe detector[14] which has a usable spectral response in the range from 8 to 12 μm. This corresponds to a range of measurable temperatures from -30 ^oC to 1500 o^C. The CMT detector is a semiconductor with small band gap sensitive to infra red radiation. The detector-amplifier-ADC system has a sensitivity of 0.08 ^oC at 30 ^oC. This is defined by the fact that an 0.008 ^oC change of temperature at 30 ^oC will result in a change of

[13] this is strictly true only for opaque objects
[14] often abbreviated as CMT

the ADC output value of 1. The scanner acquires images with 272 samples per line and 136 lines per image. The image frequency is 15 Hz with a line frequency of 2.5 kHz. At 50% modulation this corresponds to a spatial resolution of 230 elements per line. The scanner is cooled with liquid nitrogen and can operate in an ambient from -15 ^oC to +55 ^oC. The two lenses used were a 10^o and a 20^o lens with minimal focal distances of 1.2 and 0.6 meters and angular resolutions of 0.76 mrad and 1.5 mrad respectively. The signals from the detector were digitised by a 12 bit ADC and stored after a linearisation process correcting for nonlinearities of the Detector-Amplifier-ADC chain. The response function of the scanner system has been empirically determined for each combination of lens and scanner by measuring calibrated black body sources. It is described by a fit to the calibration data of the shape:

$$I(T) = \frac{R}{e^{\left(\frac{B}{T}\right)} - F} \tag{5.21}$$

In equation 5.21 I(T) is the ADC value, also called thermal value; R, B and F are the parameters of the fit and T is the object temperature in Kelvin. Typical values for the fit parameters are, R=221800, B=1434 and F=5.62.

5.3.4 Ambient Model

The detector response model and the spectral photon emittance do not fully describe the relation between the ADC reading and the object temperature because other sources of radiation and attenuation have to be considered. Figure 5.6 shows a graphical representation of the ambient model used to describe the radiation field measured by the scanner. The parameters of the ambient model are:

T_{ob} Temperature distribution of the observed surface of the object.

ϵ_{ob} Emissivity distribution of the observed surface of the object.

H relative Humidity of the atmosphere between scanner and object.

T_{at} Temperature of the atmosphere between scanner and object.

T_{am} Temperature distribution of the ambient surfaces surrounding the object.

T_w Temperature distribution of the window.

α_w Absorbtance of the window.

ρ_w Reflectivity of the window

These can be grouped into four categories which are discussed individually below.

Atmosphere

The influence of the atmosphere on the IR image is determined by three parameters:
the relative humidity H given in %,
the temperature T_{at}
and the distance between object and camera D_{ob}.
These parameters determine the transmission of the atmosphere τ_{at}. The transmission in

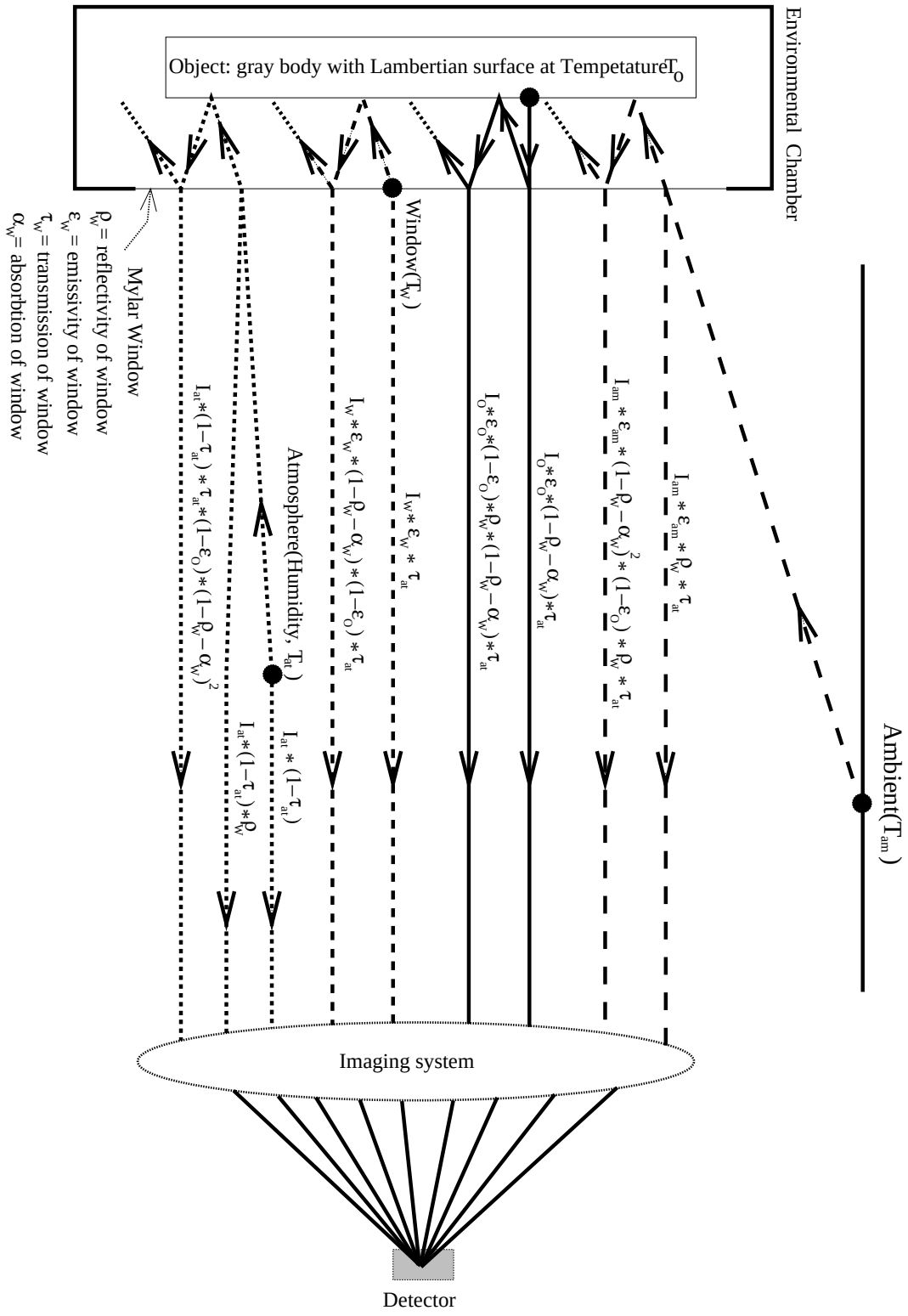


Figure 5.6: Model for the radiation sources contributing to the measured intensities. The object under study is assumed to be opaque.

the wavelength window from 8 to 12 micron is calculated as a function of the atmospheres total water content W in units of g/m^3.

$$W = H \cdot e^{\left(K_1 + K_2 \cdot \tilde{T}_{at} + K_3 \cdot \tilde{T}_{at}^2 + K_4 \cdot \tilde{T}_{at}^3\right)} \cdot 1\frac{\mathrm{g}}{\mathrm{m}^3} \tag{5.22}$$

$$\tau_{at} = X \cdot e^{\left(-\sqrt{D_{ob}} \cdot \left(\alpha_1 + \frac{\beta_1}{\sqrt{W}}\right)\right)} + (1 - X) \cdot e^{\left(\alpha_2 + \frac{\beta_2}{\sqrt{W}}\right)} \tag{5.23}$$

In the above equations K_1, K_2, K_3, K_4, α_1, α_2, β_1, β_2 and X are empirical constants the values of which are shown in table 5.2. $\tilde{T}_{at}$ is the temperature in ^oC [16] [15]. The

Parameter	Value	units
K_1	1.5587	1
K_2	$6.939 \cdot 10^{-2}$	$^oC^{-1}$
K_3	$-2.7816 \cdot 10^{-4}$	$^oC^{-2}$
K_4	$6.8455 \cdot 10^{-7}$	$^oC^{-3}$
α_1	$3.749 \cdot 10^{-3}$	$\mathrm{m}^{-\frac{1}{2}}$
α_2	$8.12 \cdot 10^{-2}$	$\mathrm{m}^{-\frac{1}{2}}$
β_1	$2.638 \cdot 10^{-3}$	$\mathrm{m}^{-2}\ \mathrm{g}^{\frac{1}{2}}$
β_1	$2.431 \cdot 10^{-2}$	$\mathrm{m}^{-2}\ \mathrm{g}^{\frac{1}{2}}$
X	1.06	1

Table 5.2: Parameters used in the calculation of the transmittance of air for IR radiation.

dependence of τ_{at} on H and $\tilde{T}_{at}$ and the resultant transmission of 1 meter of air is shown in figure 5.7 over the complete validity range of equations 5.22 and 5.23 and in a zoom over the typical operational range.

Object

The parameters describing the object are its emissivity ϵ_{ob}, its temperature T_{ob} and the assumption of an opaque object. The emissivity of the object has usually been equalized across all surfaces by spraying them with white powder as described in section 5.2. It is assumed that ϵ_{ob} does not depend on T_{ob} (grey body). The influence of the powder on the thermal properties of the objects is small. It partially acts as an insulator and sometimes, in particular on the cooling pipes which have small surface areas shifts the visible surface temperature towards that of the ambient air because the surface to the ambient is bigger than that to the pipe. This effect is small for the other components of the object.

Window

The window is described by three properties. The reflectivity ρ_W, the absorption α_W and the temperature T_W. The emissivity of the window can be calculated under the assumption of a semi transparent window from ρ_W and α_W by the following equation [15].

$$\epsilon_W = \frac{(1 - \rho_W) \cdot \alpha_W}{1 - \rho_W \cdot (1 - \alpha_W)} \tag{5.24}$$

It is assumed that the optical properties do not vary with angle which is not strictly true since there will be interference modulation from the thin foil. These interference fringes can be seen in figure 4.13 as the envelope of the absorption features in the spectrum. It is also assumed that there is no variation with T_W. The optical properties will vary with

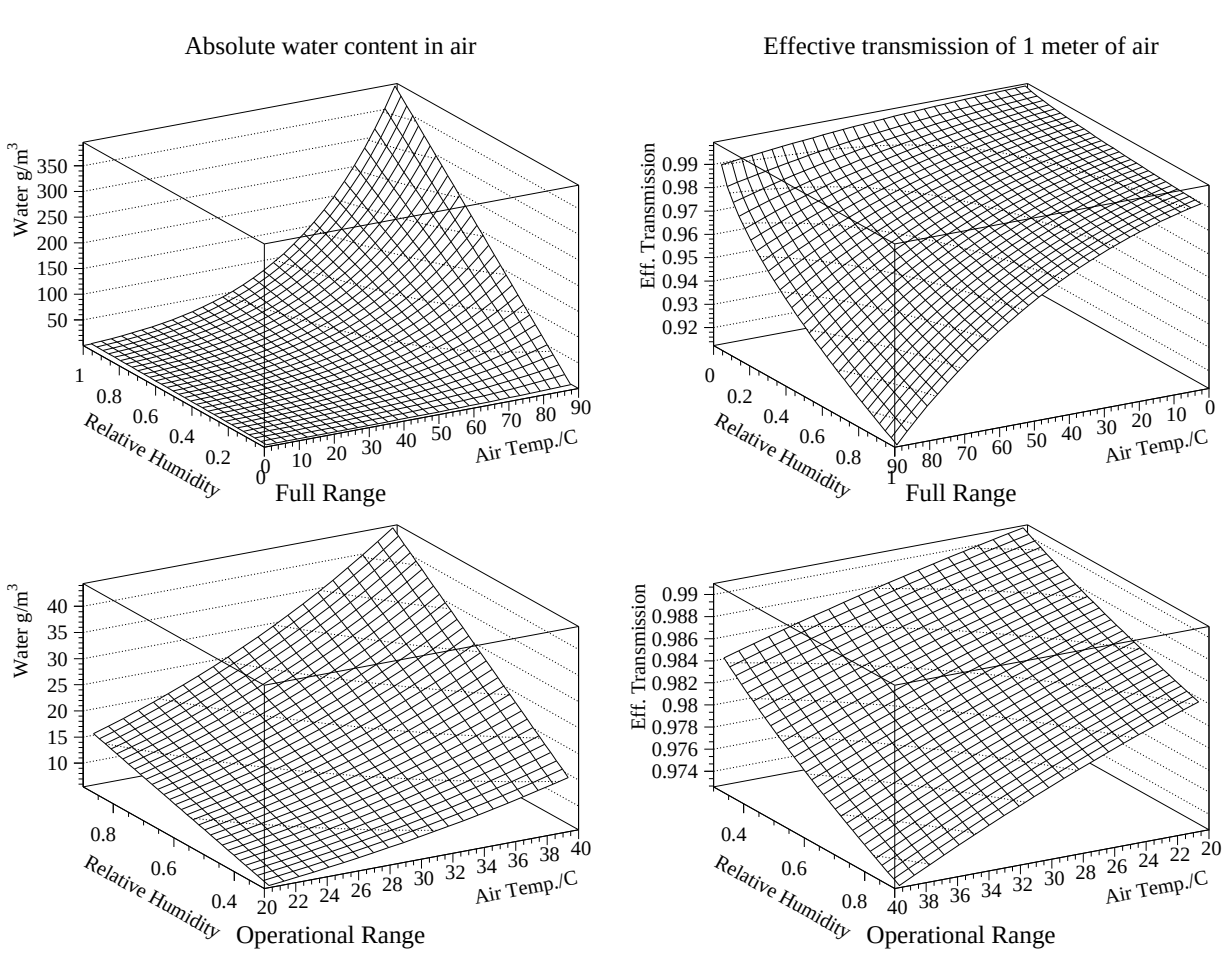


Figure 5.7: Absolute water content of air as calculated from equation 5.22 and resultant effective transmission of air from equation 5.23. Note the opposite orientations of the temperature and humidity axis in the water content and the transmission plots.

the temperature of the object observed through the window as the observed spectrum changes, but this is also neglected.

Ambient

The ambient is assumed to be at equal temperature and emissivity which is true to a good approximation because during the measurements the object is usually situated in the laser safety area where it is enclosed by two layers of black curtains and a black wooden lid. The only variation in the ambient emissivity comes from the aluminium of the optical table. Because the ambient parameters $\epsilon_{am}$and T_{am}always appear as their product in the ambient model, the emissivity is arbitrarily set to unity and only T_{am}remains as a free parameter. This means that $T_{am}$for the ϵ_{am}=1 case is usually lower than the real surface temperatures of the ambient objects.

5.3.5 Image Processing

The scanner system controller offered user friendly means of manipulating the IR images and convert them into temperature readings via an OS9 based software package called ERIKA. Because the scanner system was available only on a temporary loan basis for periods of 4 weeks and a version of ERIKA for Unix systems or PCs was not available an

image manipulation tool had to be written for a Unix or PC system. The central part of the image analysis system is the SEMPER 6 package that has already been described in section 5.2. It was operated on a PC. From a description of the binary file format used by the ERIKA system to store single images and sequences of images it was possible to write an interface that imports the ERIKA images into SEMPER. Furthermore it was necessary to read back the auxiliary information stored by ERIKA together with the images. Besides a complete description of the scanner setup they include the calibration constants R,B and F from equation 5.21. A C-program extracts all header information and writes it to a small ASCII file from where it can be imported into SEMPER. This step had to be done separately because the binary storage format for numerical values used by the ERIKA system is not compatible with the PC or UNIX representation and some byte swapping operations had to be carried out before data storage. The conversion from scanner ADC values to temperatures is calculated inside SEMPER using the model described in figure 5.6. This model is more elaborate than that implemented in the AGEMA software because it also allows for a semi transparent window in front of the object and takes into account all sources of radiation up to first order reflections in both the window and the object.

5.3.6 Calibration

The process of calibrating an IR image or sequence refers to the transformation of the raw ADC values of the image into absolute temperatures. Although the ERIKA system offers various calibration procedures these were not used for three reasons.

- The scanner system was not available for a long period of data analysis.
- There were no provisions for calibration in the presence of a semi transparent window.
- Calibrations done inside ERIKA require the object to be either of even temperature well above ambient temperature (emissivity calibration) or to be of even emissivity and well below ambient temperature (reflected ambient calibration). These conditions could not easily be created in this setup.

The acquisition and calibration of an IR image or in the more general case a sequence of images is explained as a list of processing steps below. A sequence of images is stored by ERIKA with a complete header at the beginning of each image even though parameters in the header such as ambient temperature or humidity can not be changed between images. A complete sequence can only be calibrated in one step, if the environmental conditions were identical for all images of the sequence.

1. When the sequence is acquired a set of environmental parameters relevant to the calibration can be entered by the user. These parameters are:

   - The distance between object and scanner
   - The temperature of the atmosphere
   - The humidity of the atmosphere
   - The temperature of the surrounding ambient objects

   They determine the calibration in a simple case without a window in front of the object.

2. After the sequence is acquired it is stored on disk or tape accessible for the image manipulation PC running SEMPER and the central unix cluster. This is usually done on one of the disks of the unix cluster and this file is referred to as the ERIKA sequence (image) file.

3. Resistance thermometers in the area of the image are being scanned while the sequence is taken and their data (resistance and internal time) is stored on disk in what is called the discrete data file. It is crucial that for the two asynchronous systems[15] a common time is written into the log book.

4. Those parts of the auxiliary information in the ERIKA file like the environmental parameters and the calibration values of equation 5.21 common to all images in the sequence are extracted and written to the parameter file. This is done by a C program on the central unix cluster.

5. The first raw image to be calibrated is read by SEMPER from the ERIKA file and displayed on the PCs screen.

6. On the image the positions of the sensors are graphically marked with a SEMPER program. This program writes the following data to the IR extraction file:
   Averaged ADC values in the marked are
   Errors on the mean of the averages
   Number, Position and dimension of the marked area
   The locations of the marked areas are also stored in a marker file. The marker file can be used to automatically extract information from other images without having to mark them again. In this way information from images in the same sequence can be added to the IR extraction file if they have the same environmental parameters. The marker file can also be used for later extractions on images of the same object.

7. A list of correspondences between the numbers of the marked areas and the numbers of the sensors in the discrete data files is written to a correspondence file.

8. The ERIKA internal times of the images appearing in the IR extraction file are converted to times used by the discrete temperature logging system and written to a time file. This can be done automatically by reading the relevant image headers and using the log-book information to do the transformation. The default is that the time file is created manually.

9. The discrete data file is scanned and the values of those thermometers selected in the correspondence files and their errors are written to the discrete extraction file. The values written out are linear interpolations between the two thermometry readings before and after each time given in the time file. The errors are the combined errors from the time interpolation and the individual errors of each thermometer reading.

10. The parameter file as well as the discrete- and IR-extraction files are read by a fitting program. This program tries to fit the ambient model described in figure 5.6 to the sets of IR and thermometry data. It takes into account the errors on both measurements and does a fit to data points with true two dimensional errors.

[15] The IR-scanner and the thermometry system

All parameters of the ambient model can be fixed if needed. This is usually the case for ρ_w and α_w after they have been determined once because they are not expected to vary. This approximation is true if any change of the overlap between the window transmission spectrum and the object radiation spectrum is small which is the case for the temperature range observed here. Figure 5.8 shows the results of a fit leaving all seven[16] parameters free. The input data comes from measurements of a Peltier element sprayed with the same white powder as the object and placed behind the mylar window. Currents of different magnitude and polarity were passed through the peltier so that it's surface varied in temperature over a large range. The excess power was lost through a heat sink mounted on the back of the peltier element. The parameter values of the fit are given in table 5.3. The correlation matrix[17] of the parameters is given in table 5.4. It can

Par.	Value	Error	$d\chi^2/dPar$	start	range	
ϵ_{ob}	.68507	.13026E-01	-.41130E-01	0.67	0.5	1.0
T_{am}	306.75	.37083	.93467E-01	297	295	318
ρ_w	.20925	.12685E-01	.20746E-01	0.22	0.1	0.3
α_w	.11770	.11290E-01	.96755E-02	0.14	0.01	0.2
H	.74371	.22013	.39639E-02	0.67	0.4	1.0
T_{at}	301.51	2.1391	.96375E-02	300	292	306
T_w	300.60	1.0015	.19601E-01	300	292	306

Table 5.3: Results from a fit of the ambient model to calibration data. The errors, first derivatives, starting values and ranges are also shown. The total χ^2 was 28.8 for 7 parameters and 35 points in the fit. Temperature values are given in K.

Par.	global	ϵ_{ob}	T_{am}	ρ_w	α_w	H	T_{at}	T_w
ϵ_{ob}	.99605	1.000	-.002	.377	.591	.124	.044	-.005
T_{am}	.99626	-.002	1.000	-.070	.187	-.274	-.332	-.370
ρ_w	.99637	.377	-.070	1.000	-.466	-.159	-.101	-.066
α_w	.99489	.591	.187	-.466	1.000	-.028	.107	.181
H	.99658	.124	-.274	-.159	-.028	1.000	-.278	-.274
T_{at}	.99644	.044	-.332	-.101	.107	-.278	1.000	-.333
T_w	.99625	-.005	-.370	-.066	.181	-.274	-.333	1.000

Table 5.4: Correlation coefficients of the fit parameters from table 5.3. The column global refers to the absolute value of the correlation between the χ^2 and the parameter in question.

be seen that, although all parameters are in a reasonable and expected range they are not negligibly correlated. The sum of ρ_w and α_w is consistent with a graphical evaluation of the transmission spectrum shown in figure 4.13 of chapter 4. The value for ρ_w alone is compatible with the reflections expected from the refractive index of mylar of 1.64. The two parameters τ_w and ρ_w are strongly correlated with each other and with T_{ob} as expected from their role in the model. The strong correlations and the almost linear dependence of the ADC values on the ambient temperature is the reason for the strong dependence of the fit results on the starting values and parameter range. In general data over a large temperature range is needed and as many parameters as possible should be fixed if true values for the parameters have to be found. The results shown in table 5.3 have

[16] T_{ob} is not a parameter of the model but the variable to be described by the model

[17] normalised correlation coefficients

been intensively checked. The χ^2 of the fit has been examined by varying each parameter separately[18] over its complete range shown in table 5.3 and for all parameters a distinct symmetric, parabolic minimum has been found. In cases were the fit settled into a different local minimum at least one parameter was at or very near to one of its boundary values. The total χ^2 per degree of freedom is $28.8/(35-7) = 1.02$. From this it can be assumed that the estimate of the error from the time interpolation for the discrete measurements and for the mean ADC values in the marked areas on the IR images is realistic.
If only a parameterisation[19] is needed, as in the case of the calibration of an image or a sequence a good fit with small χ^2 can always be obtained with a polynomial of 3rd order.

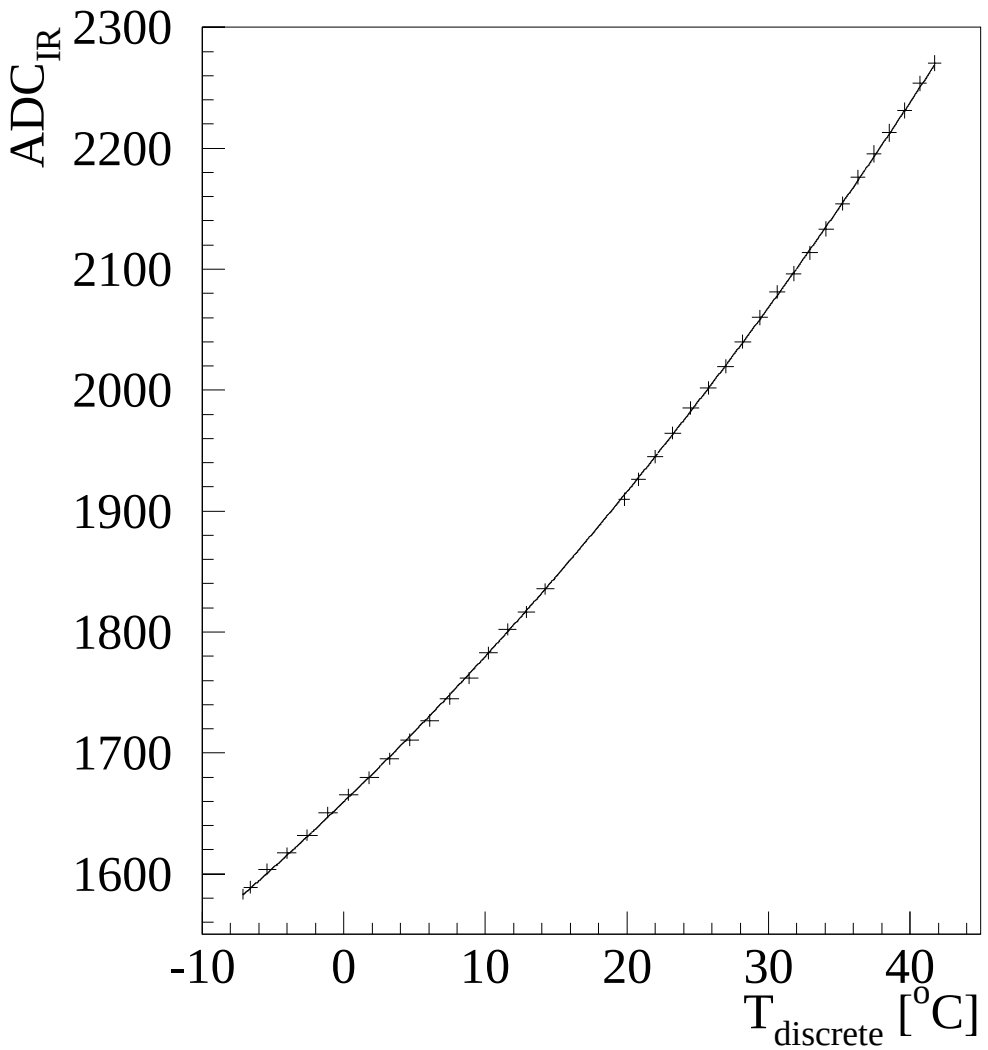


Figure 5.8: Fit of the ambient model to calibration data. All seven parameters are left free. The parameter values can be found in table 5.3 and the correlation coefficients are shown in table 5.4. The crosses represent the error bars in both directions.

[18] while keeping all other parameters fixed
[19] of the dependence ADC-value to true temperature

Chapter 6

Thermal and Mechanical Measurements

This chapter will describe the thermal and mechanical measurements of three different objects with the two setups described in chapter 5, subsection 5.2.4. Conclusions for each of the measurements, concerning the technique and the test object will be given. The three objects were, in order of time, a single silicon crystal with connection bars, an aluminium plate, and the tracker modules.

6.1 Single Crystal

Initial measurements of a single silicon crystal with connection bars (see figure 4.3) with the He-Ne-Laser setup described in chapter 5, subsection 5.2.4 are discussed in this section. The measurements had three main purposes.

- To acquire experience in the operation of an ESPI system, necessary for the construction of the next and bigger system.
- To learn techniques for the analysis of ESPI fringe patterns by analysing the relatively simple fringe patterns from a single crystal.
- Learn which distortions were incorporated into the detector by gluing the connection bars to it at an elevated temperature of 60 ^{o}C.

6.1.1 Mounting

The crystal was held in an aluminium frame which itself was positioned in a graphic plate holder. The fixation was achieved by clamping the protruding part of connection bar[1] 1 to the frame with a small screw clamp. Pivoting the detector in only one point allowed it to expand and distort freely. Because the crystal was supposed to be heated by illuminating one side with a high power halogen projector bulb, and the light needed to be restricted to the crystal area, two sets of apertures were positioned in front and behind the crystal. The faces of the apertures on the illuminated side of the crystal which were visible from the lamp were reflecting. All other apertures were matt black. Both sets of apertures were positioned as close to the edge of the detector as possible allowing a minimum of stray light to reach the inside of the measurement volume. Because the available laser power was very small the problem of stray light was more serious than with the later setup.

[1] see figure 6.1

6.1.2 Heating

On the side exposed to the heating lamp the crystal was blackened with a very thin film of black ink. This was done in order to absorb the radiation from the lamp more efficiently than its aluminised surface would have done. The front of the crystal was sprayed with white talcum powder through a fine steel mesh with a mesh pitch of 1x1 mm^2. This way the surface was made to reflect diffusely and a precise determination of the location of the connection bars behind the crystal, via orientation on the grid was possible, even though the bars could not be seen from the front. The heating system consisted of a 200 Watt halogen projector bulb and a blackened tube guiding the light to the detector backside. Radiation heating was the only easily available method of heating the detector without mechanically touching it and without changing the refractive index of the air (heating it) in front of the detector which would have deteriorated the ESPI measurements. The power of the lamp could be varied from 0 to 200 W.

6.1.3 Temperature Measurements

At the time the ESPI measurements of the single crystal were made an infra-red scanning system was not yet available and IR-measurements have been done two month later. The temperature was also measured with a PT-100 resistance thermometry system. The thermometers were read out by industrial temperature transmitters which translate the resistance of a thermometer into a current signal. These signals were than read by an 8-Channel ADC card in a PC and recorded with a LAB-NOTEBOOK[2] system. The location of the sensors on the crystal is shown in figure 6.1. The temperatures recorded during the cool-down of the detector after a 10 minute heating period at maximum light intensity is shown in figure 6.2 as a function of time for sensors 0,1,2,5,6 and 7. The corresponding IR-images, measured in a separate experiment, are shown in figure 6.3 for 0 to 24 seconds[3] after the heat has been switched of. Figure 6.4 shows the differences in temperature from figure 6.3 with respect to the first image. The IR-images have been taken after the thermometers have been removed. The thermal grease that was used to attach them can still be seen in the images. This is due to two effects. Firstly the emissivity of the surface covered with the grease is different and secondly the grease modifies the thermal properties of the detector by adding heat capacity and surface to it. The effect of the inhomogeneous emissivity can not be eliminated by looking at the difference images as they depend in exactly the same way on the emissivity as the original images. This can be clearly seen by looking at figure 6.4. Effects that can be corrected by a simple difference have to be of an additive nature rather than multiplicative as the emissivity. An emissivity equalization calibration can be done if the true temperature of all observed points is known[4] and the observed intensity is dominated by radiation from the object[5]. The online calibration of the IR-images, as described in chapter 5.3.6 resulted in the following parameters:
T_{am}=T_{at}=25 ^oC , H=60%, ϵ_{ob}=0.67, D_{ob}=50cm.
Since there was no window present in this image the window parameters (ρ_w, α_w, T_w) have been all set to 0.

[2] a product of Keithley

[3] in 2 second steps

[4] This is usually the case if the object has a homogeneous temperature distribution

[5] This is the case when the object is significantly hotter than it's ambient and if the reflectivity is low

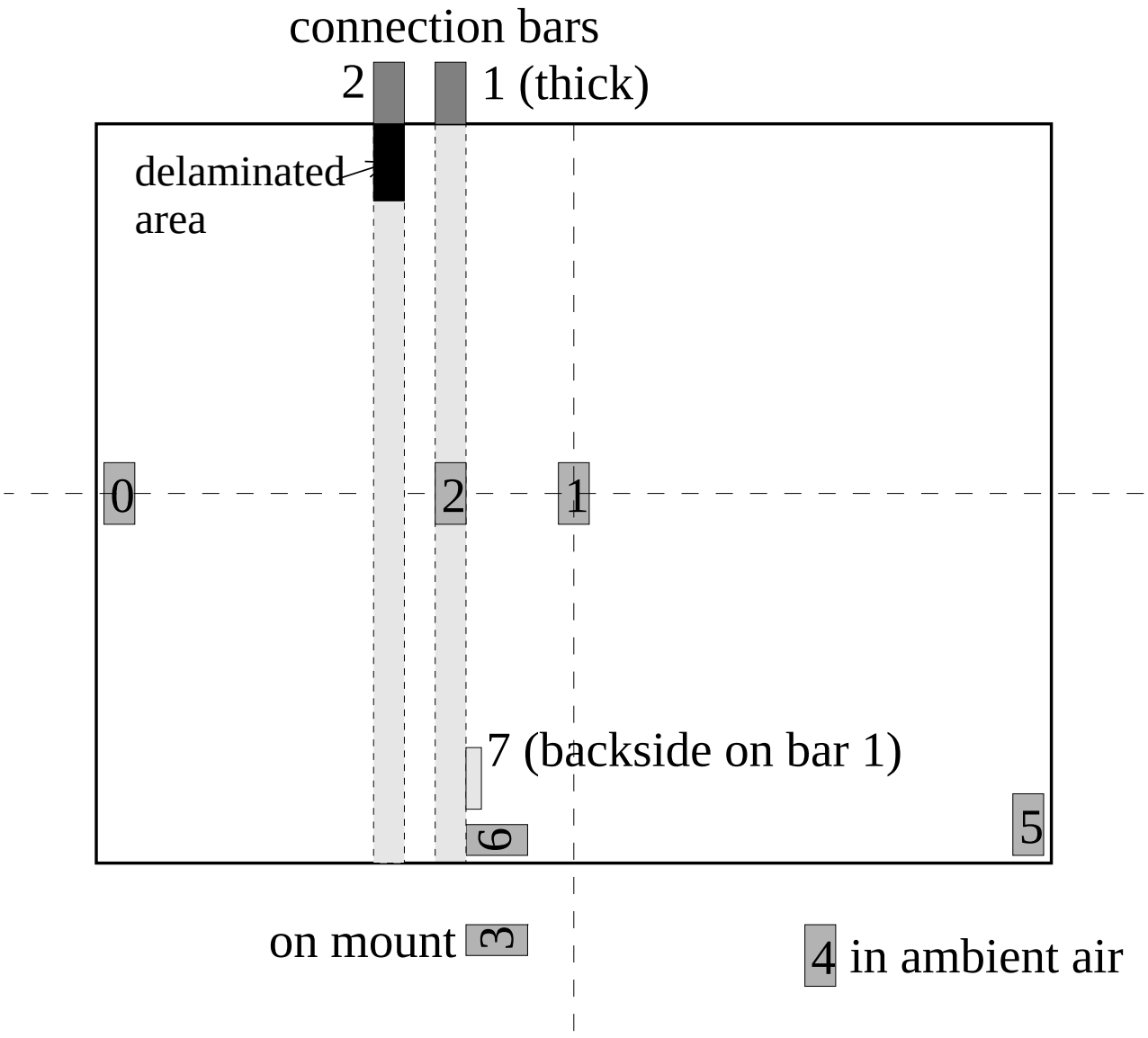


Figure 6.1: Location of the sensors on the crystal. The delaminated area under the thinner of the two connection bars is also shown. The detector is viewed from its back side, with the bars behind the detector. Sensor number 7 is attached to bar 1.

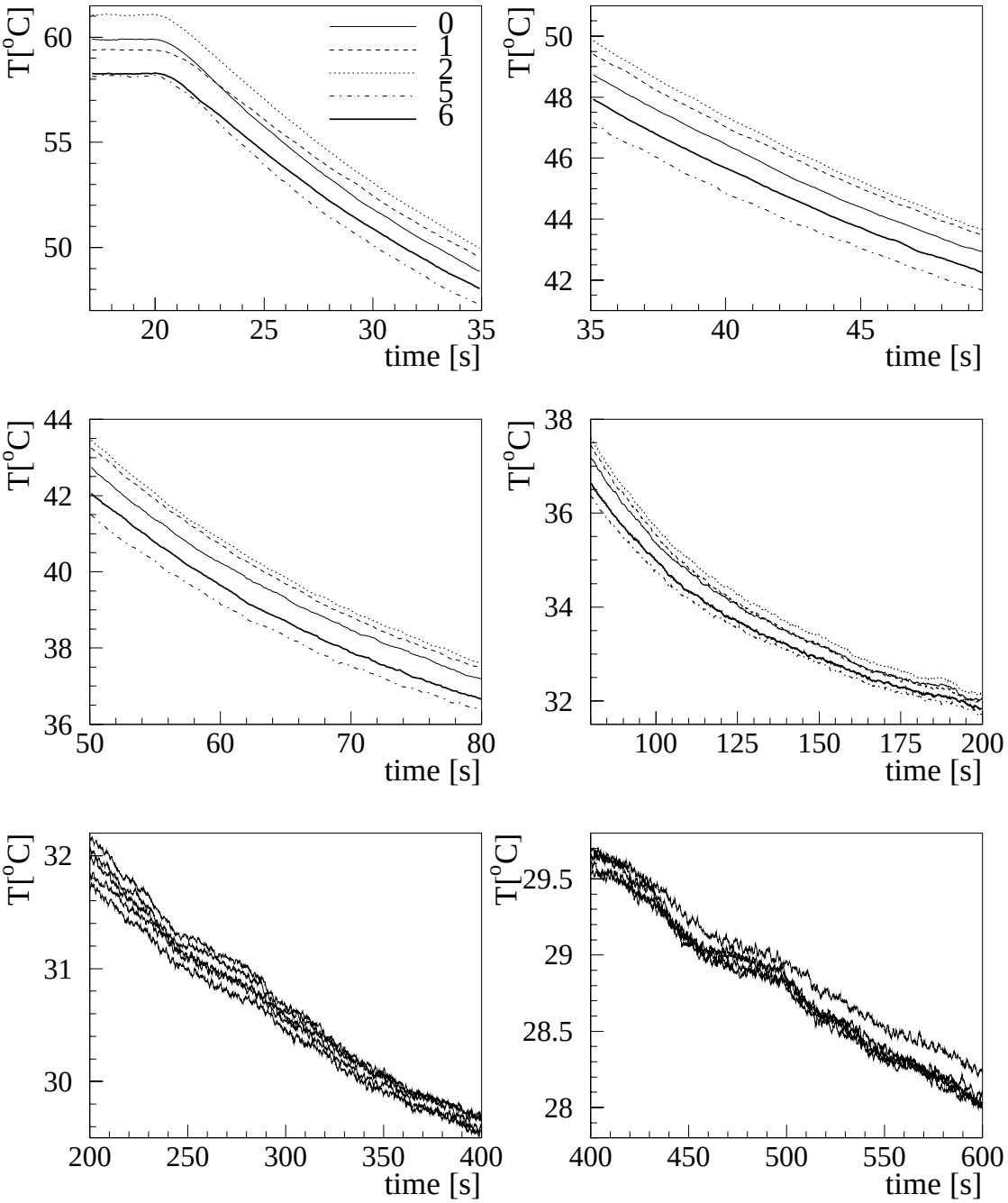


Figure 6.2: Temperature of sensors 0,1,2,5,6 on the front of the silicon detector as a function of time during cool-down. The heat has been on for 10 minutes at maximum power and switched off at t=20 s. The plot has been cut into 6 sections because of the large temperature changes compared to the differences in the sensor readings.

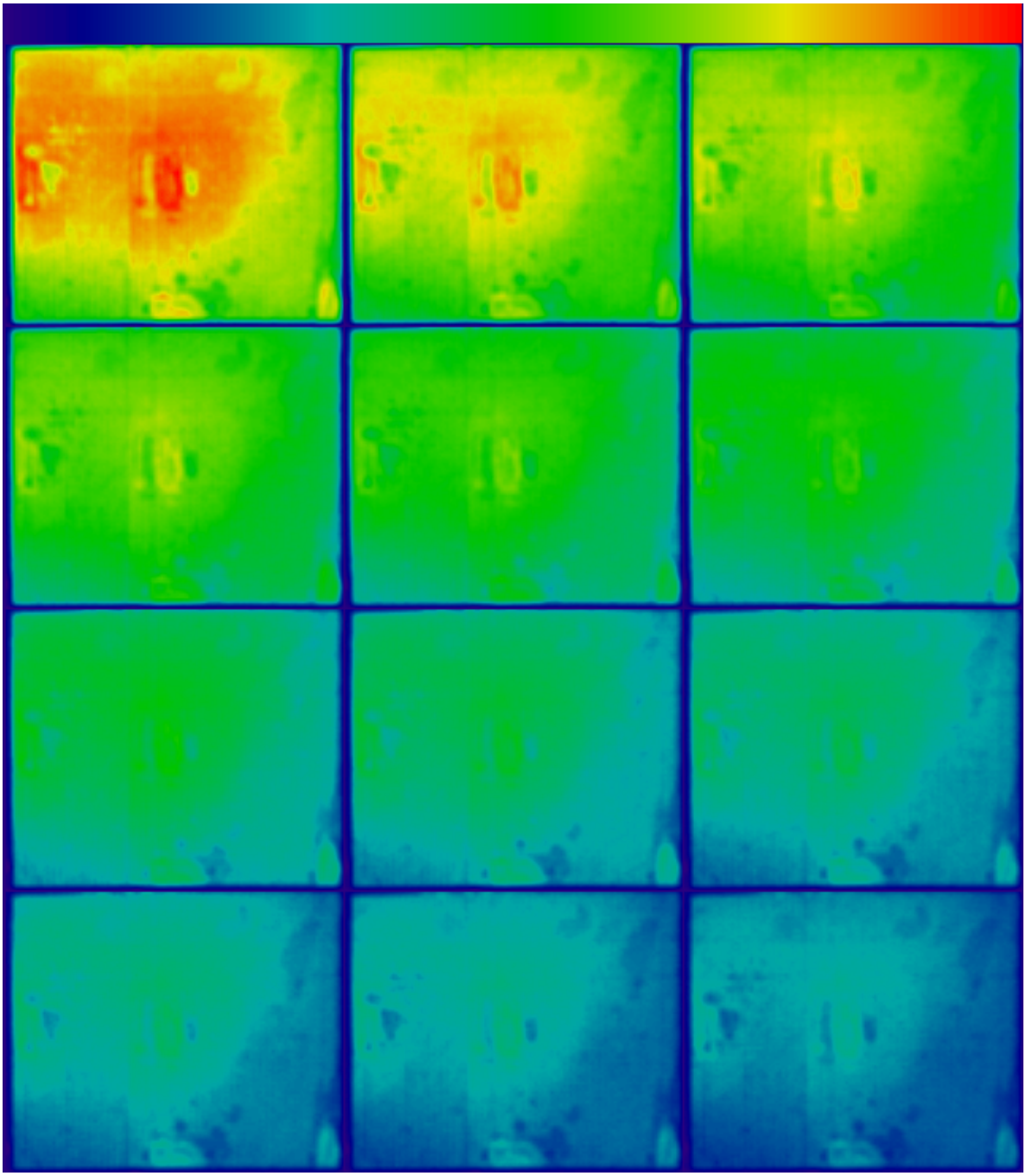


Figure 6.3: Temperature of the silicon detector during cool-down. The heat has been on for 10 minutes at maximum power. The images are taken at 2 second intervals. The first image (directly after heat off) is top left. The last image is bottom right. The colour scale at the top of the images represents temperatures from 12 to 60 ^oC.

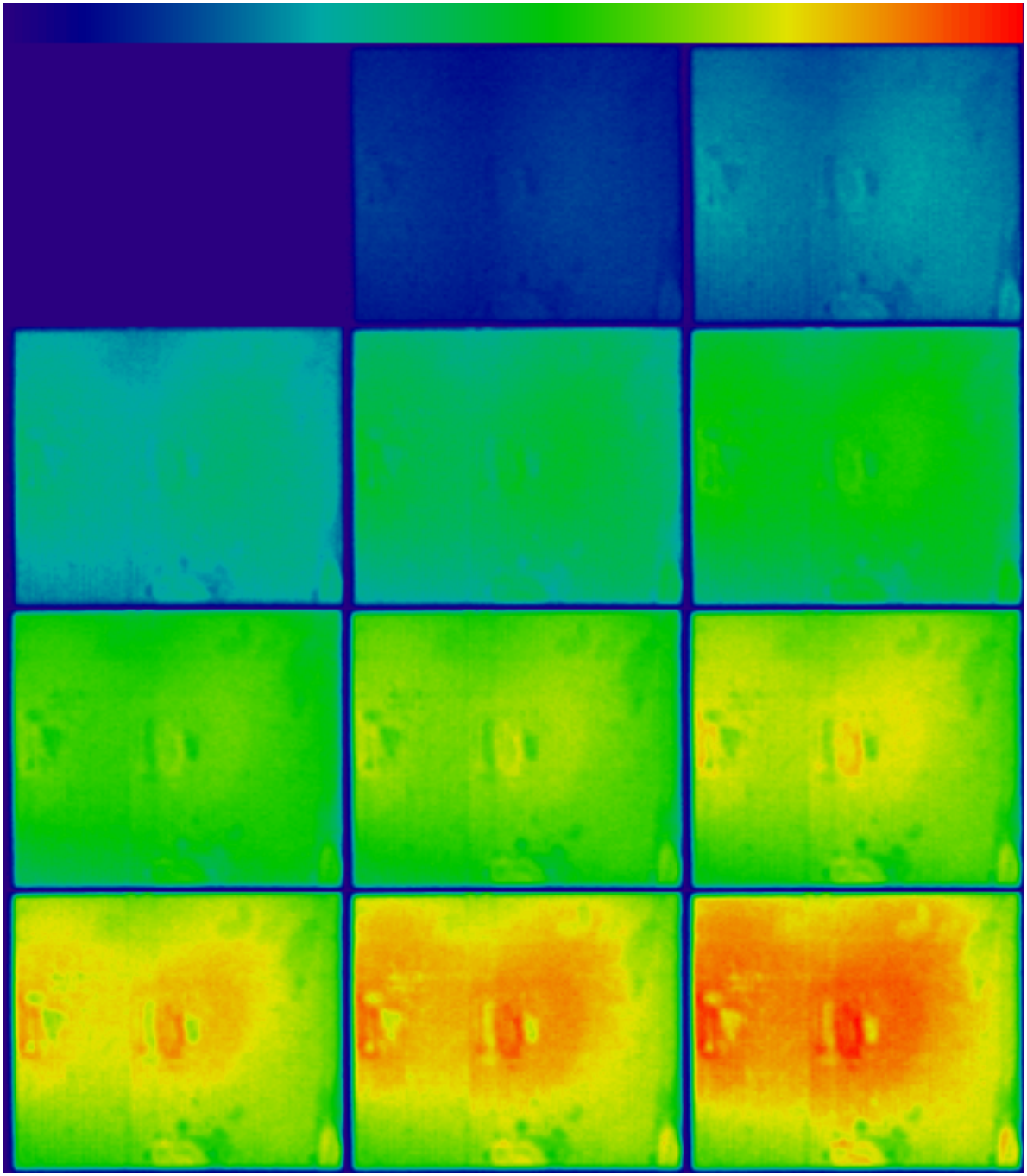


Figure 6.4: Difference of the temperature of the silicon detector during cool-down with respect to the first image in figure 6.3. The heat has been on for 10 minutes at maximum power. The images are taken at 2 second intervals. The first image (directly after heat off) is top left. The last image is bottom right. The colour scale at the top of the images represents temperature differences from 0 to 23 ^oC.

6.1.4 Distortion Measurements

Aim

The main purpose of these measurements was to prove the feasibility of the ESPI measurements of silicon detectors and to learn about the techniques of analysing and understanding fringe pattern data. A second aim was to estimate the distortion that resulted form the fact that the silicon detector had been glued to the connection bars at 60 ^oC to reduce the curing time. The idea was to continuously measure the distortion while heating the detector from room temperature until it reached the temperature where it was glued. Assuming that at this point, before the glue polymerised the shape was flat this measurement would result in a shape reconstruction from deflection measurements with unknown starting shape. An attempt to measure the shape of the crystal on an optical scanning table had been made before, but since the temperature of the table was not stabilised and the detector had to be fixed in three points to the measurement table the shape measurements had large uncertainties due to stresses that could have built up from temperature changes during the measurement.
The last and in retrospect most important target was to try and see whether a delamination of connection bar 2 could be seen in the distortion fringes. In the first 10 mm from the top of the images bar 2 is delaminated from the surface as indicated in figure 6.1 and it was hoped that a small change of fringe shape might have been visible when comparing fringes from the delaminated and undelaminated areas.

Measurements

Figure 6.5 shows the out of plane distortion of the detector in the cool-down phase after a 15 second heating period with full light power. The object geometry underlying the images is shown in figure 6.1. The detector was sprayed with white talcum powder to make it diffusely reflecting on the imaged surface. The ESPI setup was the one using the He-Ne-Laser as described in subsection 5.2.4. The heater lamp has only been switched on for 15 seconds and the detector has not been heated to its equilibrium temperature because the very hot surfaces of the detector and the mounting plates induced convection currents which disturb the images. The cooling down of the mounting plates could also have added distortions because the mounting point itself could have moved. The cool-down phase was more easy to measure because there was no stray light from the heating system and the distortions change much slower. At these temperatures[6] the detector cools down slowly, but heats up very rapidly. It was found that after heating the detector to its equilibrium temperature the distortion during the cooldown is fast enough to produce many more fringes in the time it takes to acquire an image[7] than can be resolved in one image. Furthermore the number of fringes in a cool-down period from 60 ^oC would be very large and the amount of image data to store would be prohibitive. The expected maximum distortion from the optical scanning table measurements is 180 μm which corresponds to 570 fringes. With a resolution limit of approximately 20 fringes per image this would correspond to a minimum of 29 images. In practice the number of images required would be much larger as it will not easily be possible to adjust the time between images such, that the image will contain 20 fringes. It is advisable to produce less fringes per image to have some robustness against the loss of an image. If the distortion between consecutive

[6] after only 15 seconds of heating the temperature of the detector rises by approximately 2 ^oC above ambient.

[7] approximately one second

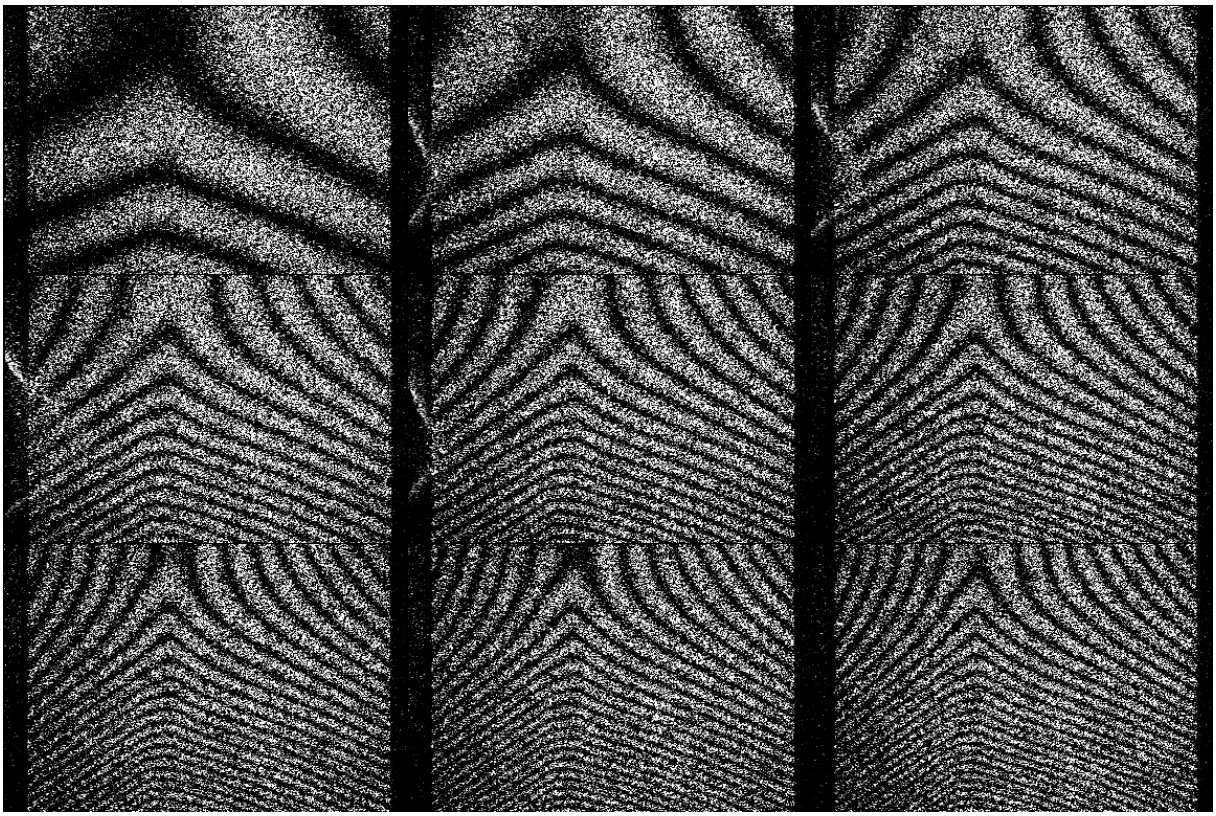


Figure 6.5: Cool-down of the silicon detector after a 15 second heating period. The images are differences with respect to the first image directly after the heat was switched of. The time with respect to the first image can be found in table 6.1.

measurements is always near to the fringe resolution limit, the loss of one image breaks the measurement sequence and images before and after the lost image can not be related. A last restriction is the increase in cool-down speed at higher temperatures. Because the detector looses heat at least proportionally to the temperature difference with the ambient, at a detector temperature of 60 ^oC the cool-down rate would be approximately 20 times faster than the one observed here. This would again lead to fringe resolution problems. The images in figure 6.5 are difference fringe patterns with respect to the first image directly after the heat was switched off. The time intervals between the images are approximately 2.1 seconds but they are not completely equal because the time it takes to store the images onto the hard disk of the computer varies. The time of each image is given in table 6.1. The fringe pattern of figure 6.5 can to a certain extend be interpreted without calculating an elevation map. Assuming that the fixation point remains in the same position the relative movements of the detector with respect to the fixation point can be read from the fringe pattern. The central idea is the fact that points along one fringe have the same displacement and although the images only show the absolute value of the displacement plausible assumptions about the sign can usually be made. For this object it seems plausible to assume that the distortion is driven by the mismatch between the CTE of the connection bar and the silicon. Because the carbon fibre connection bar has a smaller CTE than the silicon and the setup is cooling down, the silicon will decrease in length more than the bar, leading to a distortion which should bend the bottom of the detector towards the camera. A movement towards the camera will be referred to as an increase of elevation.

Going vertically down from the fixation point one finds that the elevation is monotonically increasing, as has been motivated above and the gradient, represented by the fringe density, is also increasing. This is typical of a curved distortion with a horizontal tangent in the fixation point. The fixation point tangent is a boundary condition resulting from the clamping of the connection bar.
Looking along horizontal lines in the bottom two thirds of the detector the elevation is always biggest on top of the connection bars and falls off towards the left and right edges. This can be easily found by counting the number of fringes crossed when tracing lines between the fixation point and several points on the horizontal line.
Assuming that the principle behavior on a horizontal line does not change in the upper third of the image were the fringe counting is not possible because the fringes on the left and right halves are not connected one finds that the left and the right part of the detector distort in the same direction and that this distortion is away from the camera. The upper edge of the detector can then be described as an asymmetric V-shape with the point of the V in the position of the large connection bar (the fixation point). There is very little curvature along the arms of the V as keen be seen from the very homogeneous fringe spacing along the upper edge.
Going down the vertical edges of the detector the elevation continuously increases from negative values until the zero displacement fringe is met and then increases further until the bottom edge of the detector is reached. The zero displacement fringe is by definition that going through the fixation point.
The total topology of the distortion is saddle like with only half of the saddle visible. The horizontal plateau of the saddle is in the fixation point.

# image	time/s	# image	time/s
1	0.0	6	12.4
2	2.2	7	14.9
3	4.3	8	17.3
4	7.4	9	20.0
5	9.8	10	22.5

Table 6.1: Times at which the 10 images underlying the fringe patterns of figure 6.5 were taken.

6.1.5 Resume

It could be shown that distortions of the silicon detector are observable with the ESPI method when heating the detector with a projector bulb. Several points were found to be important for the ESPI measurements.

- The detector had to be sprayed with white talcum powder to give it a diffusely reflecting surface.
- Distortions of up to 8 microns (25 fringes) could be resolved in one image.
- The change of distortion must be slow. This was found to be difficult to achieve when bringing the detector out of equilibrium with the ambient by more than a few degrees. Ideally the detector should be in thermal equilibrium with the ambient and the temperature change should be driven by changes in the ambient temperature to guarantee a controlled speed of the temperature changes. This situation is best achieved in an environmental chamber with a transparent window.

- Convection currents which induce a local change of refractive index of the air in front of the object have to be avoided. Convection appears when the object is sufficiently far from thermal equilibrium with the ambient air. This effect can also be avoided when the measurement are done in an environmental chamber.

- The shape of the detector could not be reconstructed from its distortions due to the problems explained above.

- The temperature profiles of the detector are not homogeneous, neither during the heating nor during the cooling period. This is due to inhomogeneities in the heating and to the connection bars which locally add heat capacitance to the system and influence the heat transfer from the detector to the environment. The bars can be clearly made out in the infra-red images even though they are behind the object. In figure 6.2 T0 which is close to the edge falls below T1 which is near the bars soon after the heat is turned off. This is probably due to the heat stored in the bar.

- The connection bars are very clearly visible on the ESPI fringe patterns as the location of the apexes of the fringe triangles.

- The delamination of connection bar 2 is difficult to see in the distortion fringes because the distortion is dominated by the mechanically stiffer bar 1. An indication of the presence of the second bar can be seen by looking at the apex region of the triangular fringes. The apex is broader in the region where the second bar is still in contact with the silicon and it becomes more pointed in the region where the second bar is delaminated.

6.2 FE-Test Object

In this sections measurements and FE[8] calculations of the temperature and distortion of an Aluminium plate shown in figure 6.6 are presented. The test situation was such, that the conditions could be defined with a minimal set of parameters and could easily be implemented in a finite element simulation.

6.2.1 Aim

The aim of these measurements was to demonstrate that a finite element model predicting both temperature and resulting distortion profiles from know quantities of the experimental setup could be verified by a combined ESPI and IR-measurement. The model should be calculated in two stages, using as input parameters, the measured temperature distributions and the assumptions about the heat generation. In the first case the mechanical part of the simulation relating distortions to temperature changes can be checked separately. In the latter of the two cases the measured temperatures can be used to check the simulated distributions. It is advantageous to use a test object which is mechanically very rigid and will show a slow but large movement. This can be achieved by choosing an object with large heat capacity and large coefficient of thermal expansion. Aluminium offers both of these qualities.

6.2.2 Fixation and Heating

Figure 6.6 shows how the Al plate was fixed to the pillar. It was free to rotate around P1 but clockwise rotation beyond the line P1-P2 being vertical was restricted in P2. In this way the plate was kinematically mounted and remained stress free during expansion. P1 is the point in the plate which does not move due to expansions of the plate. The front of the plate was painted matt white so that it would reflect the laser light and have a high emissivity. The plate was heated with a projector bulb from the back. The bulb was enclosed in a 6 cm inner diameter steel tube that insured that the plate was only heated in a circular region of the pipes inner diameter, centered on the back of the plate. The tube also evened out the radiation across this surface by multiply reflecting the light of its walls. The back of the plate was painted matt black so that it would absorb the radiation from the heating bulb.

6.2.3 FE-Model

The model consisted of a homogeneous Al-plate with constraints in points P1 and P2. An additional element which is not shown in figure 6.7 was inserted between the constraint point and P2. This extra element was manually expanded, simulating a tilt of the pillar. Initial simulations did not show the observed s-shape asymmetry of the fringe pattern. This is why a tilt equivalent to the observed pillar bending had to be introduced. On the back of the plate a circular central area corresponding to the illuminated area in the test object was heated homogeneously with constant power. The assumption of constant power should be justified as long as the temperature difference between the lamp and the plate does not change significantly during the heating process. This is true to a good approximation since the temperature difference between the lamp and the plate is certainly larger than 2000 ^{o}C and the maximum change in temperature is of the order of 10 ^{o}C.

[8] FE=Finite Element

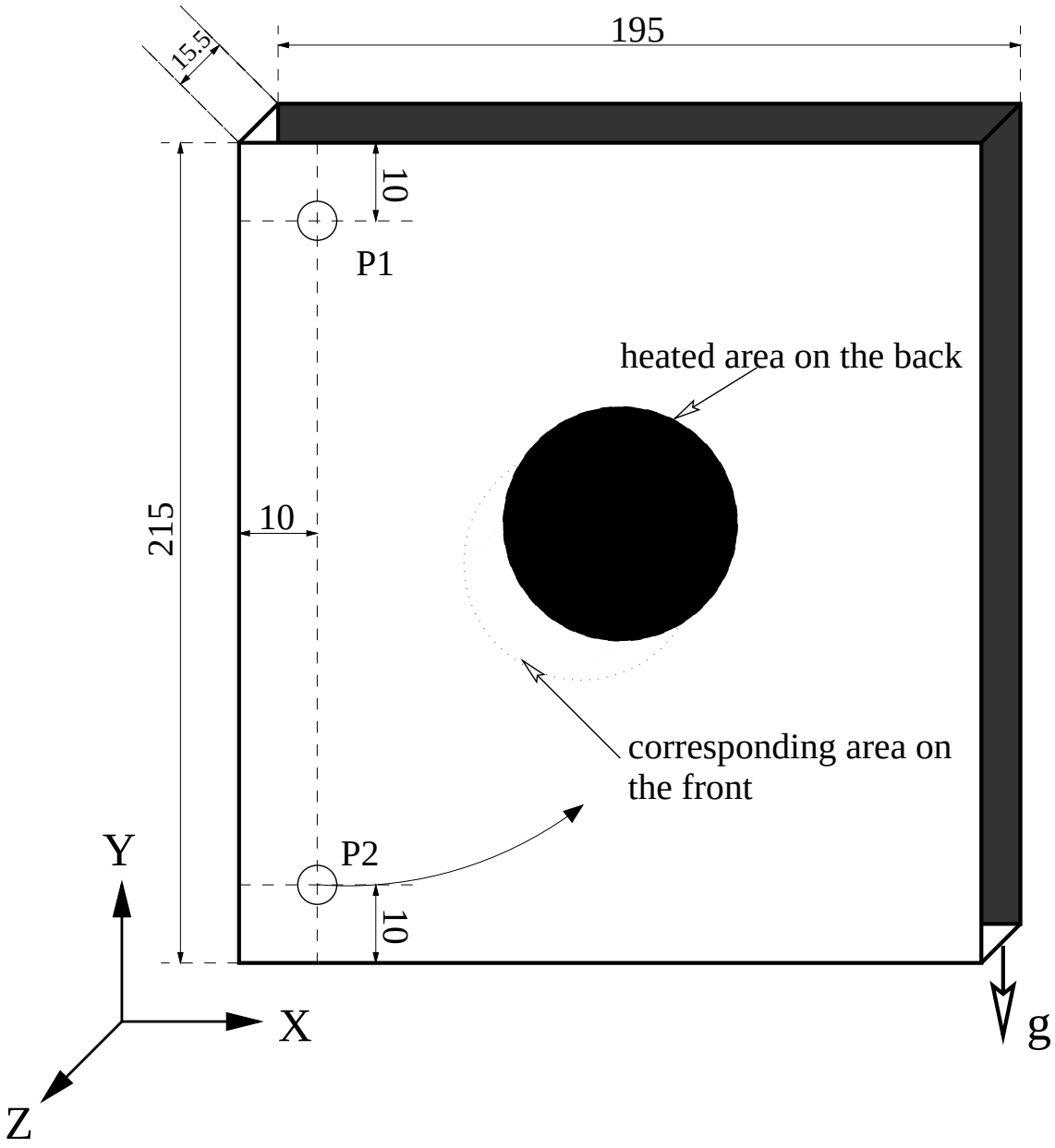


Figure 6.6: Geometry and pivoting arrangement of the Al plate measured to verify and tune finite element simulations. All dimensions are in mm. P1 is a fix point which only allows rotation around P1 in the x-y plane and P2 is a constraint which blocks clockwise rotation of the x-y plane. P1 and P2 were attached to a pillar on the left of the plate. In the center of the plate the area heated by the projector bulb is indicated.

The heating power per unit area and the pillar distortions were free parameters of the simulation and were tuned according to the temperature measurements and the measured pillar bend. Figure 6.7 shows the horizontal in plane distortions of the aluminium plate during the heating process obtained with the FE model. Figure 6.7 describes the distortion after 20 seconds of heating and is supposed to be equivalent to the second fringe pattern in figure 6.8. It has to be noticed that the fringe separation in the distortion simulations is not the same as that of the measurement. The distortion between each grey level is 0.11 μm and the units of the scale in the top right corner of figure 6.7 are 0.72 cm.

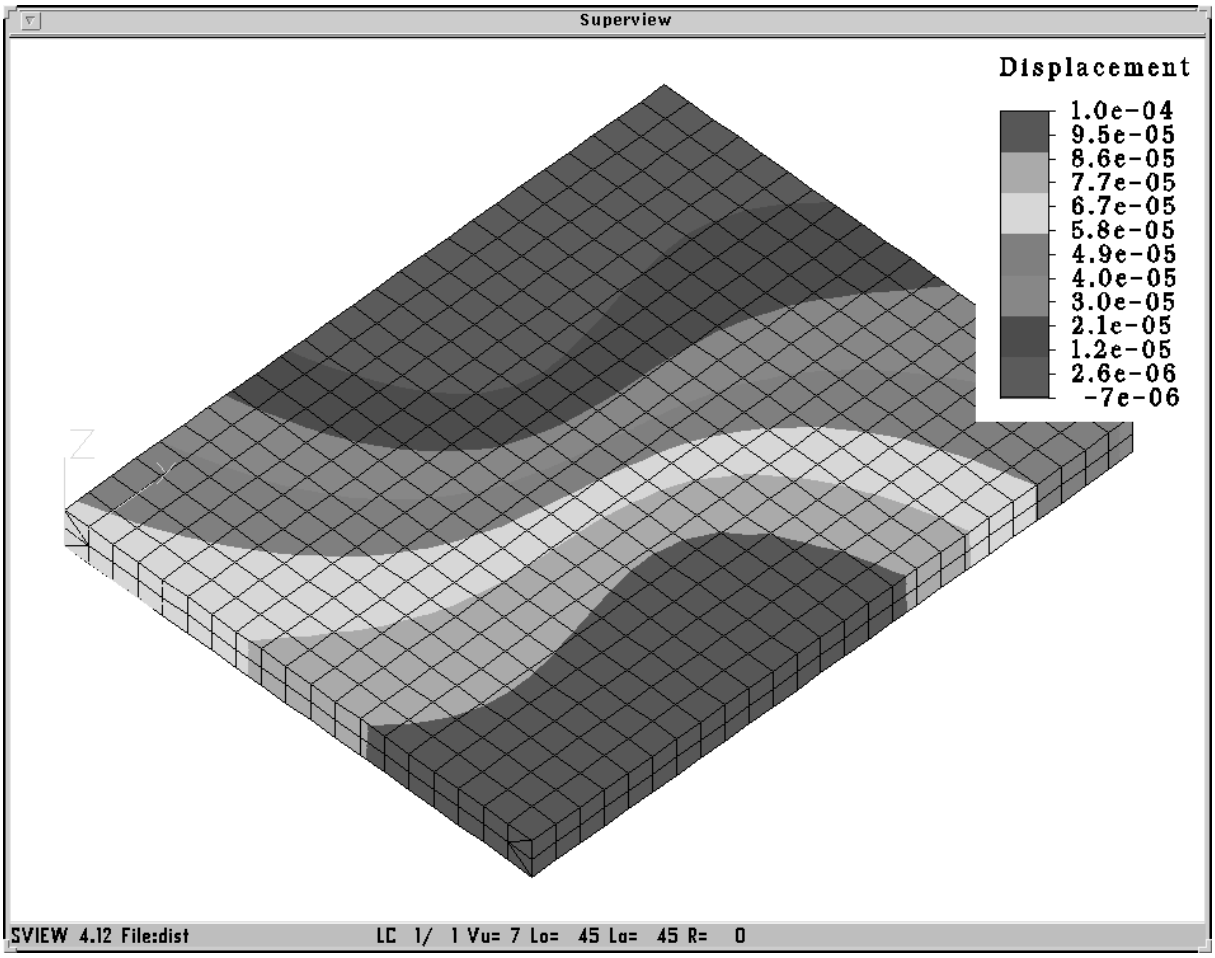


Figure 6.7: Finite element simulation of the distortion of the aluminium plate 20 seconds after the start of heating. The scale in the top right corner has units of 0.72 cm. The spacing between two grey levels is 0.11 μm. The image should be equivalent to the second image in figure 6.8. Note the different amount of distortion between fringes in the two images.

6.2.4 Measurements

The plate has been observed during the cooling and heating process. This was possible even though the heater bulb produced a lot of stray light because the more powerful Ar^+ Ion laser was used and the stray light could be tolerated. The distortion was measured in the plane of the object because this is least sensitive to effects of the heat propagating from the back to the front of the plate which may be more complicated to describe correctly in the simulations. The interferometer was set up[9] to be sensitive to distortions parallel to the horizontal axis[10]. Initially the temperatures were measured coarsely with the IR-scanner to check whether the temperature profiles on the front of the plate were symmetric around the horizontal center line. When this was found to be true only the top half of the plate was measured to enhance the resolution of the image and to reduce the distance between the object and the scanner.

[9] see figure 5.4 page 48

[10] x-axis in figure 6.6

ESPI Results

Figure 6.8 and 6.9 show a series of fringe patterns of the distortions in the horizontal direction of the plate during and after heating.

The reference in figure 6.8 is with respect to the plate before the heat was switched on, in

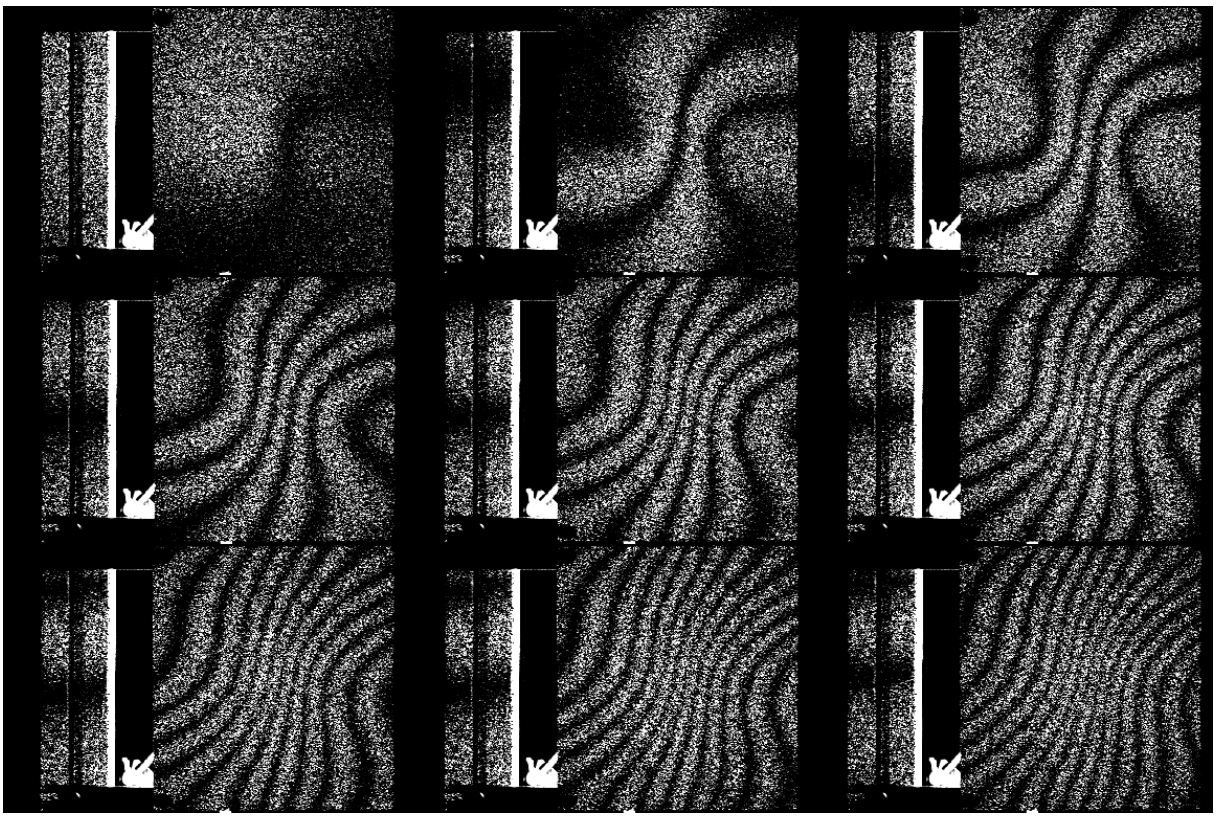


Figure 6.8: Horizontal distortion of the Al plate during heating with respect to the plate before heating. The time between the images is 10 seconds. The first image was taken 10 seconds after the heat was switched on. The distortion between two fringes is 0.26 μm. The projector bulb was operated at 46 W.

figure 6.9 it is with respect to the plate with maximum temperature. The projector bulb was run at 46 W for the heating process. The pillar onto which the plate was mounted is also shown. Although the temperature profiles were symmetric the fringe pattern are not symmetric and show a distinct s-shape. Initial simulations of the problem also showed that the fringe pattern should be symmetric. This effect is explained by looking at the fringes on the pillar in figure 6.8. The pillar is also illuminated on one side and in the last image shows three distinct fringes, corresponding to a distortion of 0.8 μm. If this distortion is included in the simulation the fringe pattern can be reproduced with a finite element simulation. The change in temperature of the plate is very symmetric and is shown on the IR-images in figure 6.11. The observed distortions during heating and cooling can be explained with a solid body rotation superimposed onto a fringe pattern from the underlying temperature changes. In order to make sure that the measured distortion is correctly explained by the tilt of the pillar, a second measurement was done, shielding the pillar from the light. The result is shown in figure 6.10. The fringe pattern in figure 6.10 is more symmetric and more similar to the initial simulation results. There is still a tilt of approximately 1.5 horizontal fringes in the image which is due to conduction of heat from the plate through the beams that hold it to the pillar. In the case of an ideal setup (no distortion in the pillar) the two points P1 and P2 would be on the same fringe, but

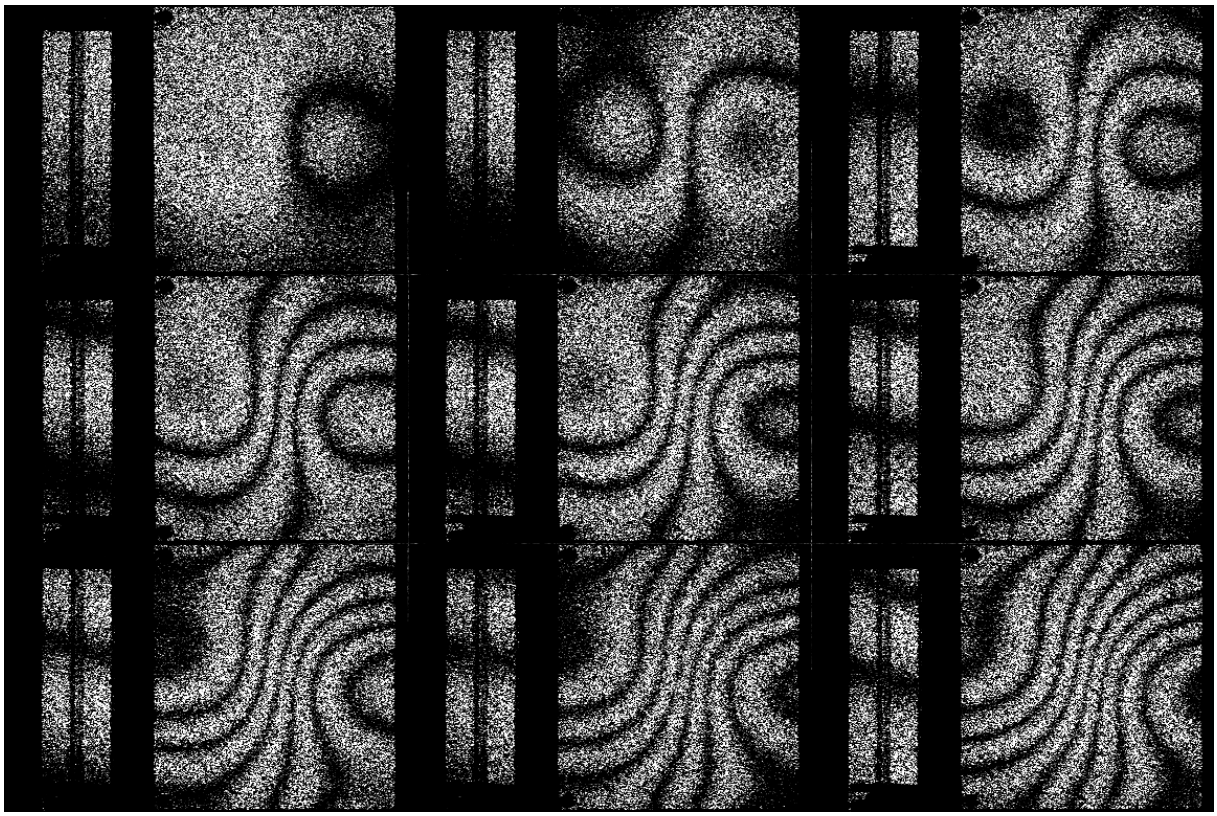


Figure 6.9: Horizontal distortion of the Al plate during cool-down with respect to the plate at maximum temperature. The time between the images is 10 seconds. The first image was taken 10 seconds after the heat was switched off. The distortion between two fringes is 0.26 μm. This is the continuation of the fringe patterns shown in figure 6.8.

in figure 6.10 they are separated by approximately 1.5 fringes.

IR-Results

Figure 6.11 shows the temperature differences of the Al plate during the heating period with respect to the homogeneous temperature before heating. Only the top half of the plate was measured and the distributions have circular symmetry to a very good approximation. The profile across the center of the distribution is shown in figure 6.12. The image has been treated with a 3×3-rank mean filter[11] to reduce noise. The small change in the temperature of the pillar that caused the s-shape asymmetry of the observed fringe patterns is visible in the last three images of figure 6.11. It less than 0.5 ^{o}C in total.

6.3 Phase Stepping

The technique of phase stepping one of the interferometer beams in order to determine the signed phase shifts resulting from the distortions of the object under study as explained in chapter 5 has been implemented for the two interferometer setups described in chaptersetupESPI.

[11] see chapte 6.3

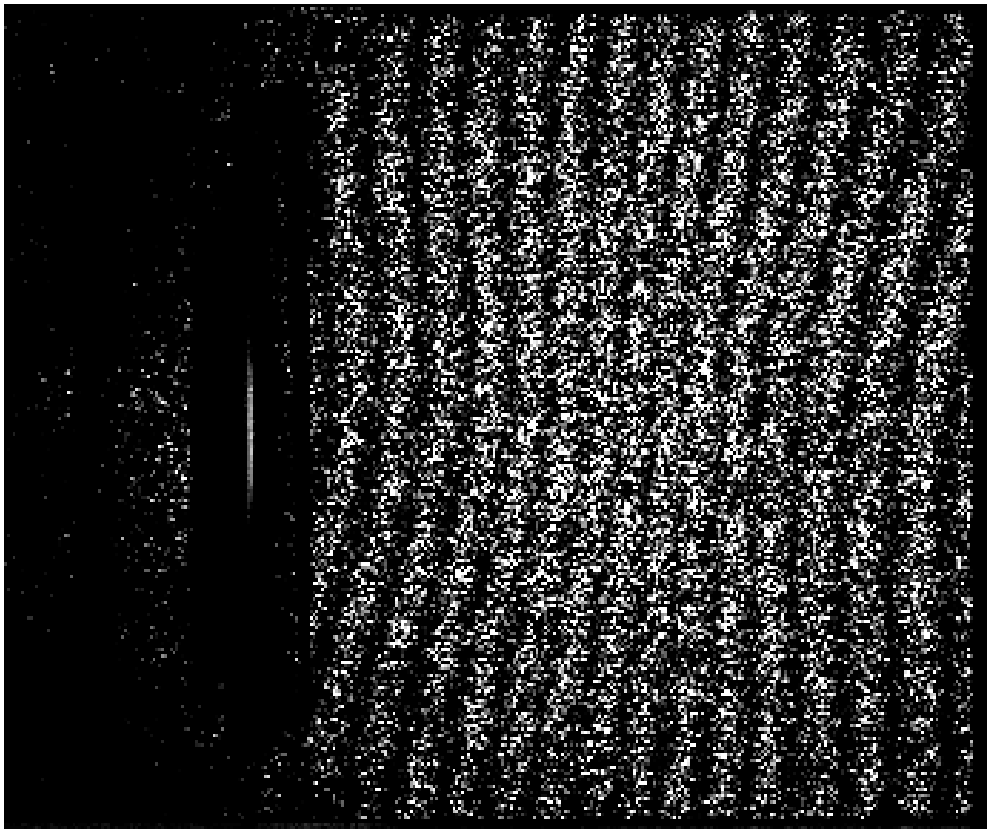


Figure 6.10: Horizontal distortion of the Al plate during heating with respect to the plate before heating. The pillar has been shaded from the light of the projector bulb to avoid external tilts of the plate. There is still approximately a tilt of 1.5 fringes (0.4 μm). The distortion between two fringes is 0.26 μm.

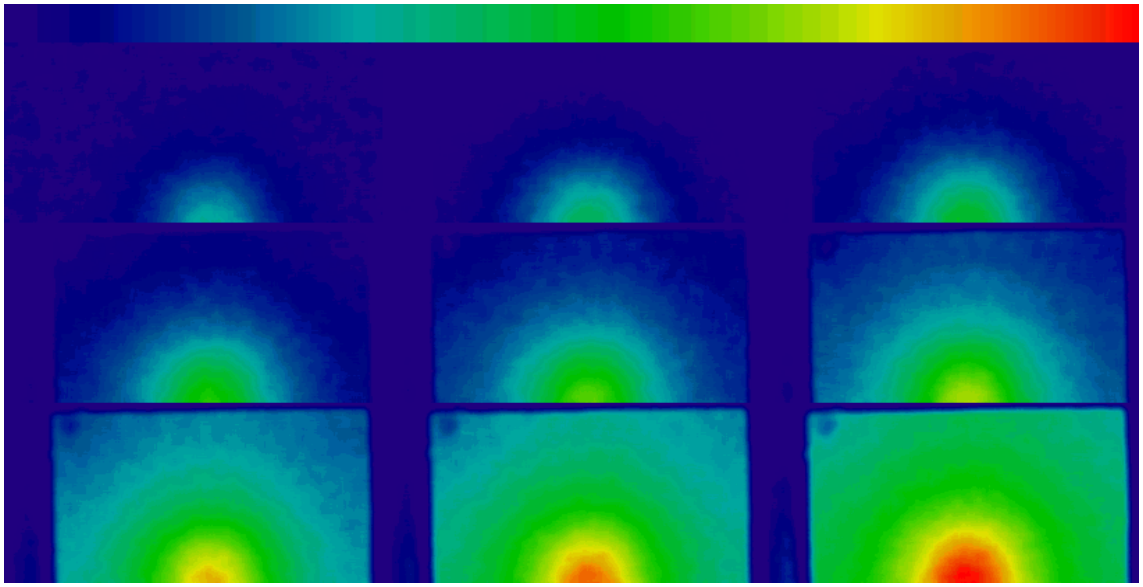


Figure 6.11: Temperature differences of the plate during heating with respect to the plate before heating. The figure shows 3 rows of 3 images each. Only the top half of the plate is shown in each image. The time between the images is 10 seconds and the first image was taken 20 seconds after the plate was switched on. The power of the heater bulb was 46 W. The colour scale on top of the image is 0 to 3.3 ^oC.

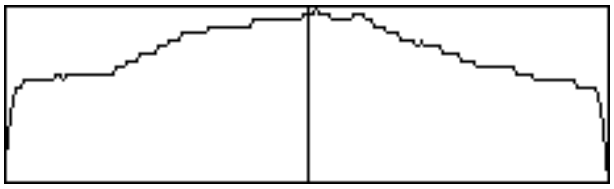


Figure 6.12: Horizontal temperature profile across the center of the Al plate from the last image of figure 6.11. The vertical scale is 0 ^oC to 3.3 ^oC.

6.3.1 Measurements

Figure 6.13 shows the carbon fibre base plate described in chapter 4, section 4.2 located inside the environmental chamber as seen from the ESPI camera (speckle pattern). The base plate was choosen to demonstrate the phase stepping method because it could easily be rotated through a very small angle. Figure 6.14 shows the horizontal in-plane distortion

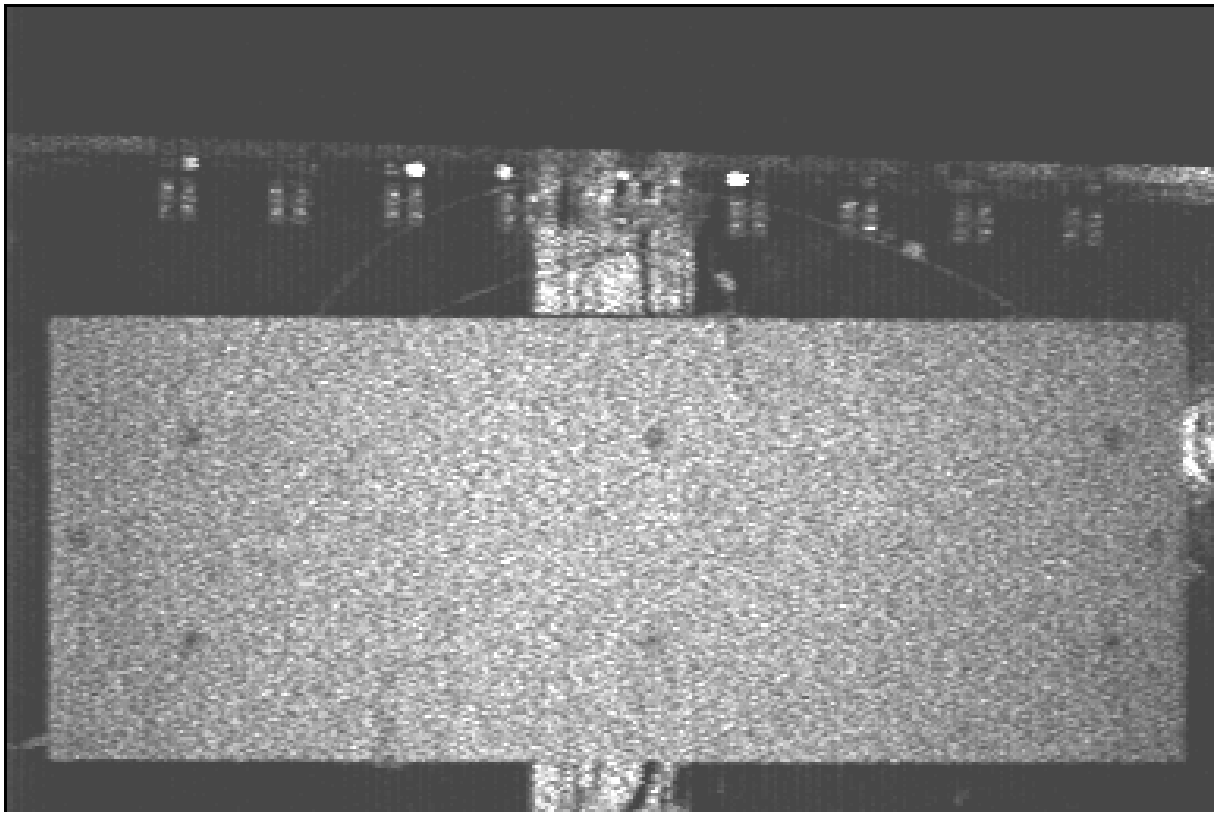


Figure 6.13: Speckle pattern of the baseplate in the environmental chamber as seen by the ESPI camera. Behind the plate, in the center of the image the supporting pillar can be seen through the transparent backside of the environmental chamber. On the top edge of the chamber the feedthrough connectors for the silicon power supply lines can be seen. Since the chamber is pivoted mechanically indepenent from the plate it can rotate around the plate center. This explains the small clockwise rotation of the chamber with respect to the plate. The six fixation points for the local support can also be seen on the plate. The bottom and side walls of the chamber are not visible.

fringes produced by small rotations of the base plate around an axis through the center of the plate, perpendicular to the imaged surface. The plate has an area of 1311 cm^2 and the total area viewed in figure 6.14 is more than 2800 cm^2. The single images underlying the sequence of fringe pattern is described in table 6.2. Between consecutive images the

plate was rotated through a very small angle by tapping one side with a pencil. The zero displacement fringe will always go through the rotation axis and for small angles the fringes will be horizontal and the displacement will be proportional to the vertical distance from the horizontal symmetry line. This explains why the fringe pattern are horizontal lines. A clockwise rotation can be distinguished from an anti clockwise rotation by knowing that the phase increases when a point moves towards the left[12] and that increasing brightness in the image corresponds to increasing phase shift. This way it can be seen that in the first fringe pattern of figure 6.14 points below the center line show increasing phase (brightness) and are therefore moving to the left, identifying the rotation direction to be clockwise.

fringe pattern	image A	image B	rotation between A and B
1	1	2	4cw=4cw
2	1	3	4cw + 11acw=7acw
3	1	4	4cw + 11acw + 5acw=13acw
4	2	3	11acw=11acw
5	2	4	11acw + 5acw=16acw
6	3	4	5acw=5acw

Table 6.2: Image pairs composing the fringe patterns of figure 6.14 and the rotations introduced between the images. The rotation is given in number of fringes where cw denotes clockwise and acw anti clockwise.

6.3.2 Noise Reduction

The images shown in figure 6.14 have not been smoothed in any form. Also pixels which did not have a good modulation[13] have not been excluded from the image as is usually done. This was done to show the true amount of noise inherent to the images. The noise increases towards the left and right of the fringe patterns because there the matching of the two illumination beam intensities is worst and thus more and more pixels will have bad modulation. The noise is most visible if a pixel represents a phase shift near to $+\pi$(white) or $-\pi$(black) and has been slightly mismeasured. In this case it may go across the boundary and turn from white to black which appears to be most visible. An algorithm to reduce the noise due to pixels crossing the phase boundary is simple. It consists of looking for pixels which only have neighbours[14] of opposite sign and relatively large absolute value and changing their value to the average value in the neighborhood. This filter is often called a rank-mean filter. Even though the fringe patterns of figure 6.14 show very little noise implementing the filter was simple enough[15] to at least test it. Figure 6.15 shows the result of rank filtering the fringe patterns of figure 6.14 with a 3x3 filter[16] in three iterations. Figure 6.16 shows a vertical profile through the phase map of the first fringe pattern of figure 6.15.

[12] This is determined by the direction of the piezo movement and which side the phase shifted beam is illuminating from

[13] see chapter 5.2.3

[14] The size of the neighborhood to be considered is the free parameter of the algorithm

[15] This filter can be implemented as a 10 line program inside SEMPER

[16] using 3x3 pixel areas

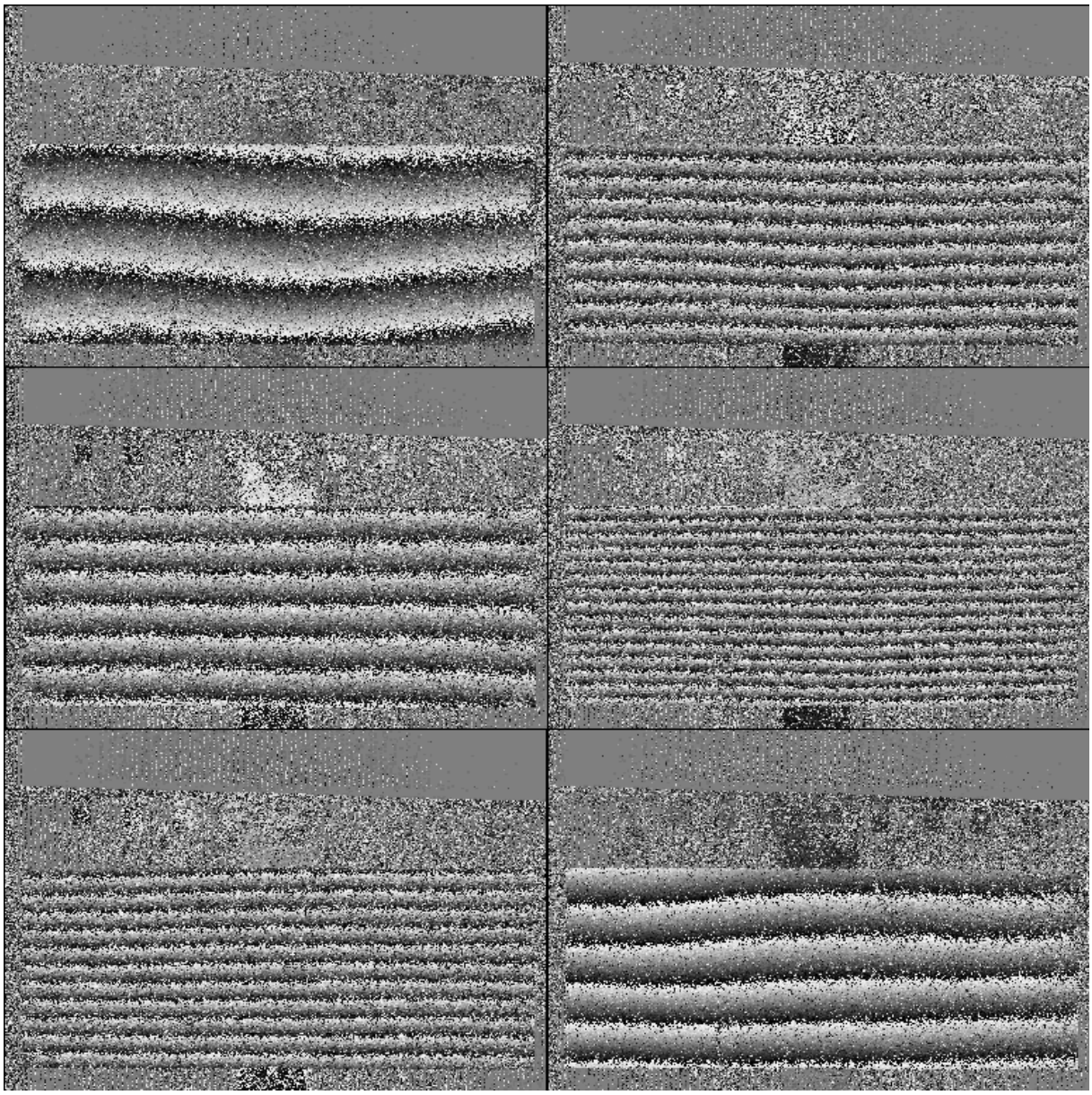


Figure 6.14: Fringe pattern showing optical phase shift obtained with the Carré method described in chapter 5. The fringe patterns are numbered by column. top left=1, middle left+2, bottom left=3, top right=4, mid right=5, bottom right=6. The underlying images and distortions are given in table 6.2. The object in view is the carbon fibre base plate inside the environmental chamber. Above and below the plate the pillar holding the chamber is visible. Inside the top edge of the chamber the feedthroughs for the silicon detector heating supplies (see figure 4.14, page 33) can be seen

6.3.3 Resolution Limit

The ratio of the rms fluctuations of the phases around a straight line fit to the dynamic range of the phases is approximately 1/20. The dynamic range[17] corresponds to a displacement of 257 nm. Assuming that this range can be resolved in 20 steps the displacement accuracy was 13 nm over a total area of 1311 cm^2. This corresponds to an angular resolution of the plates rotation of 113 nrad.

[17]The displacement between two fringes

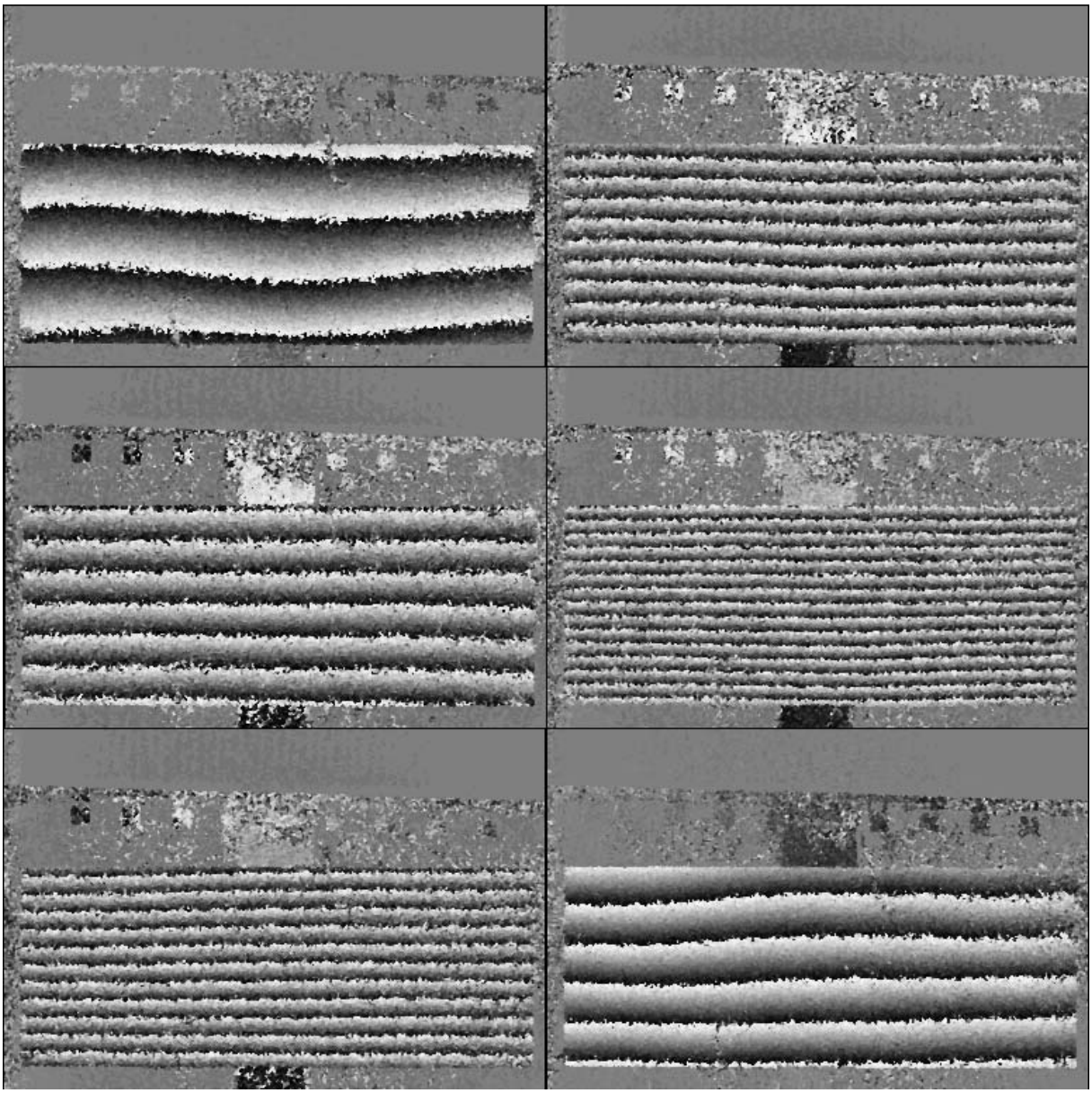


Figure 6.15: The fringe patterns of figure 6.14 after 3 iterations of a 3x3 rank mean filter.

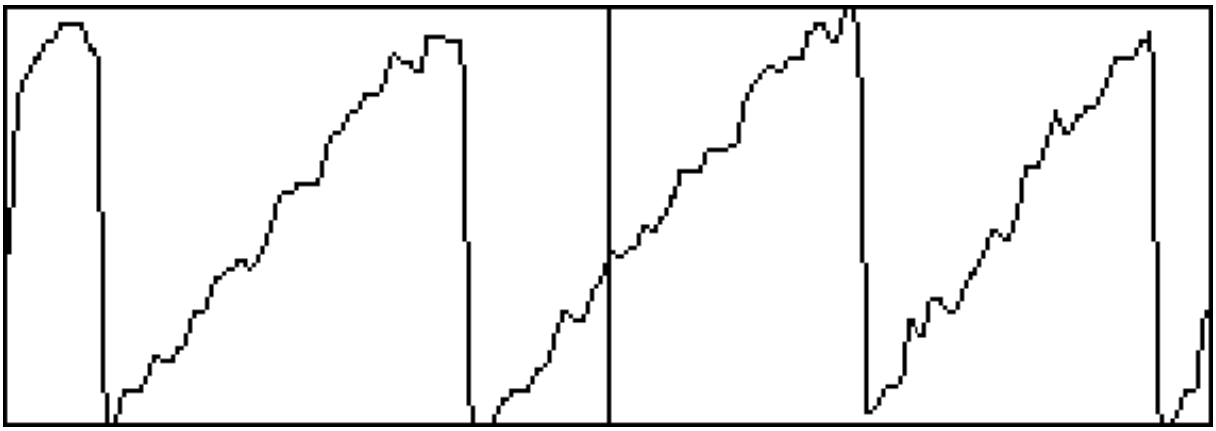


Figure 6.16: Vertical profile through the center of the phase map of the first fringe pattern of figure 6.15.

The biggest total distortion observed in one image was 15 fringes but these were by far not at the resolution limit. It is assumed that 30 fringes will be resolvable. This leaves 9 image pixels per fringe. Since the cameras effectively only had 287 pixels per column the separation between two fringes would be 5 detection elements. This is very close to the theoretically achievable minimum of 3 pixels per fringe. Looking at the feedthrough connectors in the top of the box it can be seen that the box itself is being distorted by less than one fringe in most of the fringe patterns.

The resolution of these measurements could be improved by using a faster frame store so that the time between images of one phase stepping sequence is reduced and possible phase drifts from the fibers are eliminated. The ratio of the size of one speckle and the size of one detection element, here two CCD pixels, was not 1. The speckle area in these images[18] was approximately 80 μm^2 whereas the detection element area was 252 μm^2. Thus one detection element was integrating over approximately 3 speckles and thus reducing the modulation of a pixel in the fringe pattern to 1/3 of their theoretically achievable maximum.

[18] numerical aperture 8

6.4 Silicon Tracker Modules

The detector modules and support structure described in chapter 4 have been tested in several steps. Initially the carbon fibre base plate has been observed without the local support structure. This was followed by tests of the local support structure on the plate. The modules were then mounted onto the support structure, initially without their electronic boards and finally in the complete setup. As an intermediate step the temperatures distributions of the electronic boards were measured separately before they were mounted to the modules.

6.4.1 Base Plate

The base plate was tested for distortions due to temperature changes. During the tests the base plate was mounted inside the environmental chamber on a circular steel finger protruding from the pillar through the back wall of the chamber. The finger was attached to the back of the plate in the center. An image as seen by the ESPI camera (speckle pattern) of the base plate in the chamber has been shown before in section 6.3, figure 6.13. Because the plate itself has neither a cooling system nor any heat generating components it was heated by an array of 6 projector bulbs positioned in front of the plate. In order to suppress the convection currents from the big heat load of 1.2 kW emitted from the array the lid of the environmental chamber had to be closed. Because of the very low coefficient of expansion of the plate large temperature changes had to be introduced to see any distortions. Due to the remaining convection currents inside the chamber coherence between images could not be maintained for a long time and the resulting fringe patterns are not very sharp. Figure 6.17 shows the horizontal in-plane distortions of the carbon fibre plate from the cooldown after a 10 minute heating phase at 600 W power. The images for this fringe pattern were taken 20 and 80 seconds after the heat had been switched off and the array had been moved out of the line of sight. The deflection between two fringes is 0.54 micron. To estimate the coefficient of linear expansion of the plate from these images the temperature was observed with the infra-red scanner and an average value of the temperature of the plate was calculated for each image. In the two images resulting in figure 6.17 the average temperature difference was 7 ± 1 K. The total distortion was 2.5 ± 0.5 μm and the length of the plate was 586 mm. This leads to an estimate for the CTE of the plate of 0.6 ± 0.2 ppm/K. There is an uncertainty on this number due to the out of plane distortion of the plate which was not measured and could have effected the in-plane measurement. This is estimated as a function of the maximum out of plane distortion by assuming the distortion shape to be circular. Note that the out of plane direction is defined to be along the z axis and that the long axis of the plate lies on the x-axis. In this case it it can be said that an out of plane distortion with relative magnitude $\Delta z/L_z$ will generate a relative in-plane distortion of order $\mathcal{O}(\Delta z/L_z)^2$. Even though the linear expansions of one of the skins of the plate, Δx_{skin}, necessary to generate an out of plane distortion of magnitude Δz is only $\Delta x_{skin} = L_z/L_x \cdot \Delta z = 1/58 \cdot \Delta z$, where L_z is the thickness and L_x the length of the plate, this does not compensate for the very small factor of $(\Delta z/L_z)^2$.

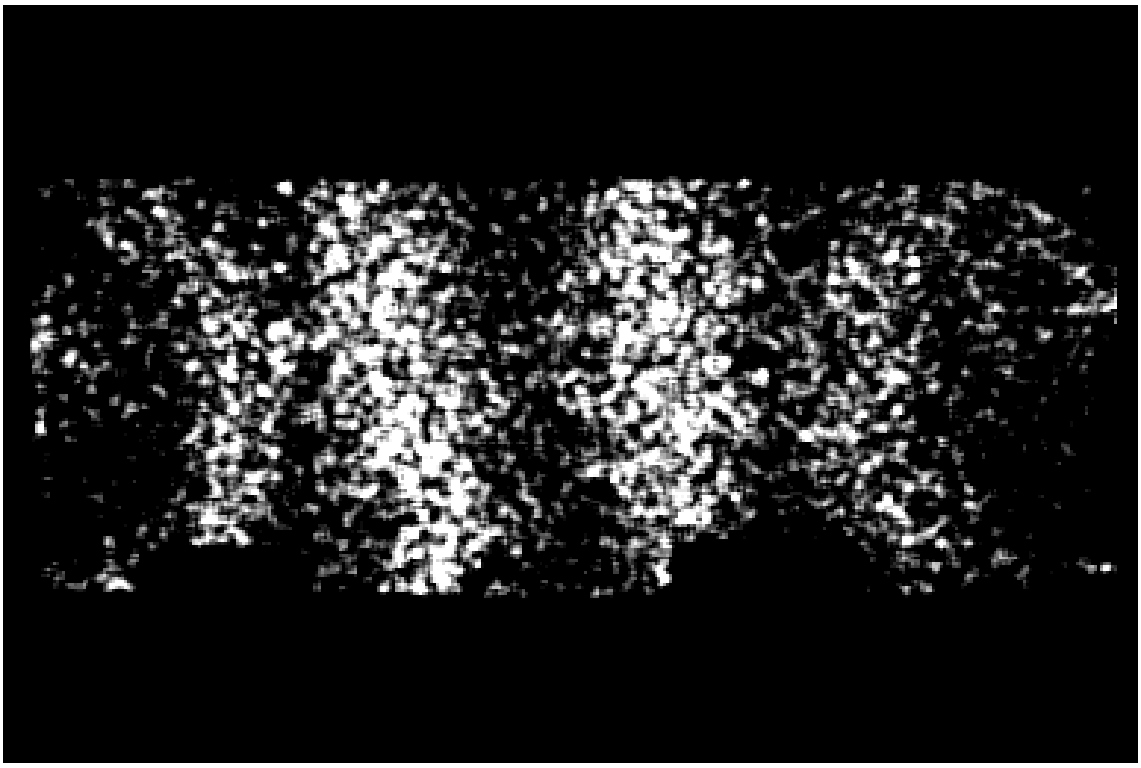


Figure 6.17: Horizontal in-plane distortion fringes of the carbon fibre plate during cool-down. The deflection between two fringes is 0.54 μm. The object can be seen as a speckle pattern in figure 6.13

6.4.2 Local Support

The main purposes of these measurements were:

- to test the stability of the local support and its mountings
- to evaluate the influence of the cooling system on the movements of the plate

The local support structure has been mounted on the carbon fibre base plate by fixing the right end with a dowel pin on the symmetry line and allowing the left end to slide in the horizontal direction by guiding a dowel pin in a polished slot. The long sides of the local support (top and bottom) have been pressed by three springs on each side onto the base plate, allowing them to move in the plane of the base plate in any direction, but keeping them always at the same elevation above the base plate. The cooling pipes were also free to move with respect to the local support on the left end and were fixed to the local support on the right end. Figure 6.18 shows an image (speckle pattern) of the left 75% of the local support inside the environmental chamber as viewed by the ESPI camera. Figure 6.19 shows the the horizontal in-plane distortions of the object shown in figure 6.18 after cooling it down from room temperature to equilibrium, with the cooling system running at a feed temperature of approximately 6 degrees. It can be seen, that the local support moved by one complete fringe (0.53 μm) to the left[19] with respect to the base plate. The contrast between the local support and the baseplate is very good. The pipes which are expanding more rapidly than the local support, due to their larger CTE and the direct contact with the coolant, go through several fringes. These can not easily be resolved because the pipes are only between 1 and 3 image pixels wide. However the

[19]The phase sign convention is opposite to that for the rotation fringes of figure 6.15

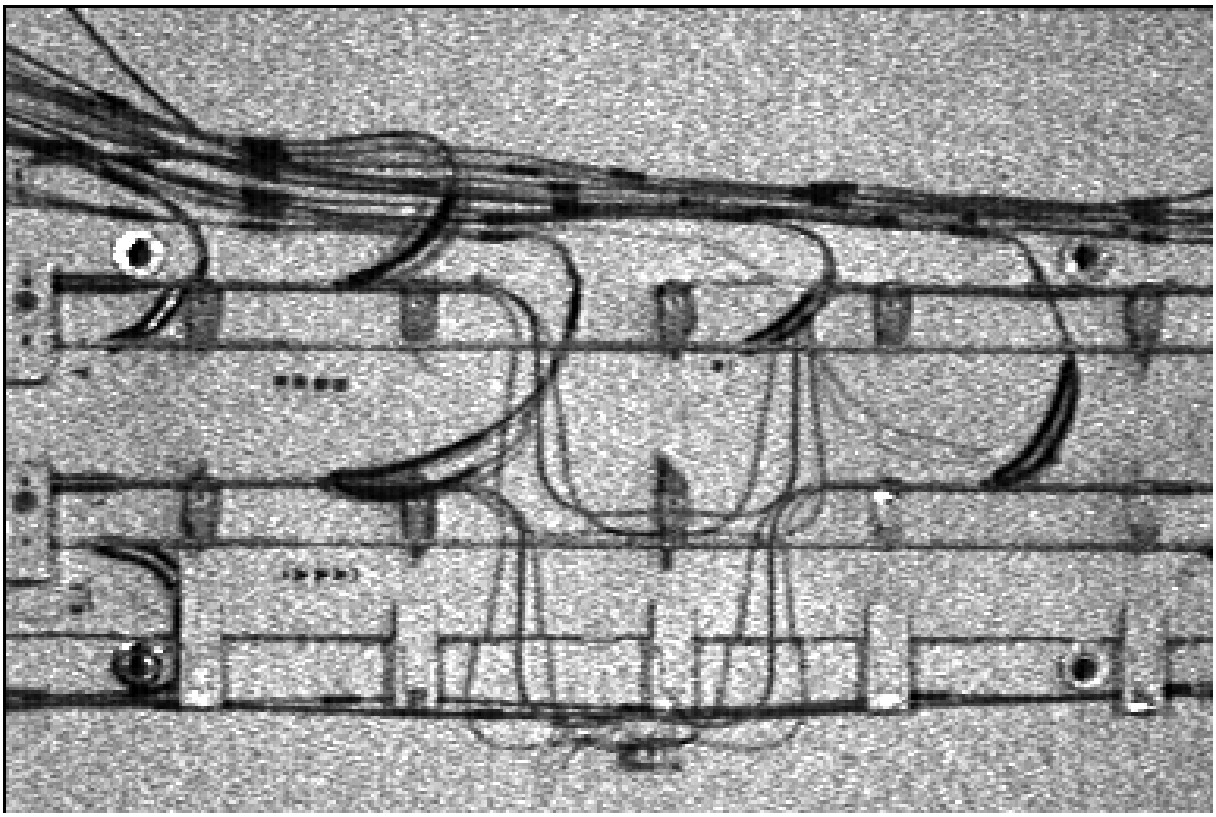


Figure 6.18: Speckle pattern of the left 75% of the local support as seen by the ESPI camera. Fringe pattern derived from distortions of this object are shown in figure 6.19. The two rows of cooling pipes can be seen with several sensors connected to them. Near the top of the image the wires from the sensors run towards their readout connections (left outside image). The fixation blocks for the pipes can also be seen on the left edge of the image. The 4 bright disks with the black cores, two on top and two on the bottom of the local support are the connection clamps holding the local support to the base plate. The environmental chamber itself is not visible. Note that there are allways two shadows for each object. This is particularly visible for the bend sections of the electronics cooling pipe.

difference between the pipe expansion and the local support expansion can be seen towards the left of the image where the pipes do not follow the long fringe of the local support. The position of some of the cables can also be seen, but do not carry phase information because the pixels on the cables have been excluded from the phase map calculation since they had very low phase modulation[20]. Figure 6.19 shows that the concept of the sliding local support and pipes works well. It also shows that some heat is taken out of the local support by the cooling system and the local support therefore distorts by minute amounts. The distortion visible on the base plate is thought to be induced as a result of the stress building up in the hoses connecting the local support to the walls of the box. Although the hose feed-throughs are sliding they can not slide without some friction. This can introduce an out of plane tilt of the base plate which will result in in a few vertical fringes across the whole of the base plate. The dark region extending from the bottom edge, slightly right of center is attributed to a distortion of the mylar window and is present in all fringe patterns that correspond to later times during the cool-down.

[20]The cables were not properly fixed and therefore were not stable during the 4 images of one phase step set.

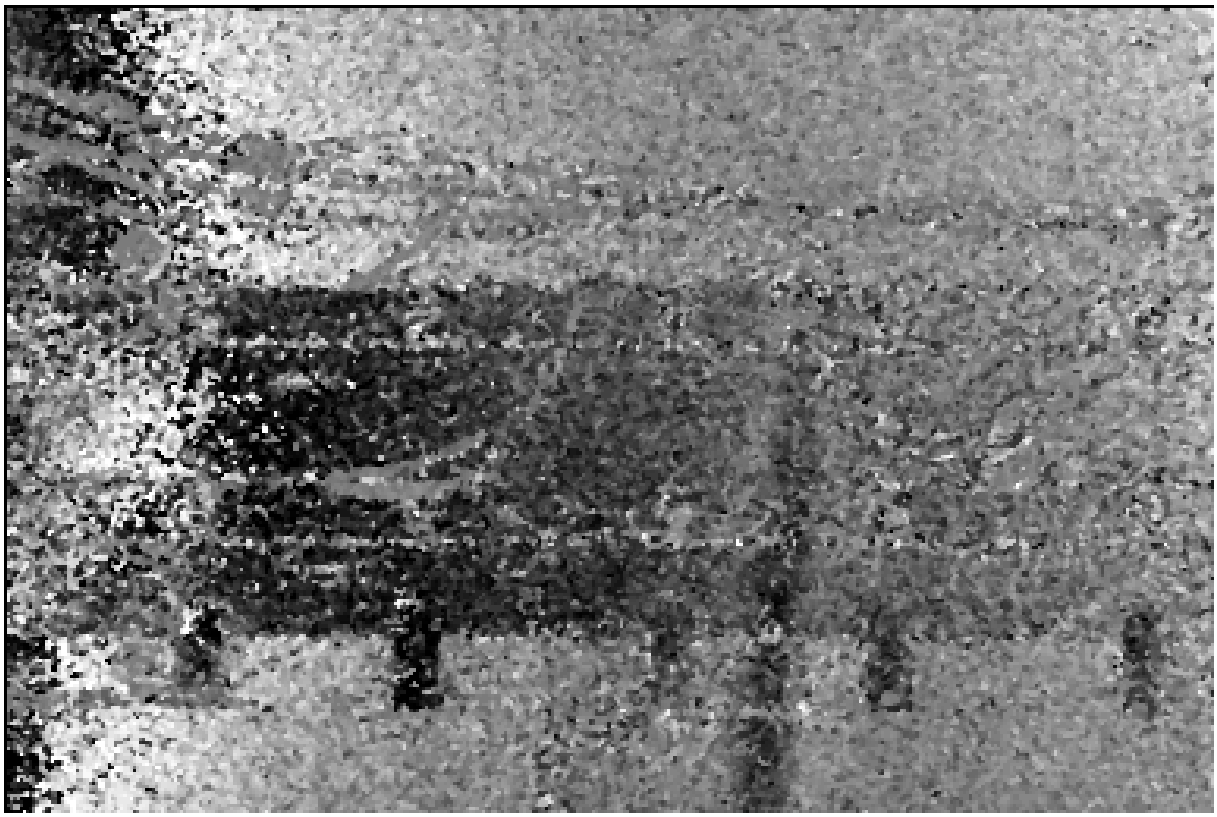


Figure 6.19: Horizontal in-plane distortion fringes on the local support and base plate after cooling down. The object is shown in figure 6.18. The fringe pattern was obtained with a Carré phase stepping method. One fringe corresponds to 0.54 μm displacement.

6.4.3 Boards

The main purpose of these measurements was to ensure that the boards were generating heat homogeneously and that they would withstand a possible failure of the cooling system after they had been built into the modules. The boards operating voltages were ramped in 0.5 V steps from 2.0 to 4.5 V. This corresponded to an increase in power from the nominal value of 0.93 W [21] to 4.7 W per board. The boards as shown in figure 6.20 were hanging vertically from their 50 pole ribbon cables. At this point the emissivity of the boards had not been homogenised by spraying them with talcum powder because some soldering still needed to be done after they had been built into the modules. Therefore the calibration procedure described in chapter 5 could not be applied. The boards were imaged from the back and the front in two separate runs and IR-images were taken when the boards were in equilibrium. In parallel to this the temperatures measured by the sensors on the boards were logged. Two boards were always measured in parallel. One made of standard glass fibre, the other of a low CTE kevlar paper material. The kevlar paper has a higher thermal conductivity and was expected to result in lower peak temperatures of the board and the chips because the board can contribute better to the cooling of the chips. This effect was found to be small. At the maximum power of 4.7 W per board the kevlar boards chips were on average 5 ^oC colder than those on the glass boards, even though they were more than 75 ^oC above ambient temperature. It was found that on the glass fibre board one chip was systematically hotter than than expected, resulting in the complete row of

[21] This assumes a total power of 1.8 mW per channel and the board would serve 500 channels

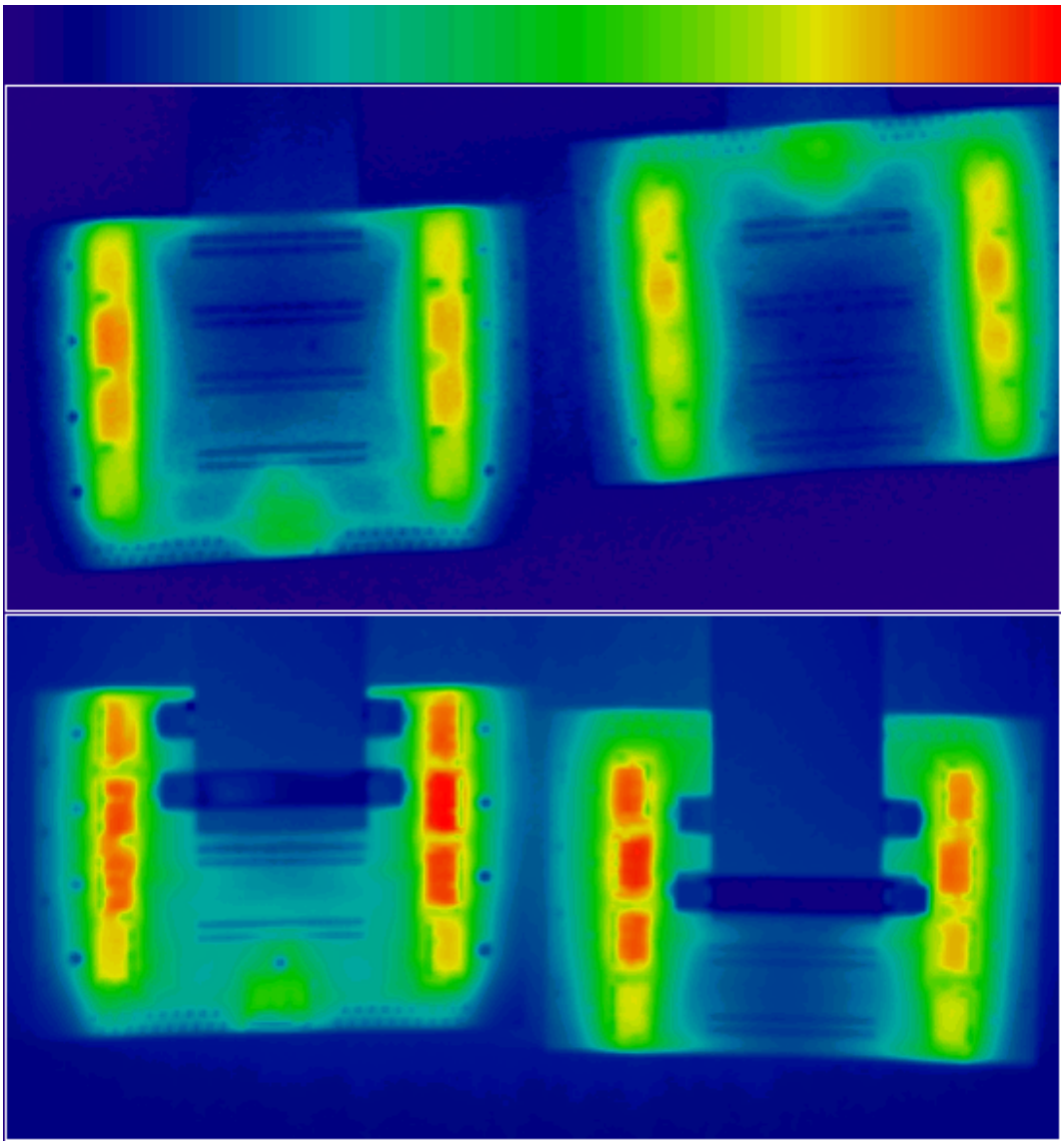


Figure 6.20: IR-images of two boards at nominal operating voltage viewed from the back (top half) and from the front (bottom half). The colour scale is in scanner ADC values and not calibrated. The scale is approximately 21 to 41 ^{o}C.

chips to heat up more than all other rows. This effect was seen from the front as well as from the back of the boards. Optical inspection of the chip in question showed no obvious damage.
It could also be seen that the boards are hotter on the top edge than on the bottom edge which is believed to be due to the upward going convection currents cooling the bottom edge more efficiently. The boards also showed no difference in their properties after they had been operated at the maximum power and had partially heated themselves up to 110 ^{o}C. Figure 6.20 shows the IR-images of the boards at the nominal operating voltage of 2.0 V from the front and from the back. The lines in the board connecting the ROC chip[22] to the power supply can be clearly made out as horizontal thin hot structures near the top (bottom) edge of the boards due to the power dissipated in them. This is an indication that the lines have been dimensioned too small.

[22]The single chip on the long edge of the board

6.4.4 Modules without Boards

There are three central points in this subsection. The first one is to ensure that the heating of the silicon detectors produces homogeneous temperature distributions by adjusting the voltages on the individual crystals.
The second point is to check the functioning of the detector part of the cooling system and see how homogeneously the detectors are cooled.
The third point is to observe the deflections of the detectors due to the silicon heating and cooling alone, without the influence of the boards which on the one hand will stiffen the modules and on the other hand represent an additional source of heat and distortion. It is also a test of the maximum observable deflection velocity because the low heat capacitance of the naked detectors will make them react very quickly to changes in the heat load.
The aim of the measurements with the fully adjusted[23] detectors was to simulate situations which could occur during the operation of the detectors inside ATLAS. These involve variations in the power dissipation of individual crystals, global power variations, changes in the cooling system parameters as well as power and cooling system cycles[24]. For all the measurements the IR-scanner has been synchronised with the ESPI-camera by triggering the camera from the ESPI PC. Out of the large sample of measurements two are described here in more detail.

Single Detector Switching

All detectors but detector three[25] have been without power and the cooling system was working with a coolant temperature of 14.1 ^oC. Detector 3 was operated at its nominal power of 0.4 W and then also switched off. The resulting out of plane distortion was observed together with the temperature changes.
The time to get into equilibrium after detector 3 was switched off was 30 s. As figure 6.21 shows, the discrete sensor on detector three[26] located at the edge of the detector farthest away from the cooling pipe, registered a temperature change of 0.74 ^oC. The other discrete sensors showed no change due to the switching of detector 3. The maximum variation of the discrete sensors on the detectors due to the heat flow from the ambient through the detectors into the cooling pipe was 1.95 ^oC which has to be compared to the temperature difference between the ambient air in the environmental chamber and the cooling pipe surfaces which varied between 8.5 and 9.5 ^oC. The IR-images and ESPI measurements for this situation are shown in figures 6.24 and 6.23. An image (speckle pattern) of the detectors in the environmental chamber as seen by the ESPI camera is shown in figure 6.22.
The image shows the row of silicon detectors below the horizontal center line and the power supply lines for the detectors. Above the detectors the free row with its cooling pipes and the cables for some of the sensors on the cooling system can be seen. The bottom edge of the local support is not visible because it is overhung by the silicon detectors. The temperature distributions show that detector 3 heats up most where its coupling to the cooling pipe is worst. The ESPI measurements show a small out of plane distortion which does not only affect detector 3, but also the neighbouring detector of the other detector

[23] voltage on detectors
[24] switching the power and cooling on and off
[25] third from the left, usually under the board.
[26] sensor number 50 in figure 4.10 and number 73 in figure 6.21

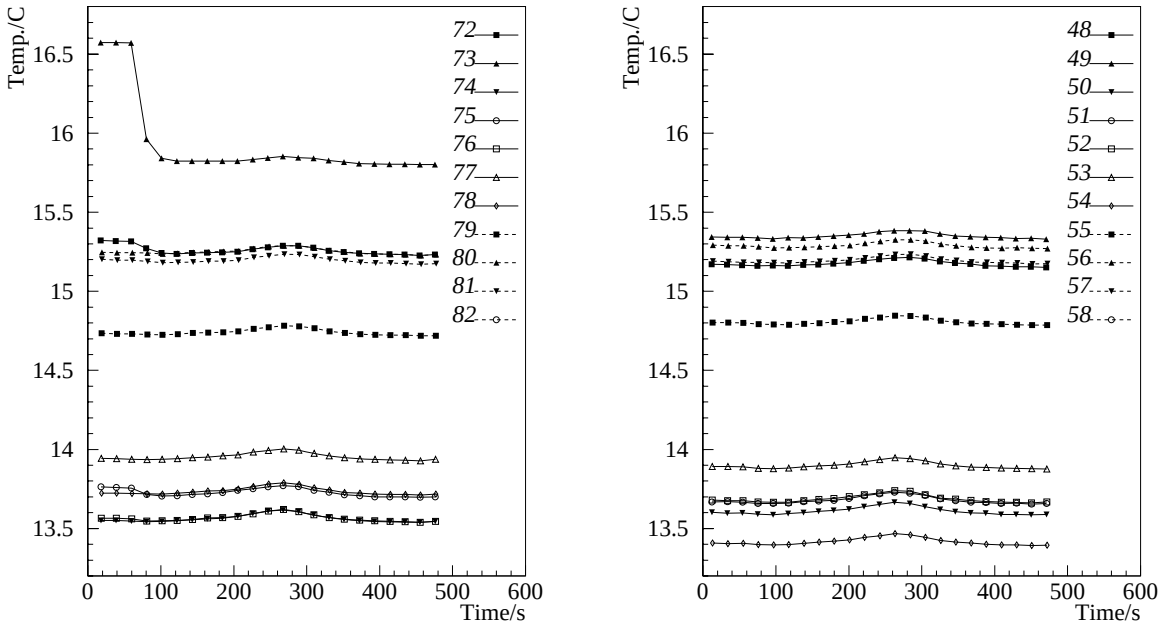


Figure 6.21: Temperature readings of the discrete sensors on the silicon detectors for module 1 (left) and module two (right) showing the switching process of detector 3.

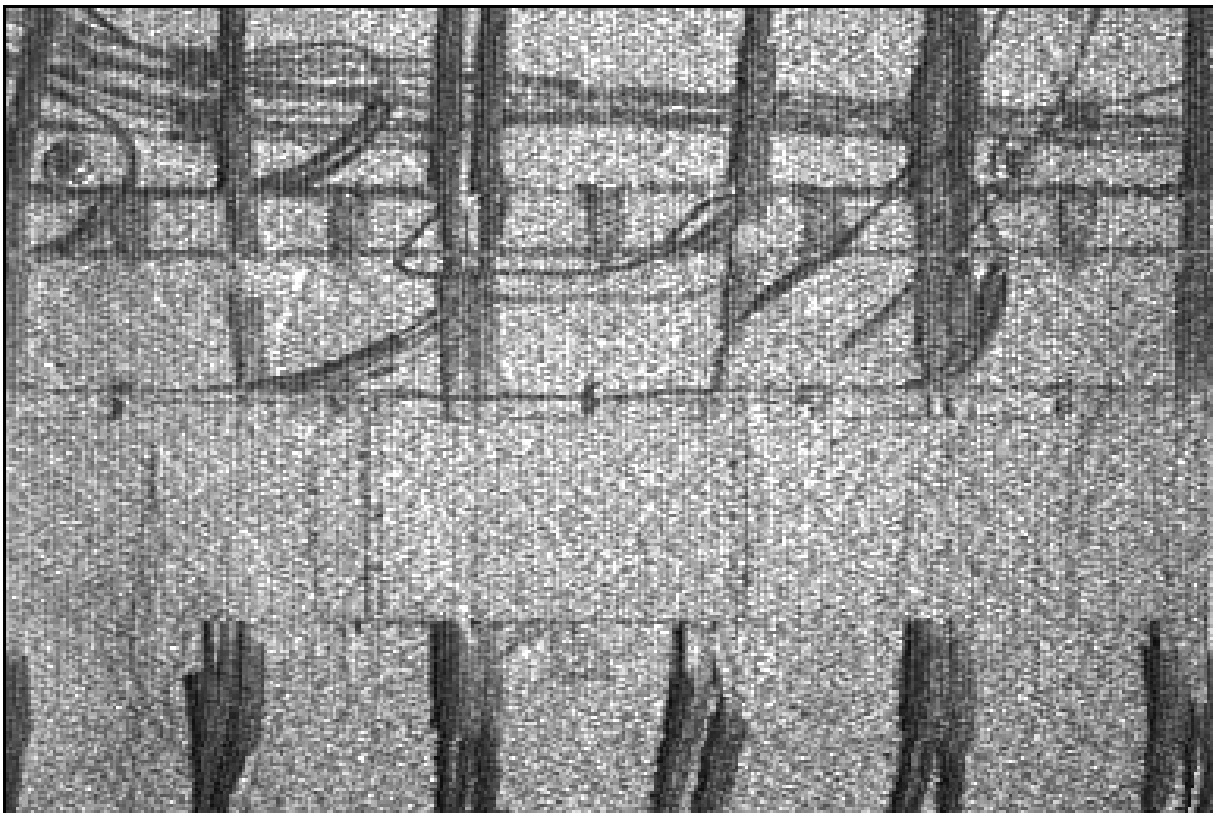


Figure 6.22: Speckle pattern of a part of the silicon detectors inside the environmental chamber as seen by the ESPI camera. Fringe pattern derived from distortions of this object are shown in figure 6.23.

pair, (detector two[27]). Because detector 3 deflects into a spherical shape it presses detector 2 down. This is the reason why detector 2 shows the set of nearly equally spaced vertical fringes indicating a rigid body rotation. The local support also shows some vertical fringes in the region of detector two. This could be due to the fact that the fixation of detector 2 or 3 is moving the local support slightly. There are not enough fringes on the image to clearly identify the source of this distortion.

[27] left of detector 3 in figure 6.23

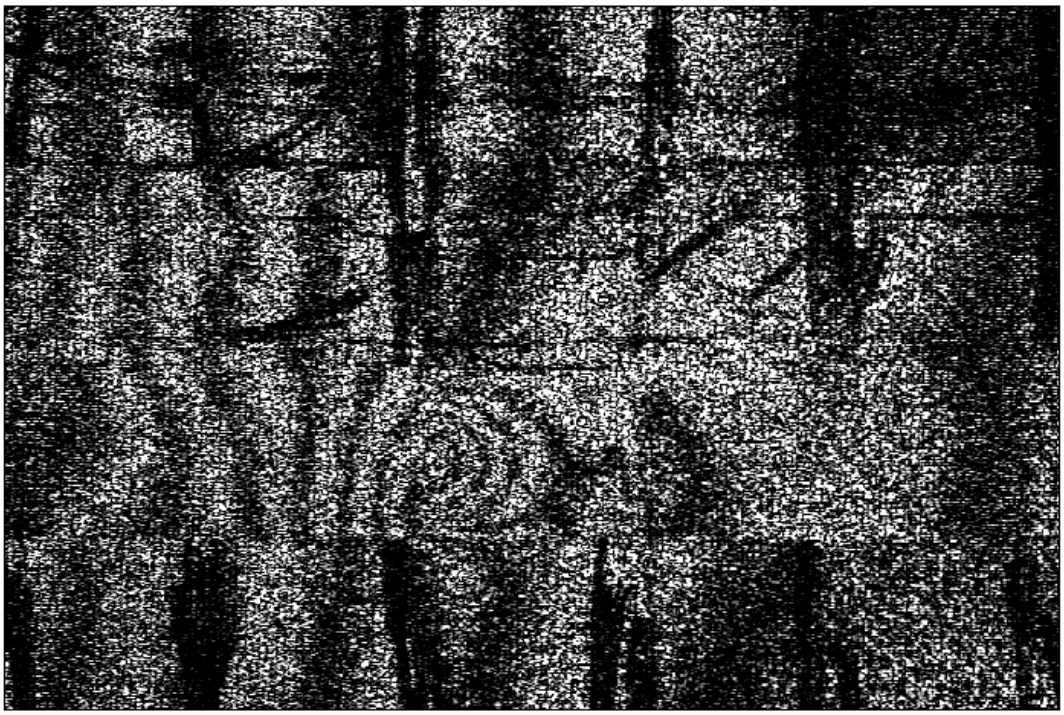


Figure 6.23: Out of plane distortion of the two modules without electronic boards mounted. The imaged object was shown in figure 6.22. The underlying temperature profiles are shown in figure 6.24. The distortion is the result of switching detector 3 (third from left) off. One fringe corresponds to a distortion of 0.26 μm.

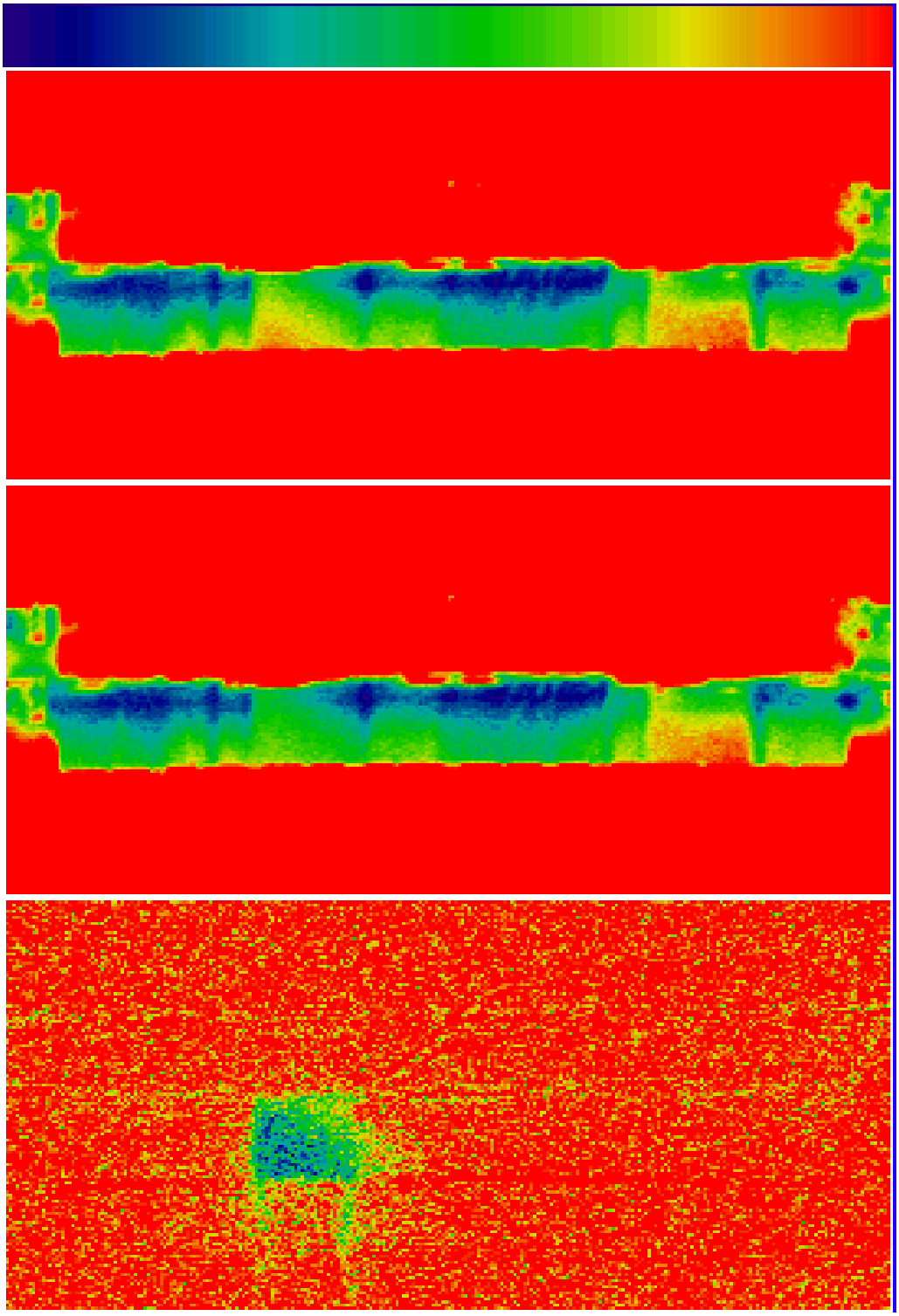


Figure 6.24: Temperature distributions of module 1 before (top) and after (mid) detector 3 has been switched off. All other detectors are without power. The bottom image is the middle image with the top image subtracted. The temperature scale on the top two images is 13.5 ^{o}C to 17 ^{o}C. On the bottom image it is -0.9 to 0 ^{o}C.

Module Switching

All detectors of module 1 (left) were operated at their nominal power of 0.4 W, those of module 2 (right) were without power. Module 2 was then switched to its nominal power. The silicon detectors of module 2 are downstream from those of module 1, whereas the boards are upstream. The cooling system was operated at an input temperature[28] of 17 ^oC and the ambient temperature in the chamber was 21.2±0.7 ^oC. Figure 6.25 shows how the discrete sensors on the silicon detectors registered the temperatures of the two modules. It can be seen that in module 2 the temperature differences on the detectors increase from

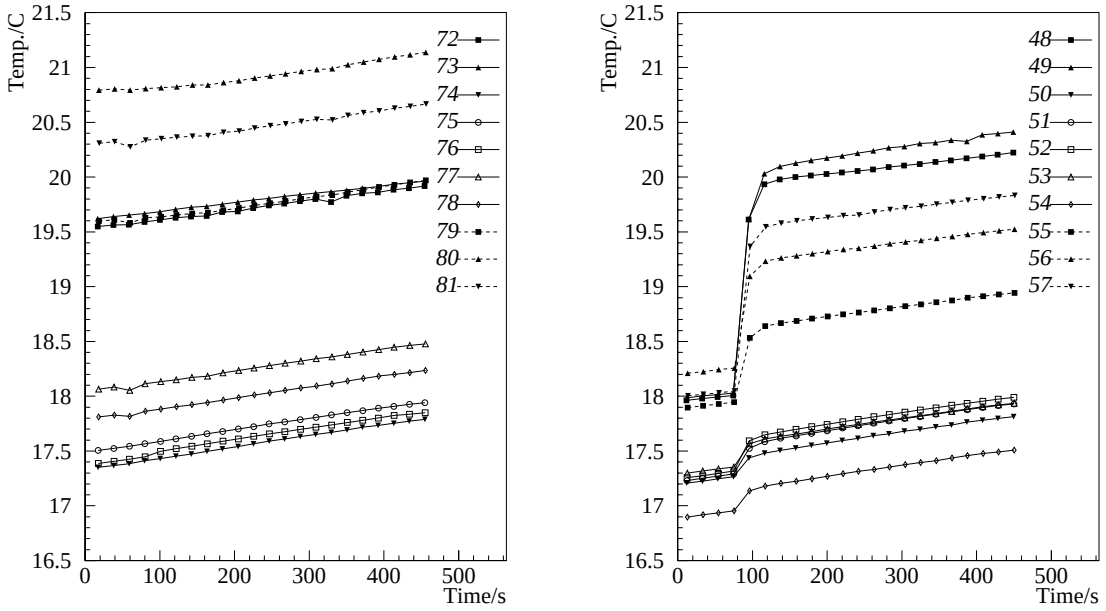


Figure 6.25: Temperature readings of the discrete sensors on the silicon detectors for module 1 (left) and module two (right) showing the switching process of module 2. The general increase of all temperatures is due an increase in the coolant supply temperature.

1.3 ^oC before the power was switched on, to 2.0 ^oC after. This allows the conclusion, that the heat load to each detector from convection with the ambient which is 5 ^oC above the coolant temperature is bigger than the load from the detector power of 0.4 W. The temperature spread of 2 ^oC across the module is within the specifications for an ATLAS module (see chapter 4) even though the total heat load to the silicon detectors is more than twice the nominal value. The IR-temperature distributions of the system before and after module 2 was switched on are shown in figure 6.26. The latter distribution allows to evaluate the silicon detector cooling. It can be seen that crystals 3 and 7 are badly cooled. This is particularly the case at the sides where they are glued to the other crystal of the pair. Both detectors are the upper detectors of their pair and the pipe has to follow their increased height with a spacer. This effect has been traced down to an increased height of the layer of thermal grease under the detectors which in term is caused by variations in the thickness of the layer of glue connecting the two crystals of a detector pair and it was found that the two pairs in question had an excess thickness of approximately 80 μm. This was caused in the gluing process because one of the two vacuum connections of the gluing jig[29] was not closed during the gluing process.

[28]The input temperature was slowly rising during the experiment

[29]see chapter 4

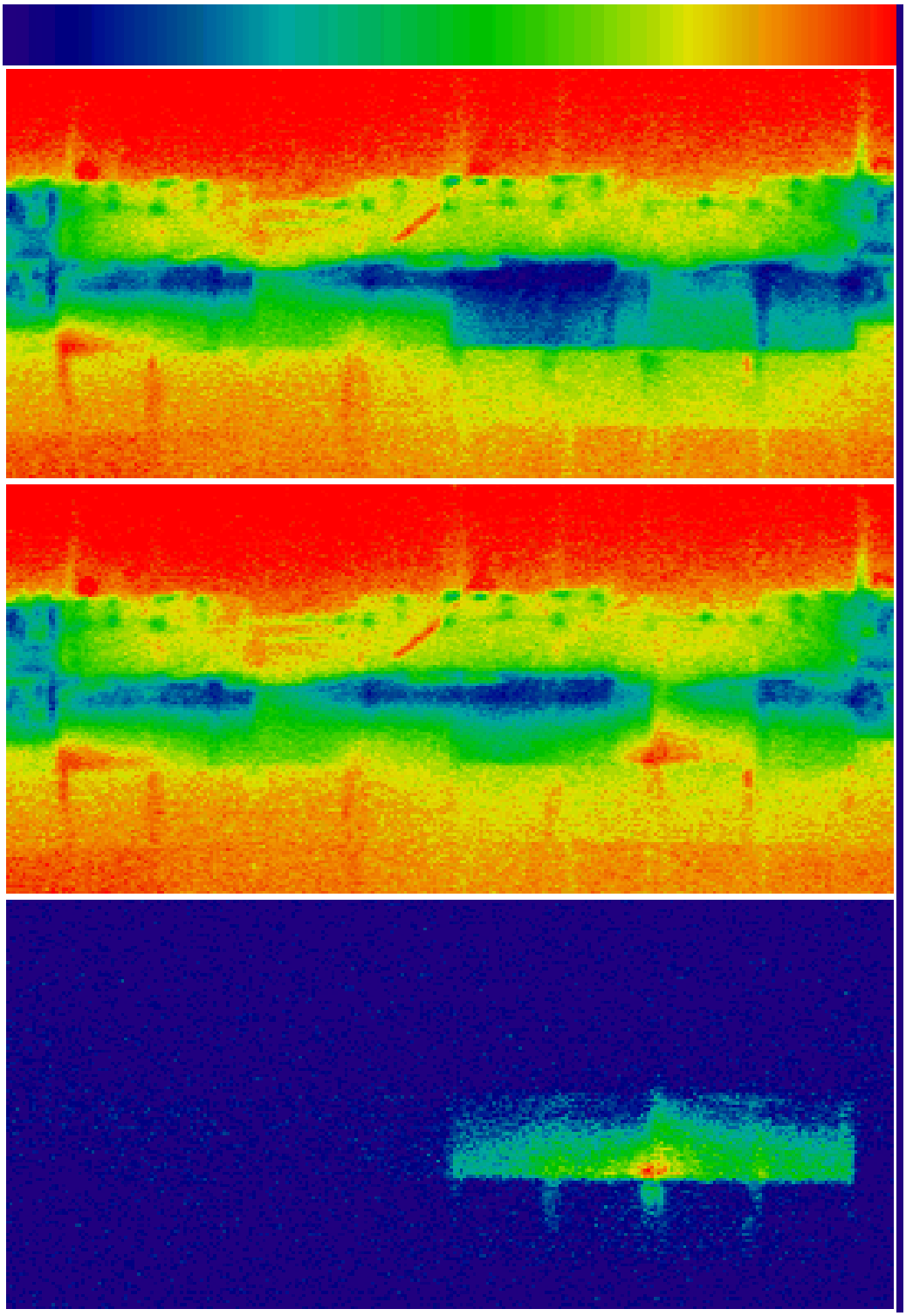


Figure 6.26: IR-temperature distributions of the silicon detectors before and after module 2 was switched on. The bottom image is the middle image with the top image subtracted. The temperature scale on the top two images is 16.8 to 21.4 ^oC. On the bottom image it is 0 to 2.6 ^oC. The detectors are numbered from left to right.

Summary

Summarising the results of all measurements of the detector modules without boards it was found that:

- The out of plane distortions were very fast and without the possibility to slowly vary the power to the detectors the distortions could only be observed for small temperature differences.
- The in plane distortions, allthough not explicitly shown as fringe pattern in this chapter were very small and could be explained by the linear expansion of the silicon detectors.
- The temperature profiles under operating conditions conformed to the specifications
- A problem in the detector cooling was found and traced back to a failure in the construction process of the detector pairs which can be avoided in future.

6.4.5 Modules with Boards

The measurements done with the complete modules had the aim to determine the deflections and temperature profiles of the modules in operational conditions. The conditions investigated were:

Start of operation: As the first measurement the cooling down of the system after switching the cooling system on is measured. This very much depends on the ambient condition before the cool-down and does not accurately represent conditions during the cool-down in the ATLAS detector. Then the power of all modules is switched on when the cooling system is already running in equilibrium.

End of operation: This test measures the reaction of the system to switching of all power.

Cooling Failure: This final test observes the reaction of the system to a complete cooling failure with all modules under nominal power.

Other tests like over powering various parts of the system have also been performed.
It is not possible to show deflection interferograms for all of the above measurements, because of the amount of data that would have to be graphically presented[30]. Because the analysis of ESPI data is very labour intensive most of the data has not been immediately analysed and only quality checks have been made during the measurements. Analysing the in plane measurements it was quickly found, that in plane distortions are well below the specified limits. Therefore it has not been attempted to fully analyse the dynamics of the distortion and how they correlate with the temperature changes. Where single images yield interesting information they are presented.

Start of operation

The total in plane distortion in the $r\phi$ direction observed during the start of operation, excluding the free cooling pipe, was less than 1.5 μm. The coolant on the input of the module was 5 ^{o}C and the ambient temperature was 25 ^{o}C. The out of plane distortions measured in a separate run were very fast, in particular on the electronic boards. The total distortion can only be estimated as an upper limit because the images were not phase stepped. The limit is derived by simply counting the maximum number of fringes across a series of images.
Out of plane measurements were made on only a part of the setup to reduce the problem of the large distortion velocity. The imaged area is shown in figure 6.27. The detectors visible on the image start on the left with the right third of detector 1 of module two[31]. In the place where the first power lines enter from top (one line) and bottom (two lines) detector two starts[32]. Then the board of module one is seen, covering the right third of detector 2 and all of detector 3. The right end of the image is marked by the left edge of the board of module 1. In the horizontal direction the image spans the length of 5.3 detectors. From top to bottom the image shows the mylar window frame[33], the base plate, the local support with the top free two cooling lines, the two modules, again a thin stretch of the baseplate and the window frame.

[30] More than 2 GB of raw ESPI data were acquired
[31] The rest of detector 1 of module two is left outside the image
[32] the transition between the detectors is usually not visible
[33] the frame is partially transparent and allows to vaguely see the power lines

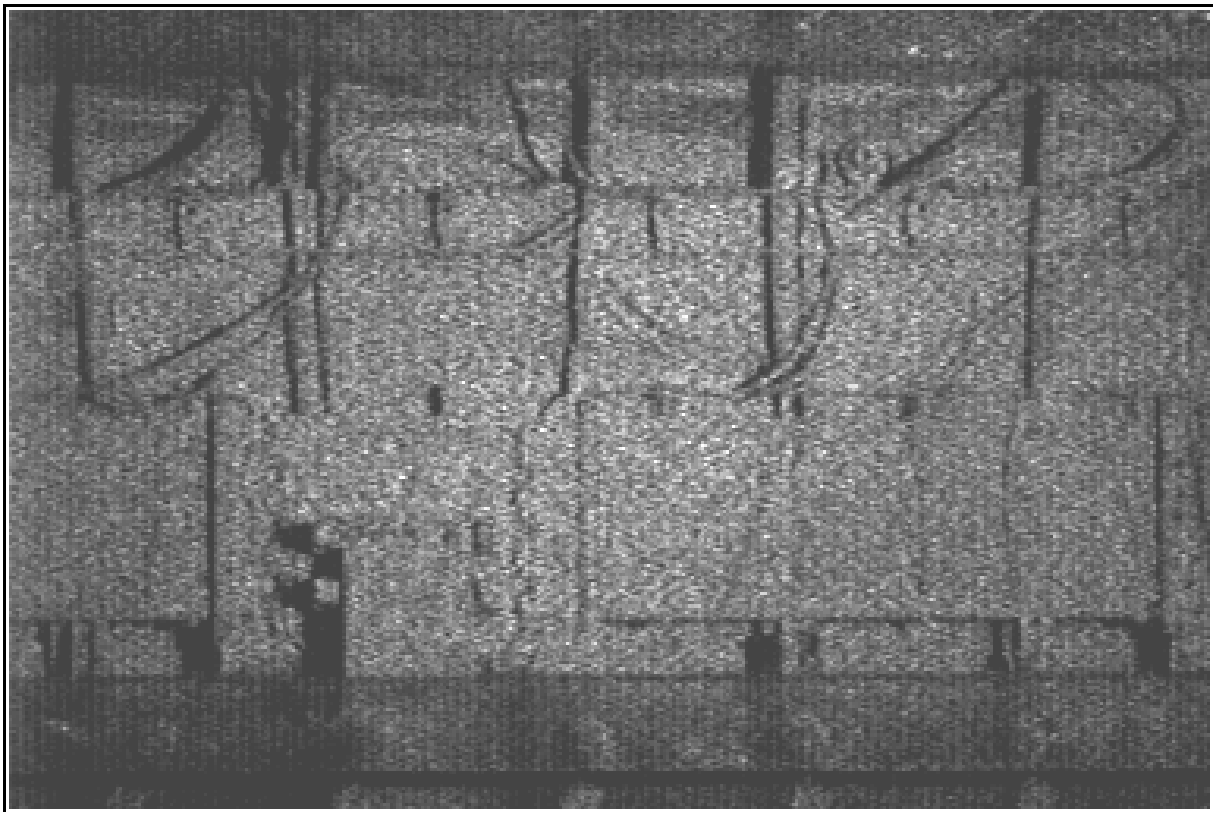


Figure 6.27: Fraction of the setup shown in figures 6.28 and 6.35 as seen by the ESPI camera with only the object beam present. The brightness has been scaled up by 50% and the dynamic range has been halved. The image is explained in subsection 6.4.5.

The out of plane distortions for the first 13 seconds of the cool-down are shown in figure 6.28. The first obvious point is, that the fringe pattern on the module is continuous and does not break[34] between detectors or between boards and detectors. This indicates that for the small distortions the flexible connection between the boards and the detectors is effectively rigid. There are also a few fringes on the local support. The fringe contrast is poor for two main reasons. The mylar window can not be stabilised against microphonic vibrations and the reflections from the window add light with a quickly varying phase. The main source of vibration in the setup was acoustically transmitted hum from the Laser power supply and the various pumps in the room. The second reason for the bad contrast is the failure of every second row of pixels in the reference beam illumination due to an interference effect with the camera face plate. This problem and the intensity variations in the beam can be seen in figure 6.29.
All areas but the boards distorted by less than 22 μm out of the plane due to the cool-down from room temperature (25 ^oC) to the equilibrium with the coolant at 5 ^oC. The typical temperatures on the silicon detectors ranged from 6 to 10 ^oC depending on the distance from the cooling pipe. This rather large difference is due to the still comparably warm ambient air in the box which was 19 ^oC close to the surface of the base plate. The ambient outside the box was 29 ^oC. Four IR-images of the cooldown phase are shown in the first four images of figure 6.32 on page 102. This situation would not happen inside ATLAS, because the gas around the detectors would be completely independent of the external temperatures. Figure 6.30 shows the temperature readings of the sensors on the silicon

[34] A break in a fringe pattern is seen as the end of a fringe

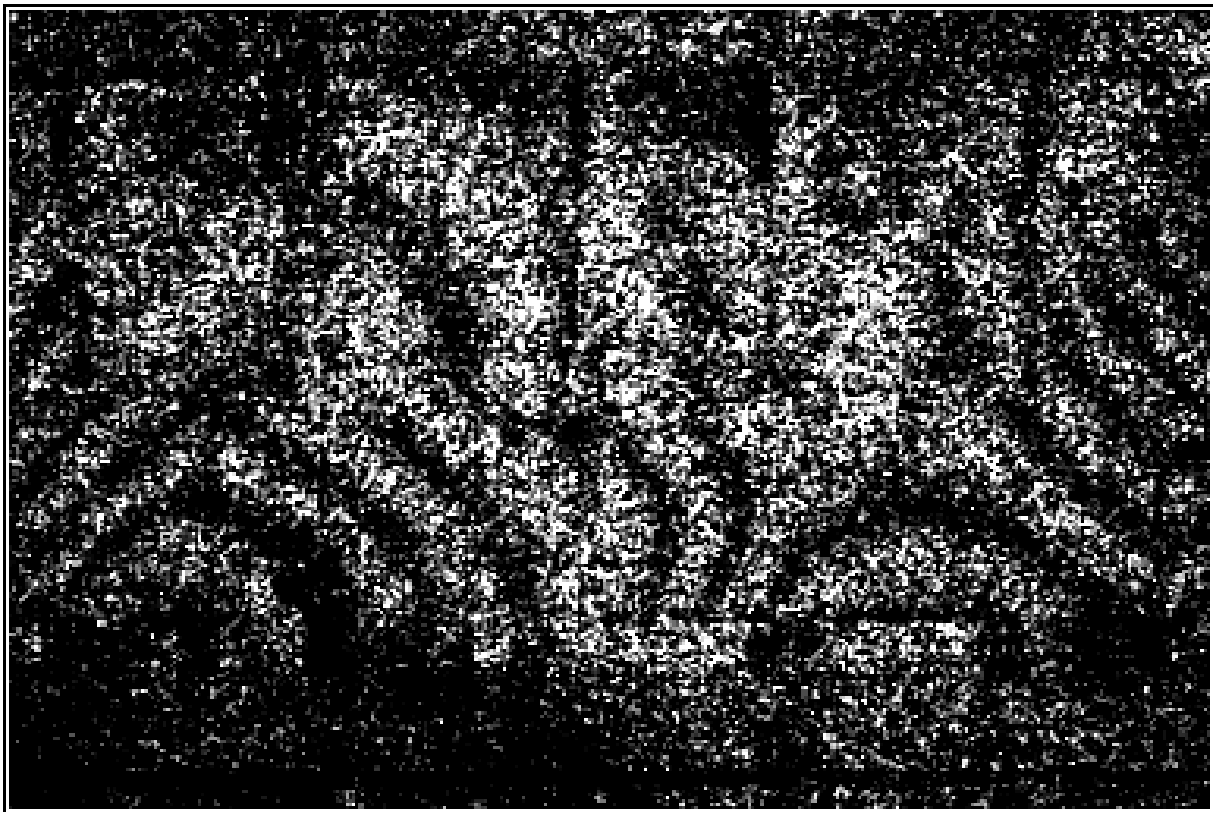


Figure 6.28: Out of plane distortion of the part of the setup shown in figure 6.27 resulting from the first 13 seconds of the cool-down. The distortion represented by one fringe is 0.25 μm. The detector temperatures changed by a maximum of 3.8 ^oC.

detectors after the cooling was switched on and how they changed when the power was switched on. In figure 6.30 two categories of sensors can be distinguished. Those which are situated directly above the cooled of the detectors area (module 1: 50,51,52,54; module 2: 74,75,76,78) and those which are separated from the cooled area by some length of silicon. It can be seen from figure 4.10 in chapter 4 that the temperatures are roughly proportional to the distance from the pipe.In the same process the boards distorted approximately 3 times more than the detectors. This could not be precisely measured because the distortion was too fast in several images. The temperatures on the boards varied from ambient temperature of 25 ^oC to 6.5 ^oC.
The subsequent turn on of the power only added 3 μm distortion to the detectors and 6 μm to the boards (no ESPI image given). As can be seen in figure 6.31, the temperatures on the boards initially spread from 6.5 to 12.2 for board 1 and from 6.5 ^oC to 13.5 ^oC for board 2. After the power was turned on the spreads increase and were 7.6 ^oC to 13.0 ^oC for board 1 and 8.1 ^oC to 16.7 ^oC for board 2. An IR-image of the modules under power with the cooling system active can be found in the 5th image of figure 6.32 on page 102. Three groups of temperatures on the boards can be distinguished. Those on the chips near the edges of the boards[35] (board 1:58-65, board 2:82-89), the one on the ROC chip (b1:68, b2:92) and those that are connected to heat sources or cooling pipes only via the board material. It is obvious that the first group stays closest to the pipe temperature and is least affected by the power increase. This is expected, as they are directly attached to the straight sections of the pipes. The temperatures on the ROC are both slightly higher

[35] the front end chips in a real board

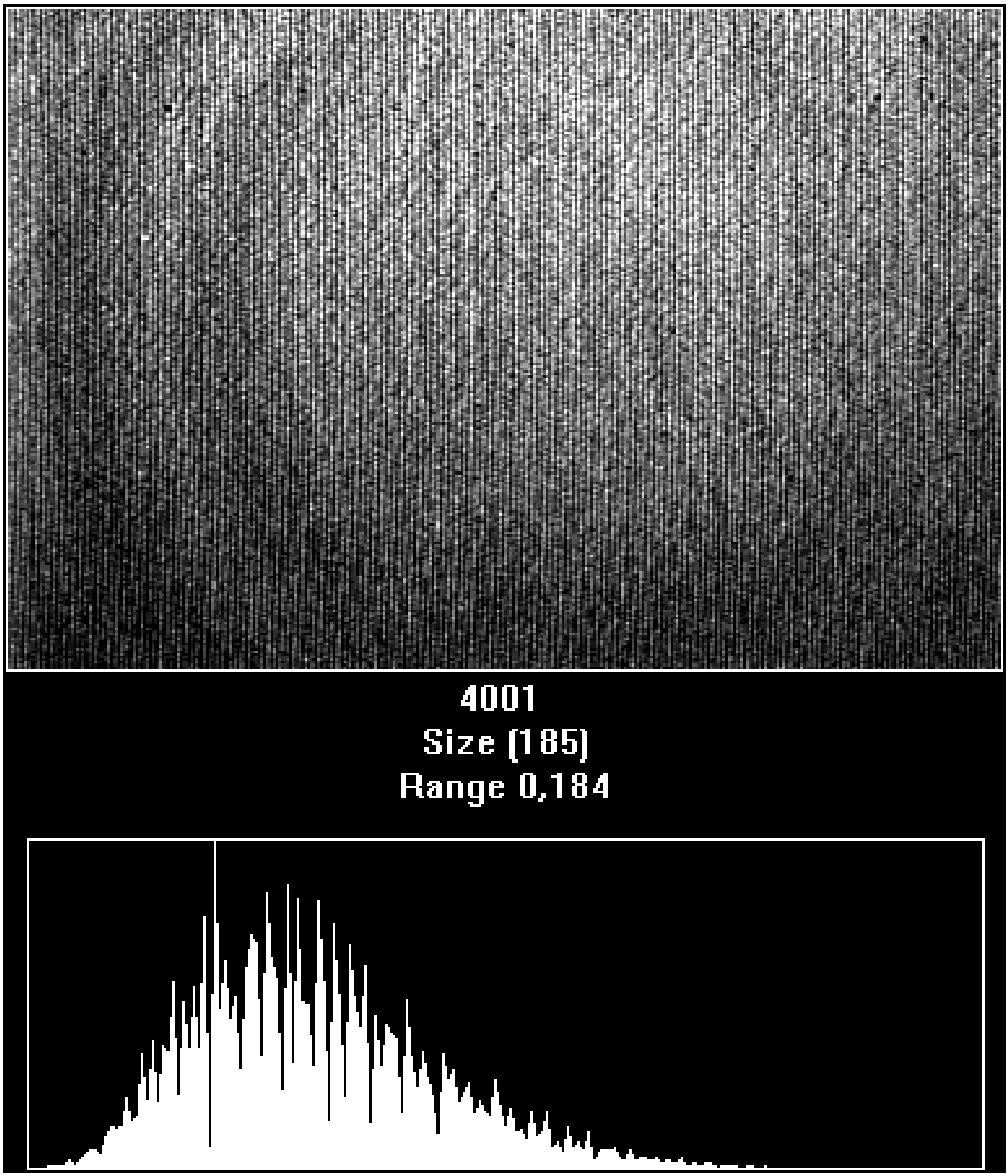


Figure 6.29: Image of the reference beam with an intensity histogram below. The vertically striped appearance is probably the result of an interference on the CCD face plate.

because the pipes bend in these places and the bending causes a distortion of the cross sectional shape of the pipe which prohibits a good contact to the chip. This is particularly obvious for board 2. A similar effect is visible on the front end chips near the pipe bends in high power operation. The temperatures of the third category show the effects of the thermal conductivity of the board material. For example sensors 93 and 69 which are located near the cable connectors show only a weak response to the power increase as they are separated from the power sources by the board material.
Figure 6.32 shows IR-images for several scenarious discussed in this chapter. A summary

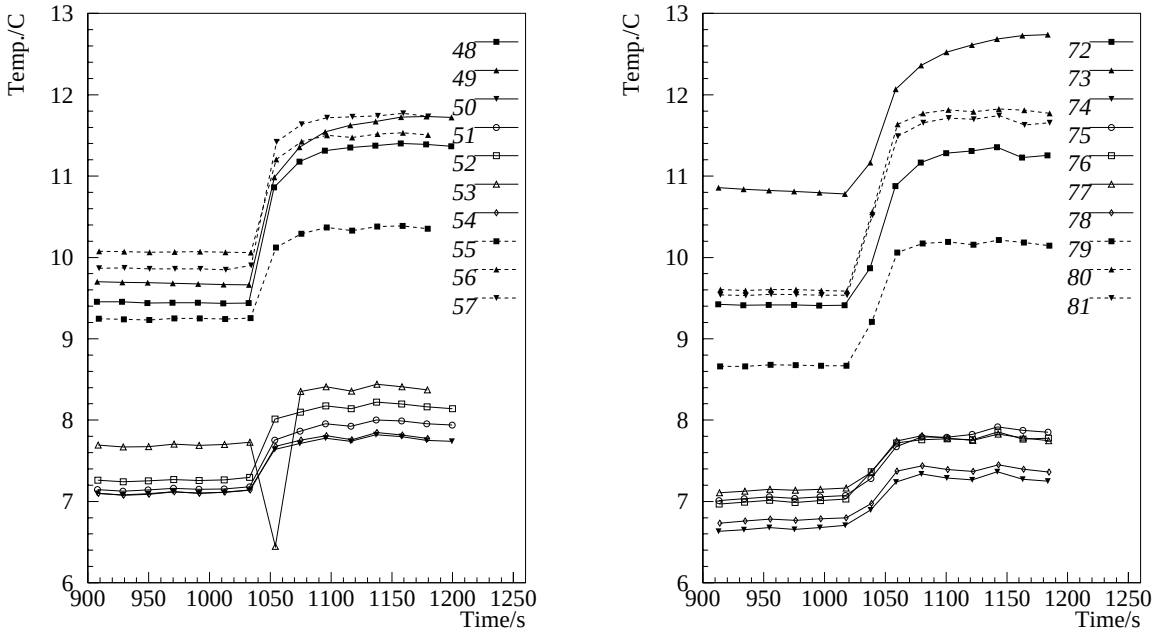


Figure 6.30: Temperature readings of the sensors on the silicon. Initially there is no power in the modules and the coolant input temperature is 5 ^oC. At t=1030 seconds the nominal power is switched to the modules.

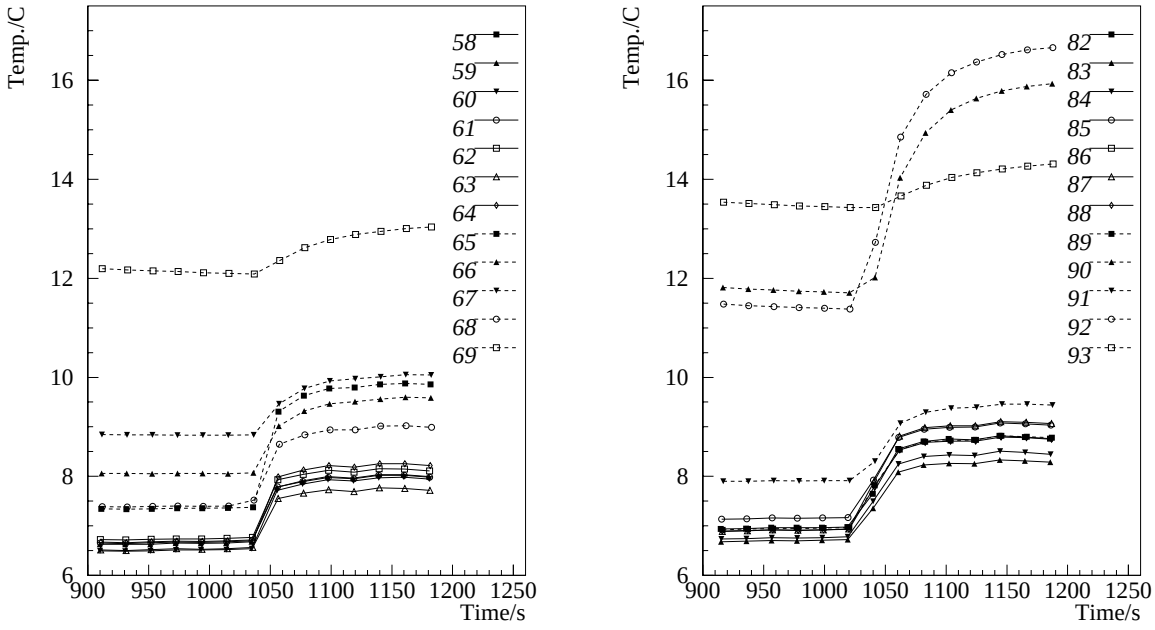


Figure 6.31: Temperature readings of the sensors on the electronic boards. Initially there is no power in the modules and the coolant input temperature is 5 ^oC. At t=1030 seconds the nominal power is switched to the modules. Sensor 86 is not on the plot because it was shorted due to a problem in board 2. Board 2(left) is the kevlar board, board 1 (right) is the glass fibre board.

of the individual situations for each image in figure 6.32 can be found in table 6.3 and are described in detail throughout this chapter.

image	phase	time [s]	power boards [W]	power detectors [W]	cooling
1	cool-down	40	0	0	starting
2	cool-down	76	0	0	on
3	cool-down	112	0	0	on
4	cool-down	281	0	0	equilibrium
5	operation	144	2*0.93	8*0.4	equilibrium
6	overpower	192	2*5.8	8*0.4	equilibrium
7	cooling failure	36	2*0.93	8*0.4	stopping
8	cooling failure	96	2*0.93	8*0.4	stopping
9	cooling failure	192	2*0.93	8*0.4	off

Table 6.3: Parameters describing the situations in which the images of figure 6.32 have been taken. The images are numbered from top left to bottom right. The time column gives the time in seconds after the start of the operation described in the "phase" column has been carried out. The "on" entries in the "cooling" column mean that the coolant in the detector is at nominal temperature. In the situation described by "cooling off" a small residual coolant flow remained.

Overpower

The IR-temperature distributions for the case of a 6.25 fold overpowering of the boards is shown in figure 6.32 in the 6th image. The increase in temperatures for the boards and the silicon detectors as registered by the discrete sensors is shown in figure 6.33. It can be seen, that the ROC chip of board 2 is significantly overheating. Unfortunately this chip can not be seen in the IR-images as it is covered by the flat cable of the of module 2 (left). The sensors on the chip, as well as the one under the chip (92 and 90) both rise in parallel. Sensor 73 on the lower side of the silicon, near to the ROC of board 2 also registers much higher temperature than the other sensors on the detectors of module 2. A similar excursion is observed on sensor 66 of board 1. This is seen in detail in the enlargement of figure 6.34 on the bottom sensor of the right row of module 1 (right). Both excursions could be traced to the problem of of the distortion of the pipe cross-section in the bending process that has already been identified as a problem during the start of operation.
The total distortions in the overpower situations are small. Out of plane distortions on the detectors are smaller than 4 μm and on the boards they are below 12 μm. In-plane distortions in the $r\phi$ direction are below 1 μm and in the Z direction they stay below 5 μm. A typical image of the out of plane distortions is shown in figure 6.35. It shows the part of the setup shown in figure 6.27. The most obvious feature in the image is the big distortion of the pipe. In spite of the small area and reduced resolution the presence of bright fringes can be clearly seen. The reduced horizontal resolution does not allow a precise count of the fringes, because they are broken by the striped pattern explained in figure 6.29.

Cooling Failure

The most severe changes of the temperature distributions are induced by a cooling failure. The IR-images of the cooling failure are shown in the last three images of figure 6.32. The

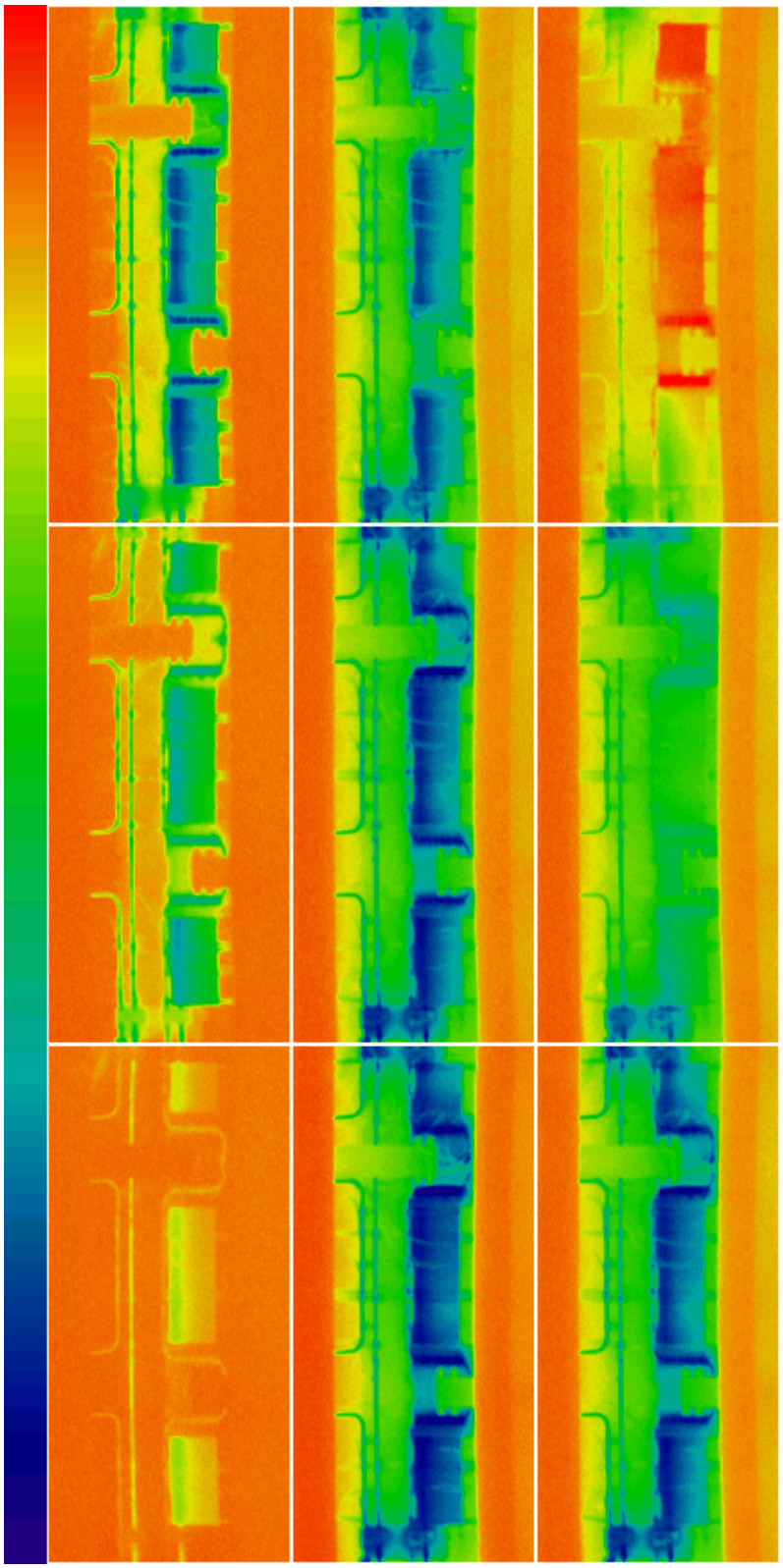


Figure 6.32: IR-temperature distributions of the full detector setup for situations described in table 6.3. The temperature scale on all images ranges from 7 to 27.3 ^{o}C. The visible object section is equivalent to the size of the mylar window.

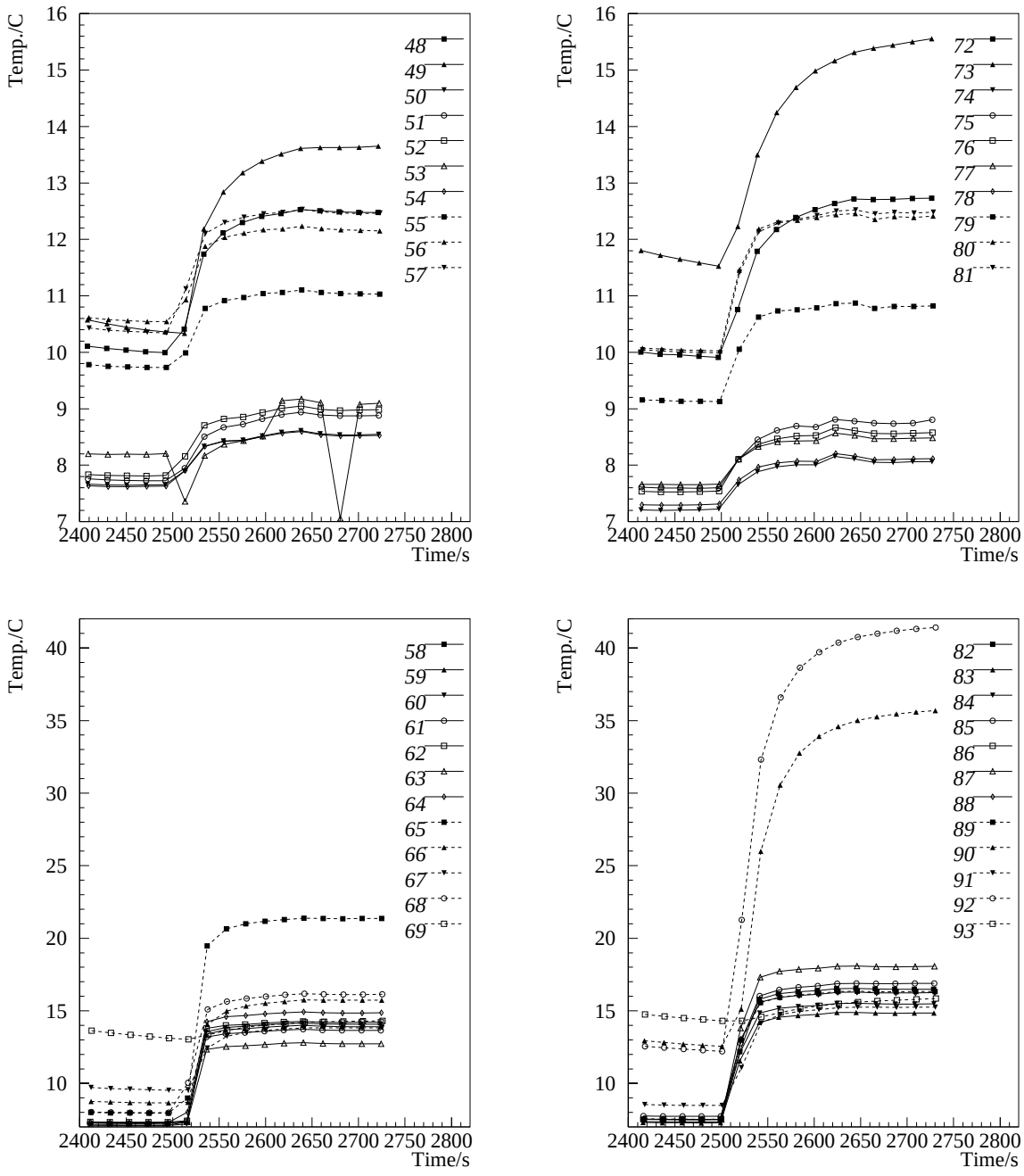


Figure 6.33: Discrete temperature readings during a 6.25 fold overpowering of the boards which were finally dissipating 5.8 W each. The top row shows the readings of the sensors on the detectors and the bottom row shows the readings on the boards. The left column is for module 1 (glass fibre) the right is for module 2 (kevlar).

temperature changes in the detectors and boards can be seen in figure 6.36. The large differences between the modules can be explained as follows.
The cooling failure was simulated by opening a large cross-section bypass valve and thus talking away most of the pressure across the modules cooling system. The bypass valve was used, rather than shutting down the system, in order to leave the cooling system circulation and preserve the coolant temperature during the test. The pressure drop across the bypass was not 0 and therefore a small residual flow below the measurement limit of 0.1 l/min remained. This flow first reached the silicon detectors of module 2, then those of

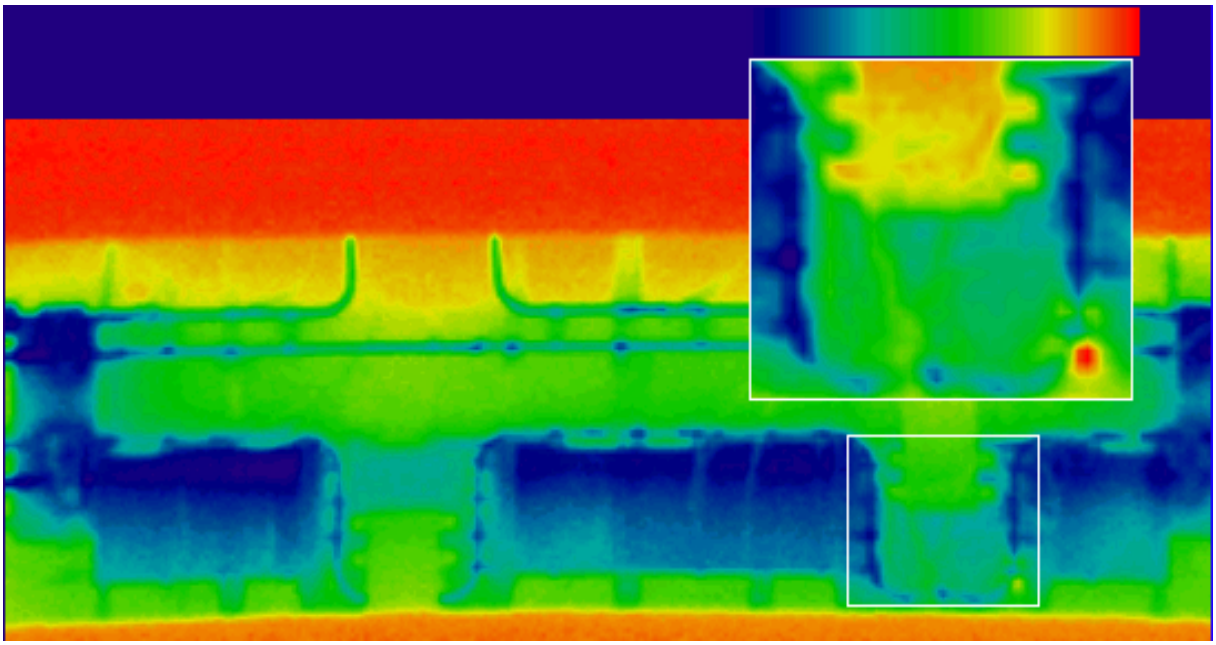


Figure 6.34: Enlargement of the 6th image of figure 6.32 with board 1 (right) enlarged and rescaled. The colour scale for the rescaled insert corresponds to temperatures from 11.2 to 21.1 ^oC , for the large image it corresponds to temperatures from 9.7 to 25.7 ^oC. The boards have been overpowered to 5.8 W.

Figure 6.35: Out of plane distortions resulting from a power increase of both modules. The object in the image is shown in figure 6.27. The power in the boards was switched to 5.8 W which is 6.25 times the nominal power. The silicon detectors were switched to their nominal power of 0.4 W each. The most prominent change is the expansion of the free cooling pipe due to the increase in coolant temperature.

module 1, after this it reaches the board of module one and finally the board of module 2. This explains why the sensors on the detectors of module 2 register a very big spread from 18 to 24 ^oC. Those parts of the detector which are upstream are colder and they get hotter the further downstream they are. When the coolant reaches the detectors of module 2 it is already much hotter and can not take away significant amounts of heat any more. Therefore the spread of the temperatures in these detectors is much smaller. The sensors on the board of module 1 see the coolant flow next and are on average colder than those on the board of module 2. They also show the bigger spread because the coolant is still able to take away some heat. On board 2 the coolant is not able to take away significant amounts of heat anymore. This can be seen by the very small spread and by the on average 2 ^oC higher temperature. The sensor on the ROC of board 2 is again the exception as it's coupling to the pipe is worse. The influence of the remaining flow can also bee seen in the last image of figure 6.32. It has to be remarked that, as can be seen in figure 6.36, the temperatures during the cooling failure were not yet in equilibrium. It can be estimated from the shape of the curves, that the equilibrium temperatures would possibly have been 5 to 8 ^oC higher than the maximum values registered here.
The out of plane distortions during the cooling failure stayed below 11 μm on the detectors and 18 μm on the boards. In plane distortions in the $r\phi$ (z) directions of less than 1.5 μm (7 μm) where measured. The out of plane distortions have a big uncertainty in this case due to the loss of some images which exceeded the fringe resolution limit.

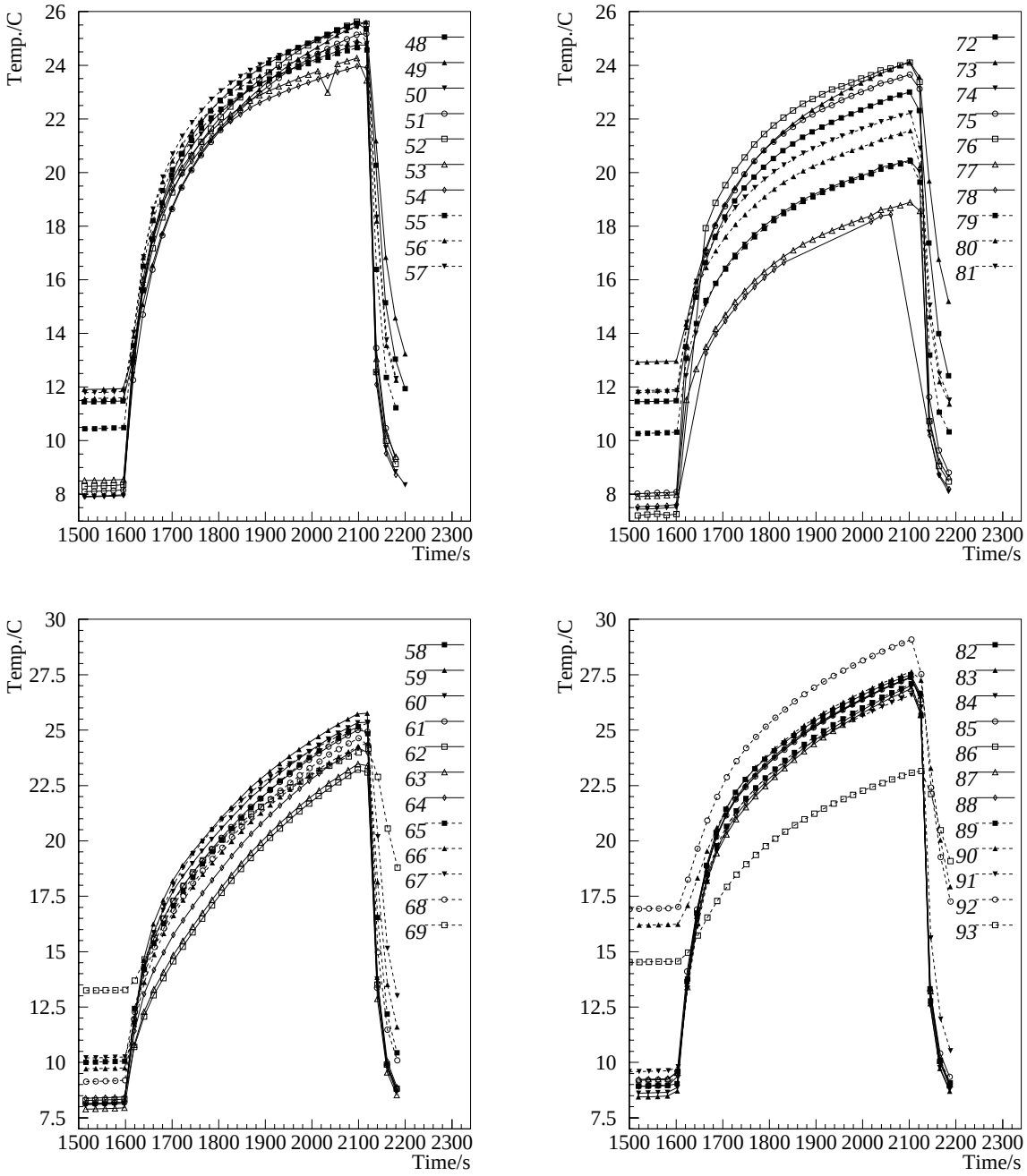


Figure 6.36: Discrete temperature readings during a cooling failure. The boards were dissipating 0.9 W each. The top row shows the readings of the sensors on the detectors and the bottom row shows the readings on the boards. The left column is for module 1 (glass fibre) the right is for module 2 (kevlar).

6.5 Conclusions

ESPI tests of the components of the silicon tracker modules have been made, increasing the complexity and size of the object under study in several steps and measuring the distortions of the tracker module layer by layer as the module has been built up. Thermal measurements have been made in parallel and have been used as input to FE-model predictions, specifying the thermal as well as the mechanical boundary conditions. A finite element model was tuned to correctly predict both temperature and deflection measurements of a test setup. A phase stepping algorithm for the ESPI has been developed and a resolution of 13 nm across a surface area of 1311 cm^2 has been measured after applying a rank mean filter to the measured phase maps.

6.5.1 Modules

The tests of the silicon tracker modules showed that the in plane distortions of the silicon detectors induced by changes in the operating conditions are of the order of a few μm in the $r\phi$ direction and always stay below 10 μm in the z direction. These values are well within the specifications for an ATLAS tracking module. The out of plane distortion are significantly bigger and faster and show a clear separation of the distortions in the boards and the silicon detectors. The boards distort by up to $\mathcal{O}(100)\mu$m but the detectors stay stable to better than 22 μm. This is an indication that the intended mechanical separation between the silicon and the boards was successful. The temperature distributions show a similar effect of separation. For an average increase in board temperature of 5 ^oC the silicon detectors only show an increase of 1 ^oC [36]. This thermal and mechanical separation allows the use of low thermal conductivity and cheap board materials. It also reduced the out of plane distortions of the detectors which, in a rigid construction of the same components would be intolerably big. In retrospect the aspect of building the module in a symmetric way, avoiding large single sided mismatches component CTE has been underestimated. It is believed that a symmetric structure has the best chances to avoid the problem of large out of plane distortions. The symmetry has to be applied to the mechanical structure as well as to the heat flow, requiring symmetric cooling. Driven by the measurements described in this work this principle has now been adopted into the latest design of a silicon tracker module for ATLAS.
The cooling system in normal operation functions sufficiently and keeps the temperature variation on the detectors below 4 ^oC. This is close to the tolerable limit. It has to be noted, that the power dissipation in the silicon detectors was assumed conservatively large (0.4 W/detector), but that the power was independent of temperature. In a real silicon detector the power dissipated would increase exponentially with temperature, doubling every 7 ^oC.
The temperature distributions on the boards show very good cooling of most FE-chips[37] and the heat can be removed directly at the source, avoiding heat transfer through the board material. Under nominal operating conditions the temperatures on the boards can be kept within 5 ^oC difference from each other[38]. It was also found that small variations in the glue thickness connecting detectors into a pair, can lead to problems with the contact to the cooling pipe. Bending of the very delicate pipes can also cause problems in the

[36] see figure 6.33

[37] with the exception of the ROC of board 2

[38] Note that the role of sensors 69 and 93 in the measurements is different to real ATLAS operation because the boards are absorbing heat from the ambient air, rather than emitting.

planarity. It is therefore considered useful to be able to equip the pipes with a machined surface where they have to be in contact with the modules. It is also advisable to increase at least one pipe dimension to achieve more robust contacts.
The measurements could be improved by the use of a real environmental chamber with an externally determined gas temperature in the chamber. The final temperature of the gas in the ATLAS inner detector cavity is not yet clear. It is **not a parameter that can be controlled** but will evolve from the boundary conditions of the module temperature distributions and the small amount of heat that can be removed with a gas flow. This evolution is not readily calculable with finite element models and can not be measured directly unless a significant fraction of a detector layer be built and operated. It is therefore crucial to measure the ratio of heat removed from the detector system via the cooling pipes to that removed by the gas, as a function of the gas temperature. This measurement was not possible because the external sources of of heat entering the system could not be accurately enough controlled. The use of an actively controlled environmental chamber together with the monitoring of the power removed by the coolant flow in a situation with known power input to the modules would allow these measurements to be taken. The real ambient temperature could then be estimated to be that, at which only a negligible fraction of the heat is removed by the gas.

Chapter 7

Pattern Recognition

This chapter describes a pattern recognition algorithm for the ATLAS inner detector and it's performance. Results from studies using the algorithms to identify single muons from 2 to 500 GeV with and without minimum bias pile-up are shown. The detector geometry is that described in the ATLAS technical proposal. The application of the algorithm to the identification of $\tilde{g}$-pairs will be described in chapter gluino.

7.1 Introduction

The silicon detectors described in the previous chapters are one essential component of the ATLAS inner detector. As mentioned before in chapter 3.2.1 the inner detector is based on the combination of two different classes of detectors, the continuous TRT and the discrete detectors, of which the silicon strip detectors are one. Having verified the mechanical feasibility of the silicon strip detectors, the most important remaining task is to show that the concept of combining two types of tracking devices does in fact allow efficient and robust pattern recognition and good track parameter measurements in the particular geometry choosen for ATLAS. This is the main motivation for the work presented in this chapter

Up to now all pattern recognition approaches using the complete inner detector of ATLAS were initiated by finding a track segment in the TRT. Since the discrete tracking detectors have much lower occupancies than the TRT a pattern recognition based on the discrete detectors could be advantageous, in particular at high luminosities. *iPatRec*, a pattern recognition program using the discrete tracking detectors only, already existed. Since *iPatRec* was not using the TRT a gross degradation of the momentum resolution and an unnecessarily high fake track rate had to be accepted. This work was intended to complement the large effort which has already gone into *iPatRec*, helping to overcome the momentum resolution limitations whilst maintaining and possibly improving the high track finding efficiency. Adding the use of the TRT to *iPatRec* was also expected to improve the stability of the combinatorial pattern recognition in high multiplicity events, by adding an independent element to the pattern recognition algorithm.

7.2 The Framework of *iPatRec*

iPatRec is a pattern recognition program for the reconstruction of charged particle tracks in the inner tracker of ATLAS. It is intended to measure track parameters in the transverse momentum range from 1 to 500 GeV. The identification is not restricted to isolated tracks but can be performed inside jets and with minimum bias pile-up as expected at maximum

LHC luminosities. *iPatRec* uses an internal description of the inner detector geometry in the form of a geometry table which is created during initialisation using the information present in the detector geometry banks. The geometry table describes the detector in terms of barrel and disk shaped layers. The layers are grouped into superlayers which are spatially neighbouring layers capable of giving space point information. The only exception from this is the TRT which by design can not deliver space point information. Those detectors capable of measuring space points are referred to as discrete tracker elements. These are:

- Pixel Detectors
- Silicon Strip Detectors
- GaAs Strip Detectors
- MSGC Detectors

The superlayers enter the geometry table in order, going out from the vertex increasing first in radius and then in η. The geometry table allows *iPatRec* to predict which detector elements should have been hit by a certain track and therefor establish the concept of holes on a track where the detector did not measure a hit, even though it was expected.

Pattern recognition in *iPatRec* is initiated by an external seed indicating a region of interest in the inner tracker. This region is called a seed road. Every seed road includes the full vertex region up to $\pm 2 \cdot \sigma_{z-vert}$ and $\pm 20 \cdot \mu$m in the transverse directions. Its extent at the outer radius is given by the precision of the seeding detectors extrapolation.
The road is symmetric in $r\phi$ and can include particles down to about 1 GeV depending on the type and minimum momentum of the seed particle.

The four seed types currently supported are:

Electron Seeds The barycenter of each cluster in the electromagnetic calorimeters provides one seed. The road width in the $r\phi$ coordinate is larger than that for muons to allow for bremsstrahlung processes.

Muon Seeds Currently the roads for muon seeds are not using tracking results from the muon toroid system due to problems with the stability of the muon tracking code. Instead the information is taken from the *KINE* bank and smeared according to the expected tracking resolution of the muon tracking system.

B Seeds The *GENZ* bank is scanned for b-quarks and a road suitable to contain a b-jet is constructed around each b-quark. This seed type has been extended to search for arbitrary particle types in the *GENZ* bank.

KINE **Seeds** A general purpose seeding routine is supplied to scan the *KINE* banks for any particle required by the user and assign suitable road parameters for the particle types.

The seeds used in this note were of *KINE* type. Once the seed is defined *iPatRec* uses a combinatorial approach to find the best combination of points from discrete tracker elements, joining these to a track segment found in the TRT. A more detailed description of *iPatRec* can be found in [17].

7.3 Global Pattern Recognition Algorithm

Track finding starts by initiating a track skeleton, which is a fit to the seed road parameters plus 2 space points from different superlayers. To account for moderate detector inefficiencies, 3 or 4 superlayers are designated to provide these space points, normally 2 at the outer tracker radius (where the road width and occupancy give the fewest pile-up combinatorials) and 2 in the 'sagitta' region (mid-radius) to give sharp Pt thresholds and to provide a reasonable extrapolation precision. These designated superlayers are referred to as skeleton superlayers. All combinatorials amongst these superlayers are considered. A skeleton is valid if:

- its p_t is inside the seed road,
- it points back to the vertex within $\pm$50cm in z and
- it could be successfully fitted with a χ^2 below a cut of typically 3 p.d.f[1]

Valid skeletons are then extrapolated into the TRT where they define a road around the centre of the skeleton extrapolation. The width of the road is given by the errors on the skeleton track parameters. It has typical (p_t dependent) dimension of 5mm in the transverse track direction when it enters the TRT. The number of TRT hits, inside the skeleton road is used to veto wrong skeletons. The minimum number of hits required is a function of eta and is shown in figure 7.1 of section 7.4.3. It has been constructed as the lower envelope of the number of straws hit by a single track which is also shown in figure 7.1
If a skeleton has passed the veto cut pattern recognition in the TRT is performed, the left-right ambiguity arising from the drift time measurement is solved. The number of hits accepted by the pattern recognition during confirmation undergoes a similar cut as the number of hits in the skeleton road. This cut can be seen as a function of η in figure 7.1 in section 7.4.3. It has been constructed to be an 80% fraction of the total number of hits found on the track. Its relation to the number of confirmations on single track events is also shown in figure 7.1
The procedure for solving the left right ambiguities is the core of the TRT pattern recognition and is described in more detail in section 7.4.
The coordinates and errors of those TRT points that have been confirmed during the left right ambiguity solving are used during the remaining pattern recognition process for this skeleton candidate. From here on there is only one hit per TRT layer in the data structure and the algorithm has effectively found a section of a track in the TRT that is compatible with the skeleton. However the TRT points are only added to the track when the pattern recognition in the discrete layers is completed and it does not add to the number of holes on a track.
The skeleton track is now built into a track 'segment' by associating the measurements in the discrete tracker layers between the skeleton superlayers which are closest to the skeleton centre.
The segment has to satisfy a χ^2 cut of typically 3 p.d.f. its p_t must lie within the skeleton road limits and it must not share more than half of it's discrete hits with a previously found track. If the segment fails the χ^2 cut the point contributing most to the χ^2 is dropped and the χ^2 is tested again.
If the segment has passed the above cuts it defines a very narrow road for space point

[1] p.d.f = per degree of freedom

finding in the high occupancy inner superlayers. Their points complete the segment. Provided that the number of holes in the completed segment is smaller than that of the best completed segment so far it is kept.
Finally a track is completed by interpolating into all superlayers that have not contributed to the completed skeleton. Space points closest to the interpolation center in these superlayers are picked up if they are inside the road and the resultant fit does not have a p_t outside the road limits. In those cases where a superlayer could not contribute a space point but still has a single coordinate measurement[2] this measurement can be picked up in a second pass of extrapolations.
A quality is calculated for each track which is the sum of the χ^2 p.d.f and the number of holes in the discrete detectors. A final track then has to pass the following cuts:

- At least 2 of the 3 innermost layers must be present
- The sum of holes and hits shared with a previous track from any other skeleton must not exceed half the number of hits (discrete layers only).
- The number of holes on the track must be smaller than that of the best track for this skeleton so far (discrete layers only).
- The number of holes must be less than a cut of typically 7 (discrete layers only)
- The number of hits must be larger than a cut of typically 5 (discrete layers only)

The segmentation procedure during the track finding serves to reduce the combinatorial overhead and allows for deviations from a pure helical trajectory due to material effects such as bremsstrahlung and multiple scattering.

7.4 TRT Pattern Recognition Algorithm

This chapter describes the elementary mathematics of the TRT pattern algorithm in section 7.4.1 and the implementation of the method in section 7.4.2. The main task of the algorithm is to find a section of track in the TRT which can be associated to the skeleton track found in the discrete tracker elements. It aims to associate points uniquely, trying to solve the ambiguities arising from the drift time measurements in the TRT and associate appropriate errors to each point. If a unique choice of drift time direction is not possible the straw centre position is associated and the error is adjusted to the measured drift time value.

7.4.1 Assumptions

The trajectory of a particle of charge q, transverse momentum P and impact parameter b inside a homogeneous magnetic field parallel to Z is described by the following equations

$$\Phi = \Phi_0 + \frac{R}{2 \cdot R_{curv}} + \frac{b}{R} \tag{7.1}$$

$$Z = R \cdot \tan\left(\frac{\pi}{2} - \Theta\right) + Z_o \tag{7.2}$$

[2]Space point finding can fail if one coordinate measurement has a large uncertainty due to a wide cluster or if the matching coordinate measurement was inefficient.

where the radius of curvature is given by:

$$R_{curv}[m] = \frac{P_t[\,\mathrm{GeV}]}{0.3 \cdot q[\mathrm{e}] \cdot B[\mathrm{T}]} \tag{7.3}$$

using the usual definitions of the coordinates R, Φ, Θ.

The pattern recognition in the TRT is based on the following assumption:

1. At the radii of the TRT (63-107 cm) the contribution of the impact parameter term in equation 7.1 is small compared to the P_T dependent term and equation 7.4 holds.

$$\Phi \approx \Phi_0 + \frac{R}{2 \cdot R_{curv}} \tag{7.4}$$

   This is true for all particles with impact parameters smaller than 1 cm, down to transverse momenta of 0.5 GeV at B=2T.

2. The skeleton road provides a valid estimate of the true track momentum. Here valid means that the skeleton road must contain the true track.

3. The drift time ambiguities within the measurement of the ϕ coordinate are treated as two possible values ϕ_l and ϕ_r at the same radius where:
   $\phi_r = \phi_{straw} - L_{drift}/R_{straw}$ and
   $\phi_l = \phi_{straw} + L_{drift}/R_{straw}$.
   Here L_{drift} is the drift distance in the TRT straw tube which is calculated from the measured drift time.

The latter is an approximation which is strictly true only for infinite momentum tracks. A correction for low momentum tracks can be easily implemented. Instead of changing the radius of the measurement an effective drift distance can be calculated which represents the intersection of the particle trajectory with the cylinder of radius of the straw center (R_{straw}). The new drift time d' can be expressed as:

$$d' = \frac{d}{\sqrt{1 - \left(\frac{R_{straw}}{2 \cdot R_{curv}}\right)^2}} \tag{7.5}$$

For the lowest momentum tracks considered in this study (2 GeV) this correction changes the drift distance value of a straw at 1m radius by only 1.1%. Even for the biggest possible drift distance of 2mm this corresponds to only 22μm being only 14% of the drift distance resolution. In view of the smallness of this corrections an implementation in *iPatRec* was not considered necessary.

7.4.2 Implementation

For a given skeleton momentum P_t^{skel} a Φ' is calculated for each hit and both drift time alternatives:

$$\Phi' = \Phi - \left(\frac{1}{P_t^{skel}} \pm n \cdot \delta_{\frac{1}{P_t^{skel}}}\right) \cdot R \cdot C \approx \Phi_0 \tag{7.6}$$

$$C = \frac{0.3 \cdot q \cdot B}{2} \tag{7.7}$$

Where $\delta_{\frac{1}{P_t}}$ is typically $0.004\,\mathrm{GeV}^{-1}$ and n is in the range [0,8] for P_t=1 GeV falling to [0,2] for $P_t \geq 12\,\mathrm{GeV}$.

The values of Φ' are histogrammed for different values of n until a peak is found in the histogram or the peak search is abandoned. A peak is defined as the biggest sum of two adjacent bins if it has more than 12 hits. The bin width is 0.5 mrad.
If a peak is found on the edge of the P_t range the central value of the search is shifted towards this edge and the search is continued until a peak is found away from the edge or no peak is found.
The values of $R \cdot \Phi$ for all points falling into the histogram peak bins are then fitted to a straight line and points with a distance of more than 200 μm from the central fit value are discarded. The fit is repeated with the remaining points only. Now any hit in the TRT within 400 μm of the new central fit value is added to the track. The re-fit is done using $R \cdot \Phi$ rather than Φ because the measurement error of the TRT in $R \cdot \Phi$ is constant with radius, whereas the errors on Φ increase as the radius decreases. This leads to an inaccuracy of the histogramming method for small radii.
The errors assigned to each hit inside the final selection band depend on the uniqueness of the drift time assignment. The errors for the $R \cdot \phi$ coordinate are:

- The drift time resolution of 160 μm if the hit is uniquely selected, i.e. the other drift time value was outside the final selection band.
- $\sigma_{R\cdot\phi} = \sqrt{(160\mu\mathrm{m})^2 + L^2_{drift}}$ for hits which are not uniquely selected.

7.4.3 Hit Distributions

As described earlier in section 7.3 the cuts on the required number of hits inside the TRT are a strong function of η. They reflect the overall tracker layout optimisation and have been improved by the latest designs. These cuts will have to be retuned when the detector layout is finalised. With a more final detector layout it is straight forward to tune these cuts more tightly to the hit distributions. This was considered unnecessary because the layout still differs substantially from a possible final version. Figure 7.1 shows the distribution of hits and confirmed hits for 500 GeV muons as obtained without any cuts. They also indicate the position of the cuts as a function of η. Comparing figure 7.1 with the eta dependence of the momentum resolution as shown in figure 7.3 it can be seen that the decrease of resolution power in the region of $1.3 \leq \eta \leq 2.0$ can not be explained by loss of TRT confirmed hits but probably results from the loss of the last silicon strip barrel measurement.

7.5 Results

The analysis described in this chapter has been done for muons only but will be extended to electrons with the same parameter range. Certain performance variables were evaluated over the spectrum of processed data given in table 7.1. The performance variables given in this chapter are defined below.

- Signal track efficiency ϵ_t is defined as:
  The ratio of the number of reconstructed events over the number of events with valid seeds, where a valid seed can be defined in two ways; either any seed with

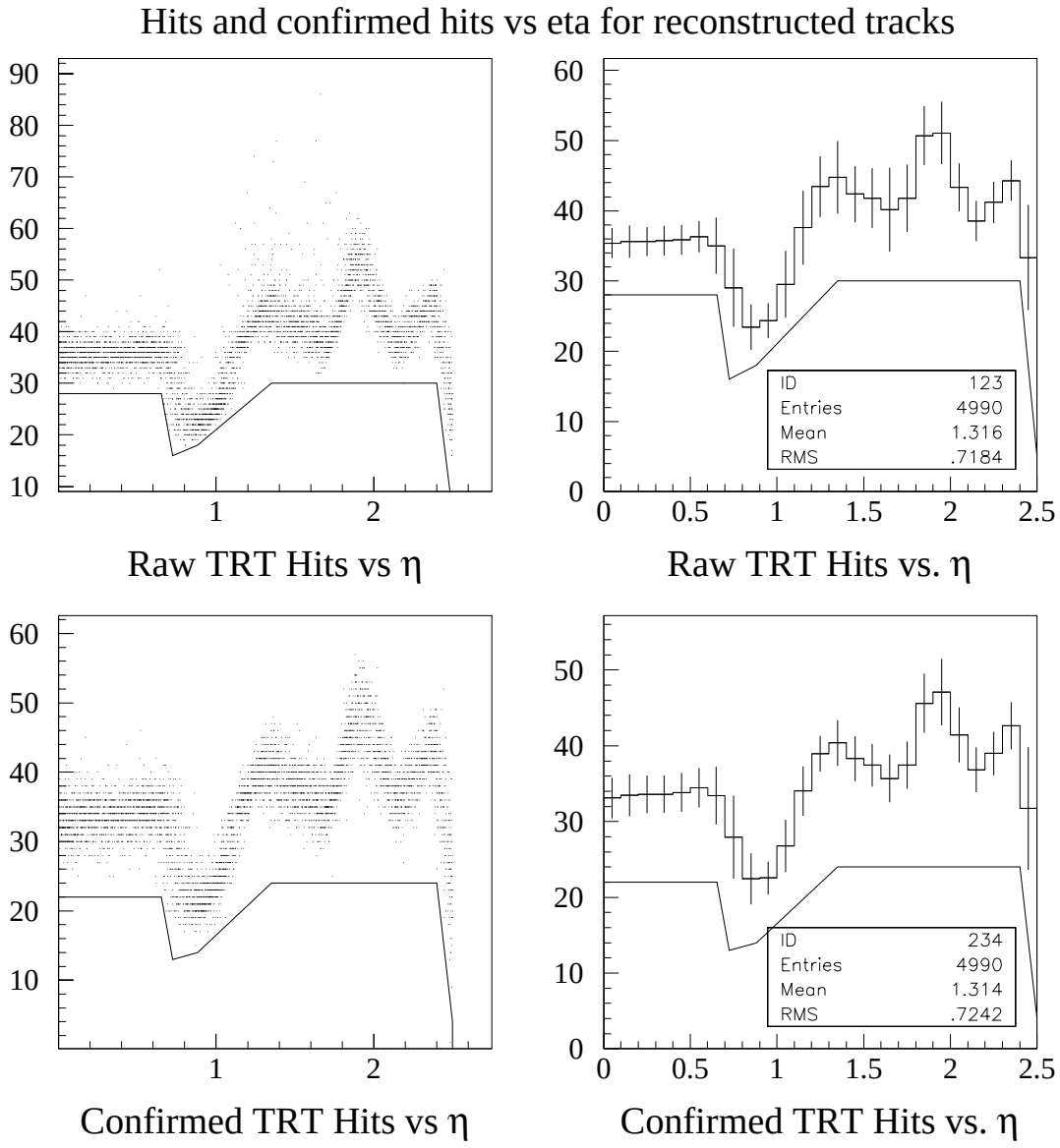


Figure 7.1: Number of TRT-hits in the skeleton road and number of confirmed hits in the fit of reconstructed muon tracks of 500 GeV p_t. The right half is a profile histogram of the left half showing the mean value of the distributions. The error bars are the RMS of the distribution on the left half. No pile-up was added and no cuts on the number of hits were applied. The cut function used for later analysis is also shown

L$[cm^{-2}s^{-1}]$	500 GeV	100 GeV	40 GeV	10 GeV	2 GeV
0.0	5K,μ	5K,μ	5K,μ	5K,μ	5K,μ
$0.29 \cdot 10^{34}$	1K,μ	0	0	0	0
$0.57 \cdot 10^{34}$	1K,μ	0	0	0	0
$1.00 \cdot 10^{34}$	3.96K,μ	1K,μ	1K,μ	1K,μ	1K,μ

Table 7.1: Number of processed events as a function of background luminosity and primary particle momentum. The primary particles were μ's for all events.

non zero momentum, hereafter referred to as a 'loose seed', or a seed matching the momentum of the signal particle[3]. to better than 1%, referred to as a 'tight seed'. An event is reconstructed if at least one track with $p_t \geq 20\,\mathrm{GeV}$[4] is found in an event with at least one valid seed.

- Transverse momentum resolution σ_{P_t} (for vertex constrained fits) is defined as: $\sigma_{P_t} = \sigma(P_t^{true}/P_t^{found} - 1)$ for the best track in events with tight seeds. The best track is that with the lowest quality parameter described in section 7.3.
- Impact parameter resolution σ_{A_0} (for non vertex constraint fits) is given in μm
- Initial azimuthal angle resolution σ_{ϕ_0} (for vertex constraint fits) is given in rad.
- Initial polar angle resolution σ_{θ_0} (for vertex constraint fits) is given in rad.
- Resolution of the Z coordinate of the particle vertex σ_{Z_v} (for vertex constraint fits) given in cm.

[3]The signal particle is the particle entering the reconstruction via the primary input stream of SLUG. In this study signal tracks were single muons of fixed p_t.

[4]For signal tracks of p_t=10 GeV and p_t=2 GeV this threshold was lowered to 1.5 GeV

7.5.1 Resolutions

All five track parameter resolutions have been studied in this work with emphasis on momentum- σ_{P_t} and impact parameter resolution σ_{A_0} . These are the most important parameters for many physics studies and are discussed first. All results in this subsection have been obtained using vertex constraint fits, except for σ_{A_0} . This was done because the impact parameter resolution with vertex constraint can not be used in cases where the particle does not come from the primary vertex, since it is drastically better than the resolution without vertex constraint. This is not the case for any other track parameter, as long as the actual vertex is not more than $\mathcal{O}(1\text{cm})$ displaced.

Momentum Resolution

Figure 7.2 shows the transverse momentum resolution as a function of η for single muon tracks. The resolution for 500 GeV tracks crosses the 30% line at η=2.3. A degradation of the resolution in the region of $1.3 \leq \eta \leq 2.0$ can also be seen. This loss of resolution is present down to a p_t of 40 GeV. At lower p_t multiple scattering dominates this detector effect.

The influence of luminosity on the resolution as a function of η can be seen in figures 7.3 to 7.5. The resolutions are significantly better than those quoted in [5] for two reasons:

- All points found by the TRT pattern recognition have been added to the fit, whereas only one master point was used to represent the TRT hits in [5]
- Perfect alignment of the subdetectors in the inner trackers has been assumed.

The biggest degradation of resolution appears in the region of $1.3 \leq \eta \leq 2.0$ where the two 500 GeV samples differ by up to 6% corresponding to a relative difference of 30% or a 6 σ separation. The degradation becomes less severe as the momentum decreases and is only marginal in the 10 GeV sample.

In trying to determine whether the degradation is due to pattern recognition problems or is fundamentally limited by the detector performance, it is useful to check whether the momentum resolution remains Gaussian. Figure 7.6 shows the fraction of events with values of $\sigma(P_t^{true}/P_t^{found} - 1) \geq 3 \cdot \sigma$. It is important to restrict the sample to a detector region where the resolution does not vary with η. This is why figure 7.6 has been restricted to $|\eta| \leq 1$. However no significant increase of the fraction of events outside 3σ is observed. Keeping in mind that in principle a Gaussian fit to $\sigma(P_t^{true}/P_t^{found} - 1)$ is only strictly valid for events from the same detector position figures 7.7 to 7.8 show the energy dependence of the momentum resolution degradation for samples including events from all eta and $|\eta| \leq 1.0$. Although the absolute values of the resolution for the all η sample are significantly higher, since they include the poorer performing regions of large η, figure 7.8 shows that the influence of luminosity on the resolution is estimated similarly by both samples.

For 500 GeV tracks figure 7.9 shows an increase inχ^2 per degree of freedom of 0.090 due to pile-up. The values for other momenta can be seen in table 7.2. The fact that the average χ^2 is less than 1 shows that the errors associated to the detector measurements are partially overestimated. Assuming an average TRT occupancy of 20% at a luminosity of $1.0 \cdot 10^{34} cm^{-2} s^{-1}$ [4] and an average of 36 TRT hits per track, 7.2 pile-up hits per track are expected. In 50% of all cases the pile-up hits will have a smaller drift time than the signal track hits and therefore overshadow them. If the centroid of the TRT hit histogram is still a good estimator for the true track direction then the cutoff of 400 microns around

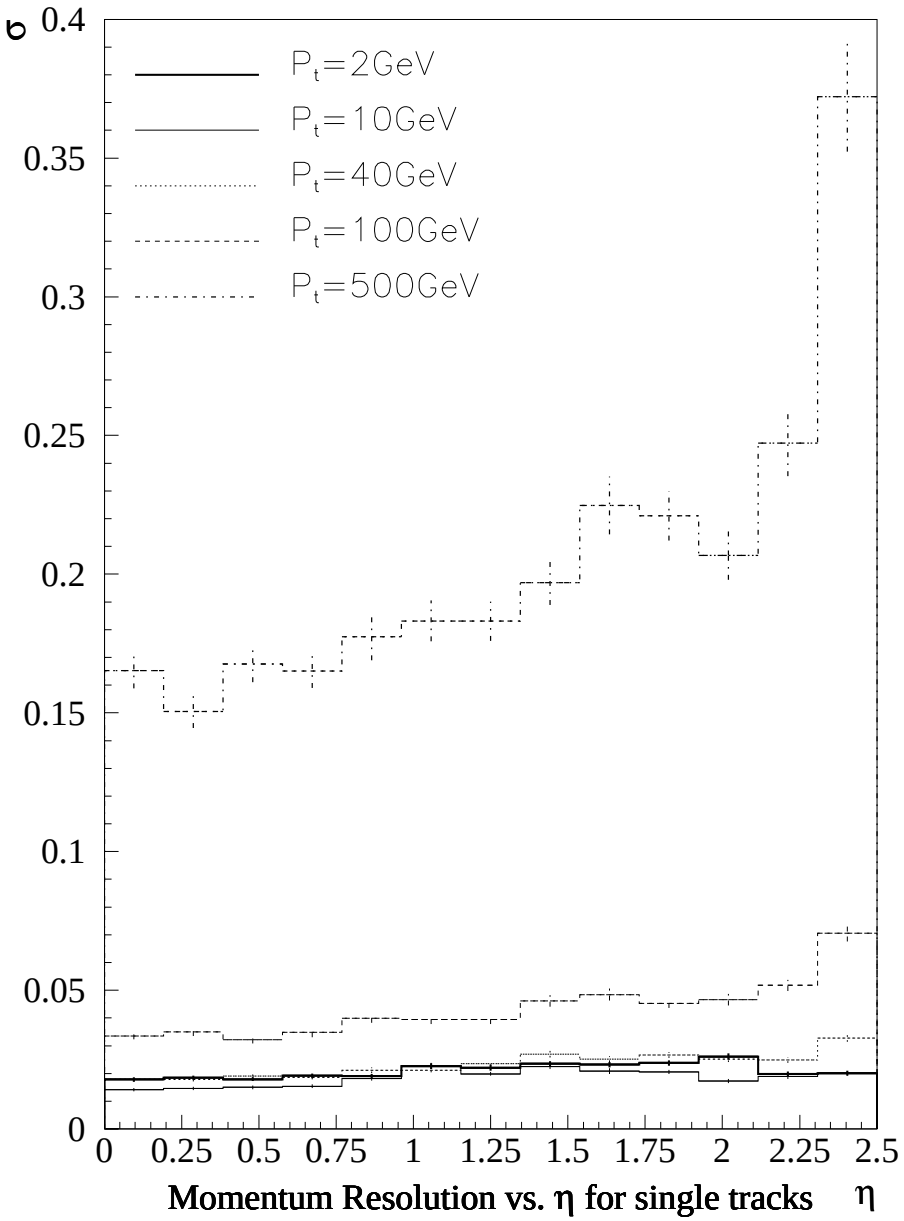


Figure 7.2: $\sigma(P_t^{true}/P_t^{found} - 1)$ as a function of η. The values are taken from Gaussian fits in each eta bin. The fit has been restricted to approximately $\pm 3\sigma_{P_t}$ around the centre of the distribution. The bin width in η is 0.19. Each momentum sample represents 5000 tracks.

L$[cm^{-2}s^{-1}]$	500 GeV	100 GeV	40 GeV	10 GeV	2 GeV
0.00	0.7186	0.6895	0.6767	0.6662	0.7554
$0.29 \cdot 10^{34}$	0.7554				
$0.57 \cdot 10^{34}$	0.7817				
$1.00 \cdot 10^{34}$	0.8090	0.7823	0.7798	0.7777	0.7817

Table 7.2: χ^2 per degree of freedom for muons of various momenta with and without pile-up

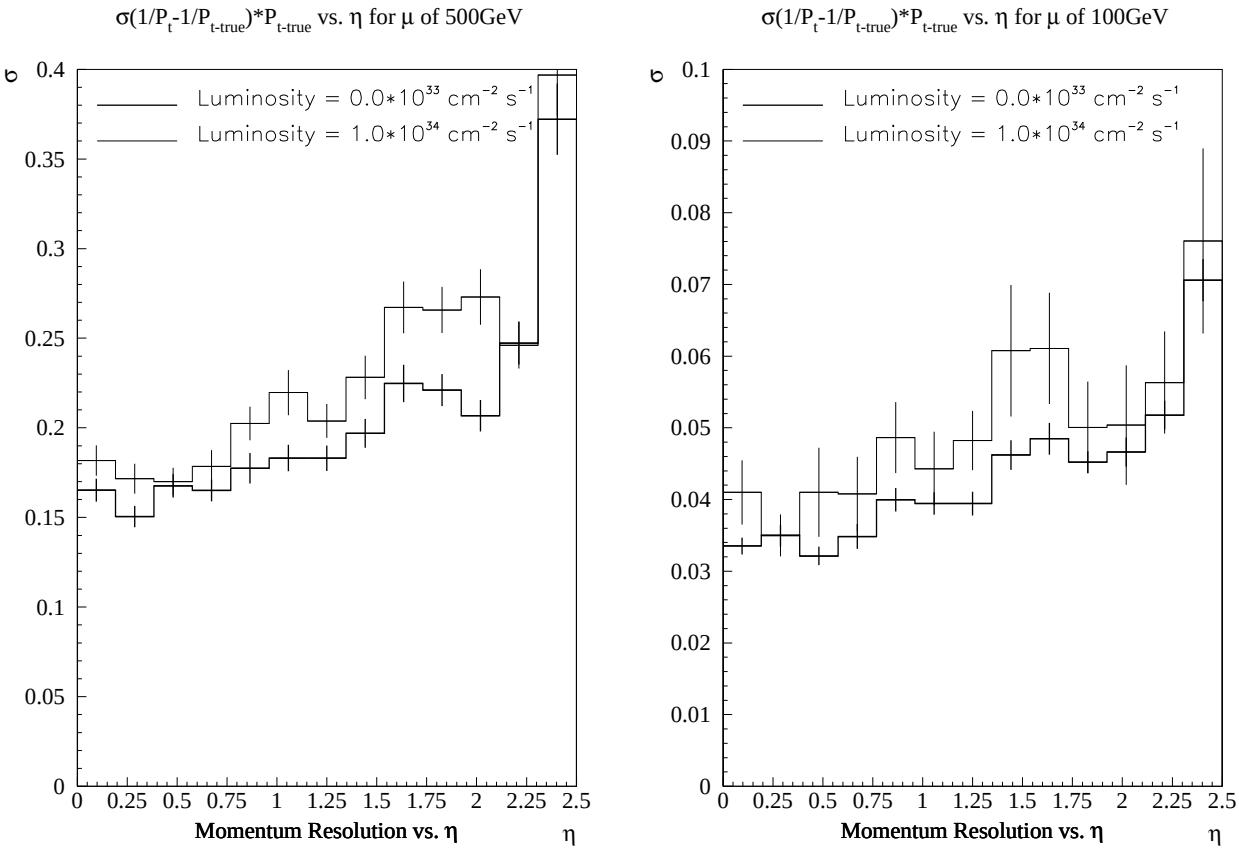


Figure 7.3: $\sigma(P_t^{true}/P_t^{found} - 1)$ as a function of η for 500 and 100 GeV muons with and without pile-up corresponding to a luminosity of $1.0 \cdot 10^{34} cm^{-2} s^{-1}$. The bin width in η is 0.19. At high luminosity the samples represent 3965 tracks at 500 GeV and 956 tracks at 100 GeV.

the centroid will limit the χ^2 contribution of each pile-up hit to 2.08.[5] This would give a total χ^2 contribution from pile-up of 7.5. On tracks with an average of 79 degrees[6] of freedom this leads to an increase in χ^2 per degree of freedom of 0.095. This result is compatible with measured increase of 0.09.

Impact Parameter Resolution

All results shown in this subsection have been obtained from none vertex constraint fits. The impact parameter resolution has been determined as a function of η and luminosity and is shown in figures 7.10 to 7.12 for muons of 2 to 500 GeV. The degradation of σ_{A_0} due to luminosity is not significant when compared to the errors. In the central rapidity bin the resolution is of the order of 30 μm for 500 GeV muons. The resolution is clearly multiple scattering dominated for the 2 GeV sample where it is around 100 μm at η=0. The results agree with those shown in [6] for low p_t over all η and for high p_t at η=0. The high p_t samples show a stronger deterioration of the resolution with η than shown in [6]. The resolution has been parameterised as function of p_t and η for $p_t \geq$10 GeV as follows:

[5] A box distribution of width 800 microns with $RMS = 800\mu\text{m}/\sqrt{(12)} \approx 230\mu\text{m}$ is assumed and normalized to the single straw measurement accuracy of 80 μm .

[6] Per track 4 space point like measurements with 3 coordinates and 36 TRT measurements with 2 coordinates are assumed

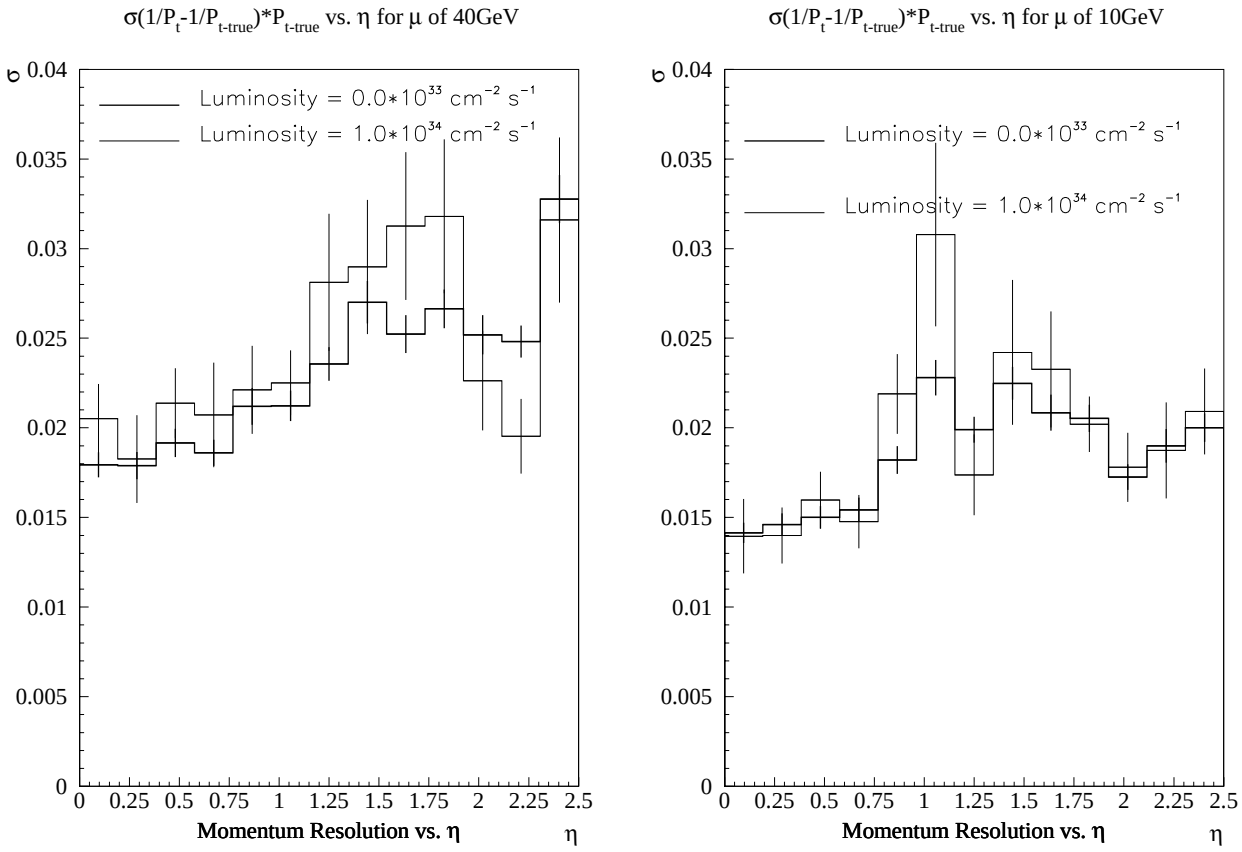


Figure 7.4: $\sigma(P_t^{true}/P_t^{found} - 1)$ as a function of η for 40 and 10 GeV muons with and without pile-up corresponding to a luminosity of $1.0 \cdot 10^{34} cm^{-2} s^{-1}$. The bin width in η is 0.19. At high luminosity the samples represent 930 tracks at 40 GeV and 935 tracks at 10 GeV.

$$|\eta| \leq \mathbf{1.8}: \quad \sigma_{A_0} = a(\frac{1}{P_t}) + b(\frac{1}{P_t}) \cdot \eta \tag{7.8}$$

$$|\eta| > \mathbf{1.8}: \quad \sigma_{A_0} = c \tag{7.9}$$

where the parameters a,b and c have been expressed as functions of inverse transverse momentum as:

$$\begin{aligned} &\mathbf{P_t} \geq \mathbf{40}\,\mathrm{GeV}: \\ &\quad a = 27\mu\mathrm{m} \\ &\quad b = 11.65 + \frac{174.8}{P_t[\,\mathrm{GeV}]}\frac{\mu\mathrm{m}}{\mathrm{GeV}^{-1}} \\ &\quad c = 40\mu\mathrm{m} \end{aligned} \tag{7.10}$$

$$\begin{aligned} &\mathbf{10}\,\mathrm{GeV} \leq \mathbf{P_t} < \mathbf{40}\,\mathrm{GeV}: \\ &\quad a = 27 + \left(93.3 \cdot \left(\frac{1}{P_t[\,\mathrm{GeV}]} - \frac{1}{40\,\mathrm{GeV}}\right)\right)\mu\mathrm{m} \\ &\quad b = 11.65 + \frac{174.8}{P_t[\,\mathrm{GeV}]}\frac{\mu\mathrm{m}}{\mathrm{GeV}^{-1}} \\ &\quad c = 40 + \left(346.7 \cdot \left(\frac{1}{P_t[\,\mathrm{GeV}]} - \frac{1}{40\,\mathrm{GeV}}\right)\right)\mu\mathrm{m} \end{aligned} \tag{7.11}$$

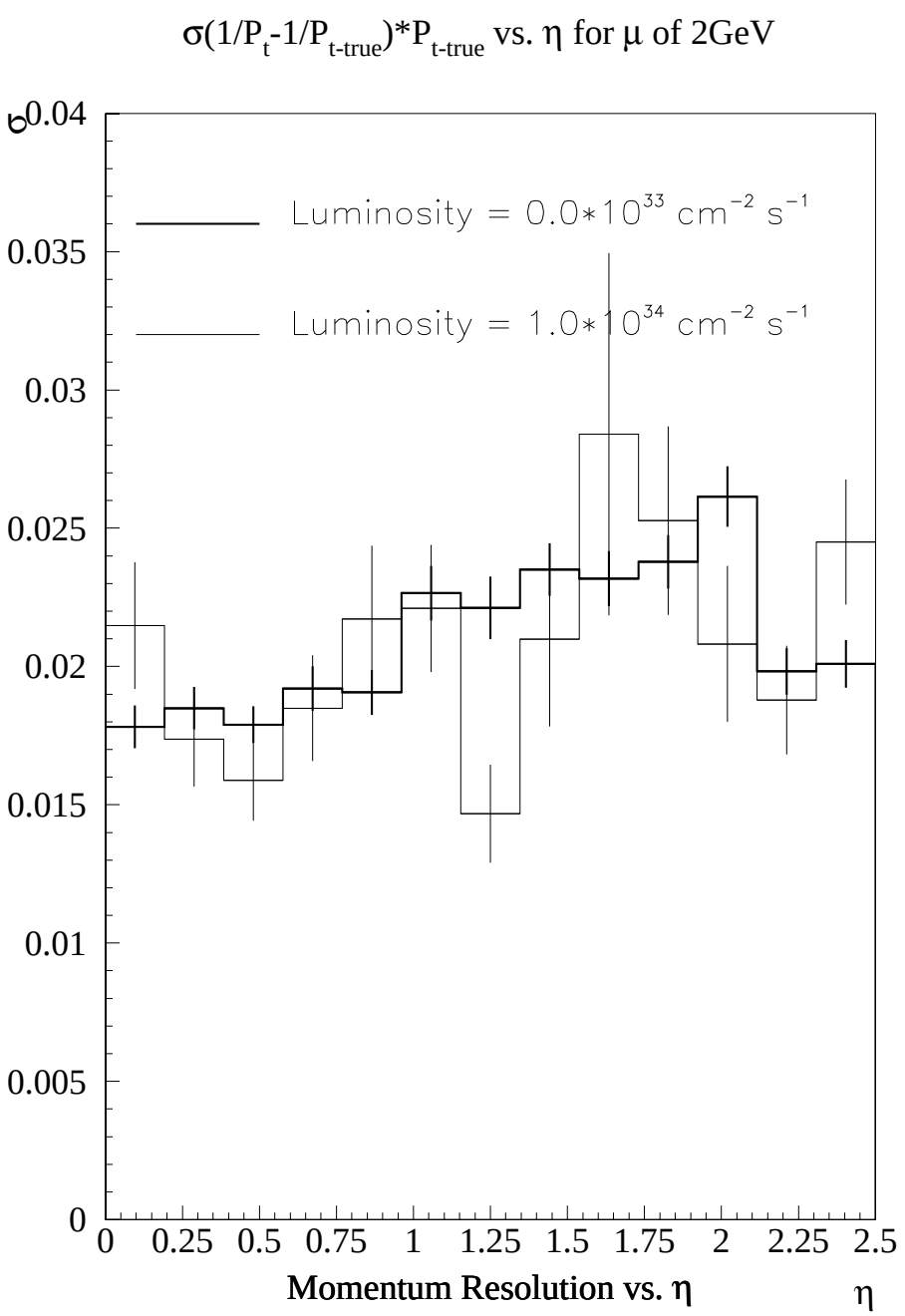


Figure 7.5: $\sigma(P_t^{true}/P_t^{found} - 1)$ as a function of η for 2 GeV muons with and without pile-up corresponding to a luminosity of $1.0 \cdot 10^{34} cm^{-2} s^{-1}$. The bin width in η is 0.19. The high luminosity sample represents 912 tracks.

For $\eta \leq$1.8 this parameterisation represents a linear increase of σ_{A_0} with a common offset of 27 μm for all transverse momenta above 40 GeV. For transverse momenta between 10 and 40 GeV the offset is extrapolated linearly in $\frac{1}{P_t}$ between 34 μm at 10 GeV and 27 μm at 40 GeV.
In the region where η >1.8 the resolution is assumed to be constant with η. Its value decreases linearly in $\frac{1}{P_t}$ from 66 μm at 10 GeV to 40 μm at 40 GeV. For transverse momenta above 40 GeV it remains constant at 40 μm .
The parameterisation has been obtained as a conservative estimate from figures 7.10

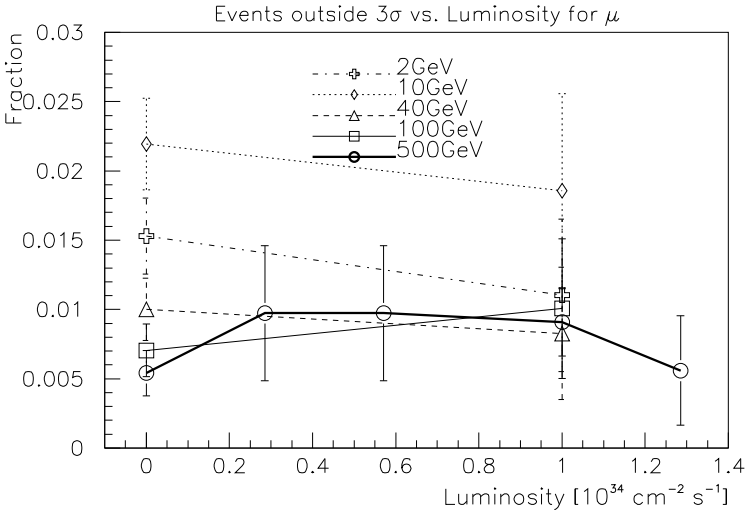


Figure 7.6: Fraction of events with $\sigma(P_t^{true}/P_t^{found} - 1) \geq 3 \cdot \sigma_{P_t}$. Events are restricted to be within $|\eta| \leq 1.0$. The expected value for a perfect Gaussian is 0.27%

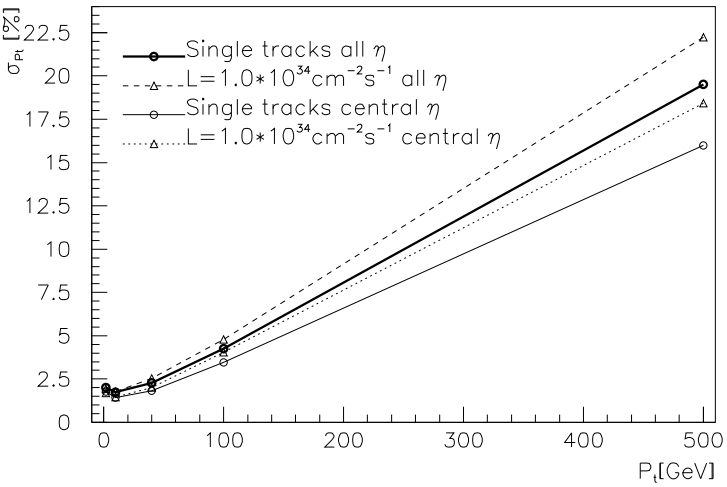


Figure 7.7: Momentum dependence of the momentum resolution σ_{P_t} for single muon tracks with and without pile-up. The lower pair of curves corresponds to a sample restricted to $|\eta| \leq 1$. The higher pair includes events from all η

to 7.12. The p_t-dependence of a and b in equation 7.8 has been obtained from a fit to the offsets and slopes in figures 7.10 to 7.12 in the region of $|\eta| \leq 1.8$. The fit is shown in figure 7.13. It does not contain the points for p_t=2 GeV since these are not relevant for the use of the parametrisation in chapter 8

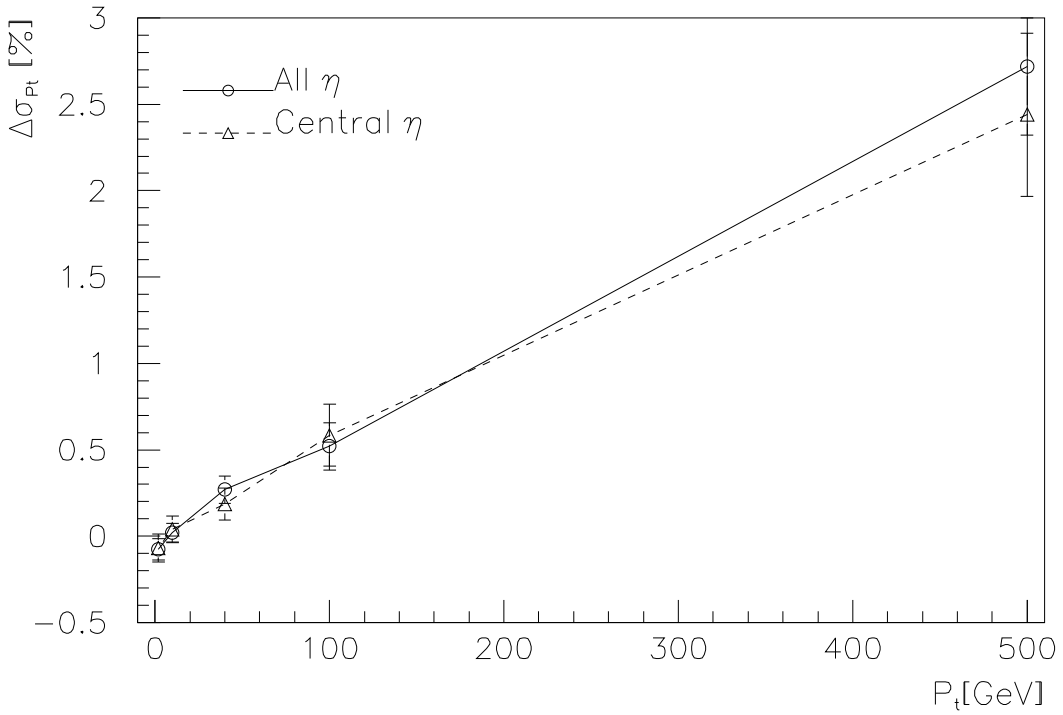
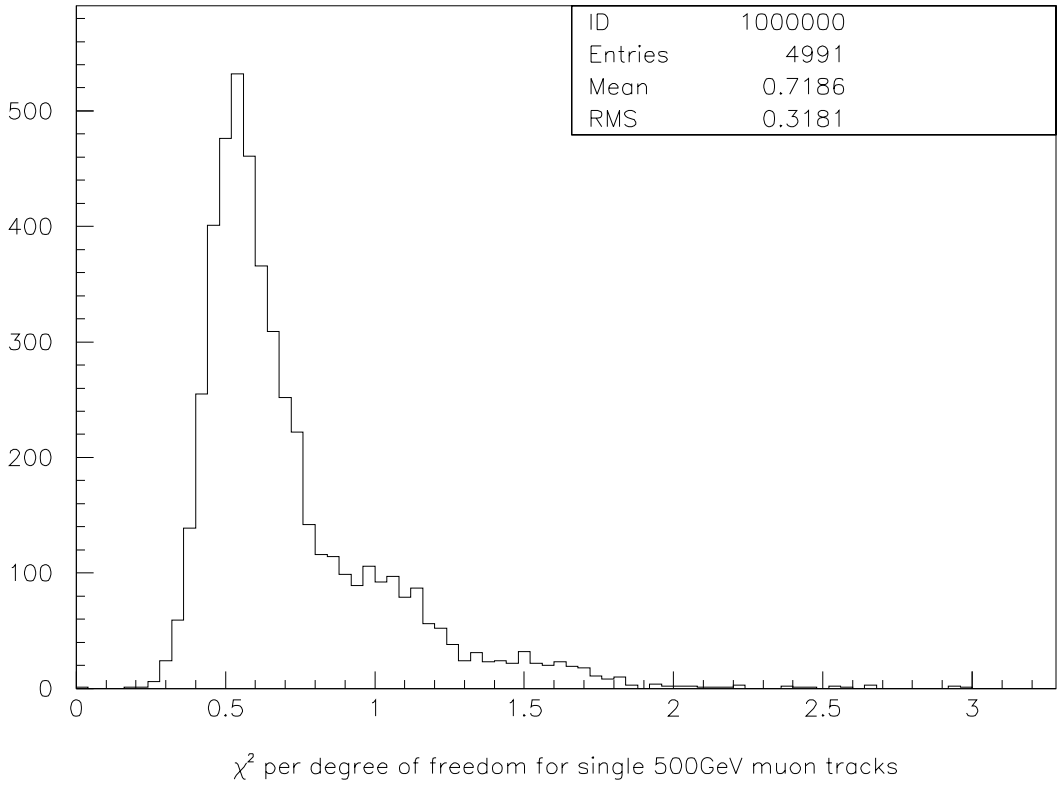


Figure 7.8: Momentum dependence of the degradation of momentum resolution σ_{P_t} through pile-up for muon tracks. The two pairs of curves corresponds to a sample restricted to $|\eta| \leq 1$(central eta) and one for all η.



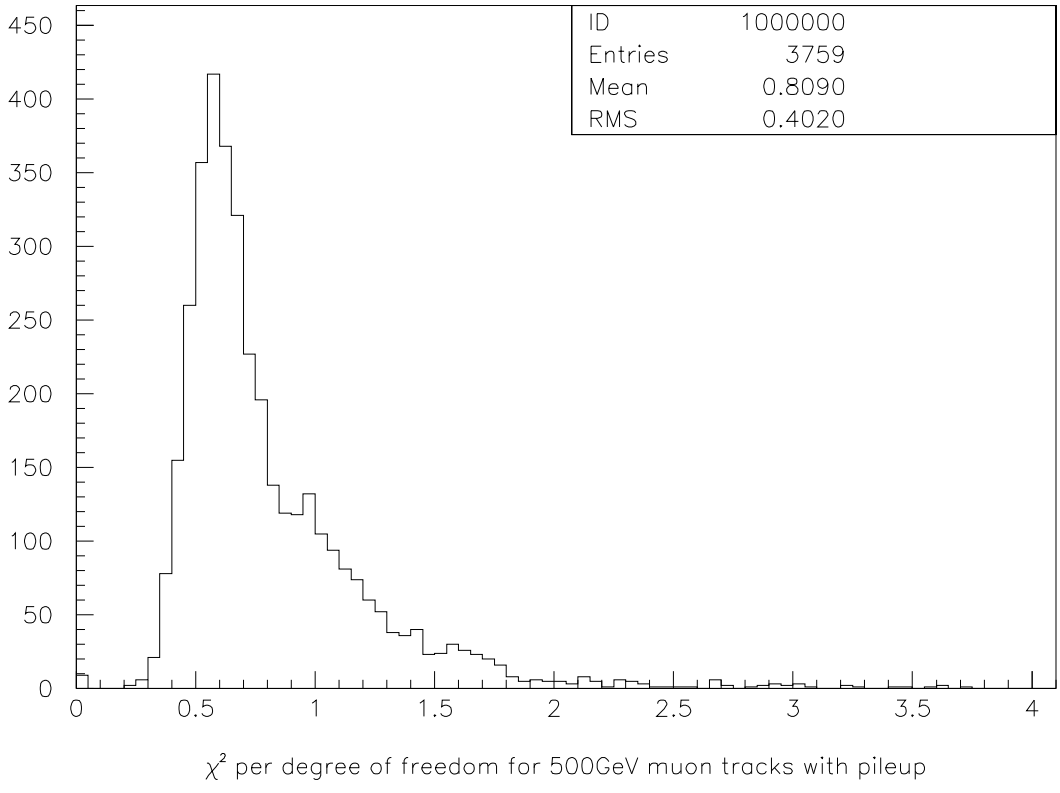


Figure 7.9: Distribution of χ^2 per degree of freedom for vertex constraint fits to 500 GeV muon tracks. The lower distribution is measured at L= $1.0 \cdot cm^{-2}s^{-1}$.

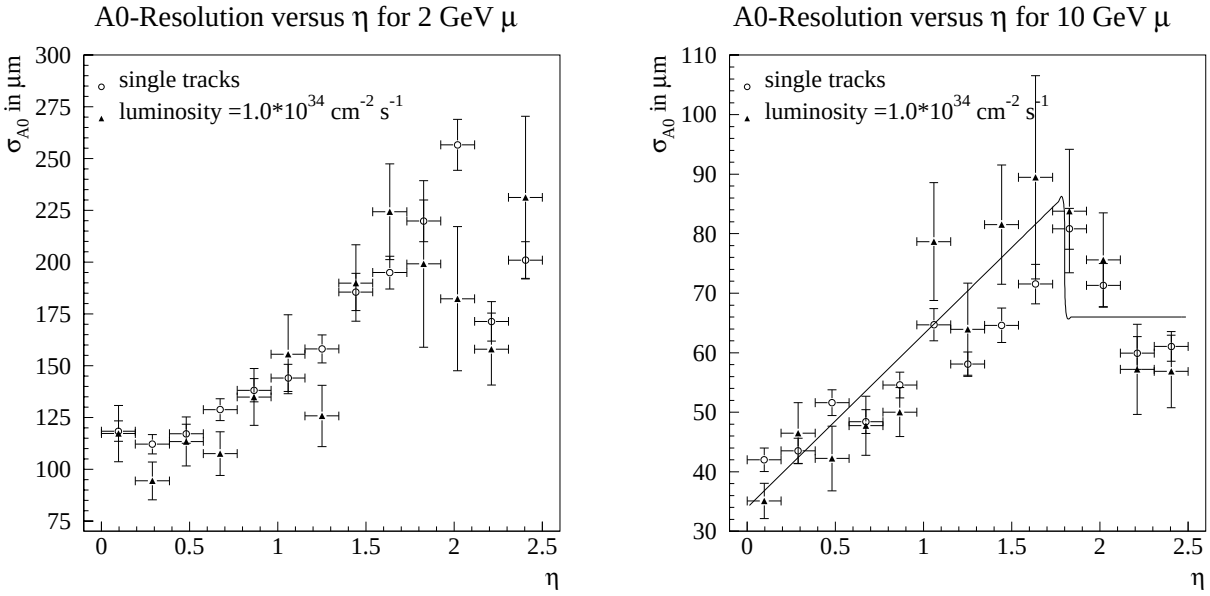


Figure 7.10: Impact Parameter Resolution as a function of η for muons of 2 and 10 GeV transverse momentum. Each plot compares the resolution for single tracks with that for a luminosity of L= $1.0 \cdot cm^{-2}s^{-1}$.

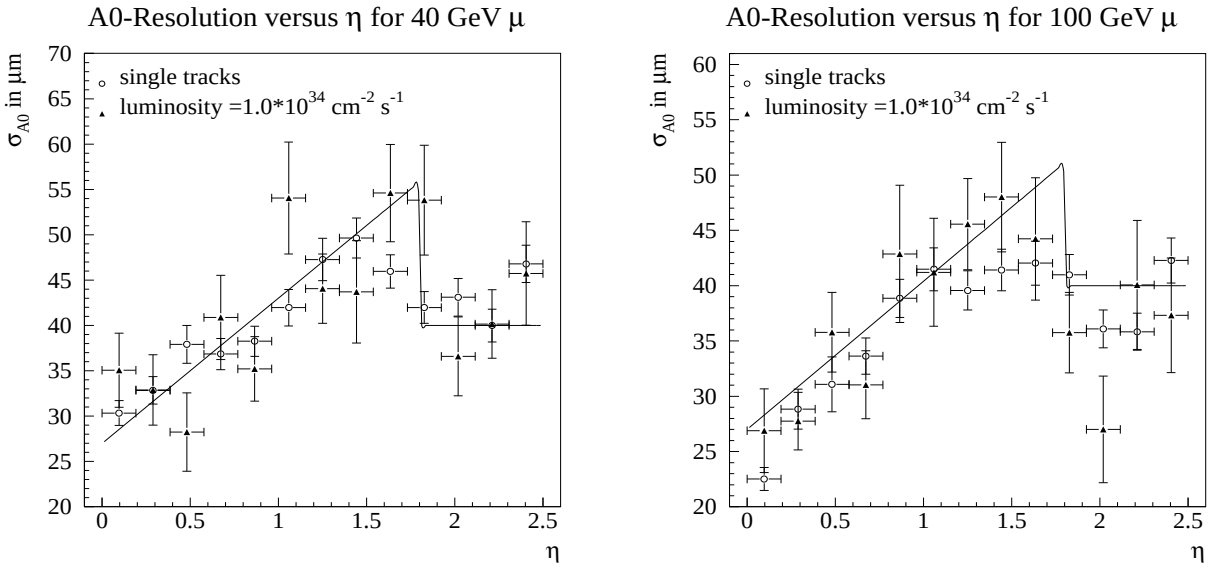


Figure 7.11: Impact Parameter Resolution as a function of η for muons of 40 and 100 GeV transverse momentum. Each plot compares the resolution for single tracks with that for a luminosity of L= $1.0 \cdot cm^{-2}s^{-1}$.

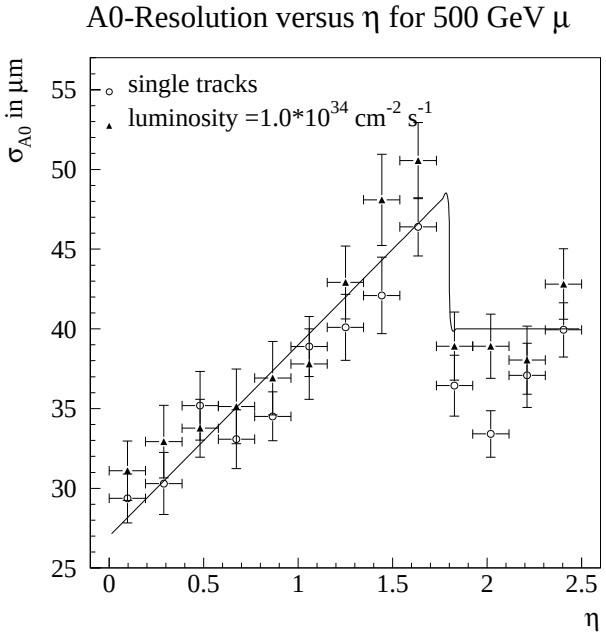


Figure 7.12: Impact Parameter Resolution as a function of η for muons of 500 GeV transverse momentum. Each plot compares the resolution for single tracks with that for a luminosity of L= $1.0 \cdot cm^{-2}s^{-1}$.

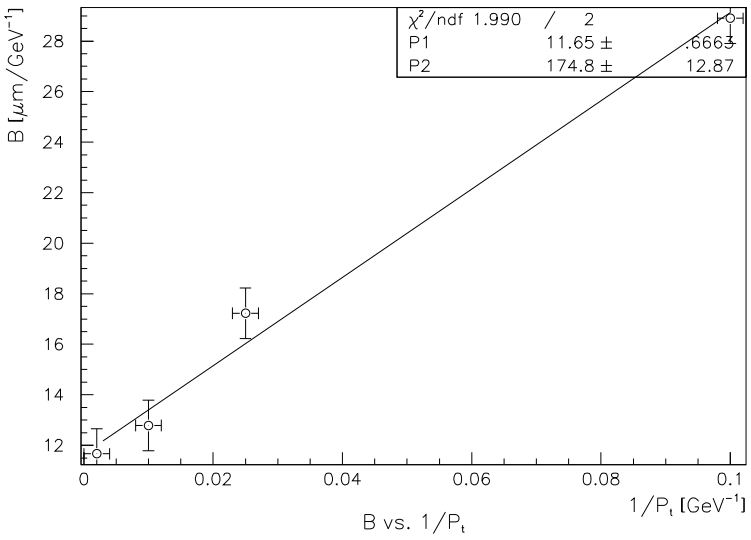


Figure 7.13: Fit to the transverse momentum dependence of the Impact Parameter Resolution slope b as defined in equation 7.8.

Azimuthal Angle Resolution

σ_{ϕ_0} has a strong transverse momentum dependence and is anti correlated with σ_{P_t} . Figures 7.14 to 7.18 show the eta dependence of σ_{ϕ_0} for muons from 500 to 2 GeV transverse momentum. The degradation of the resolution with luminosity decreases with transverse momentum as the multiple scattering increases. At p_t of 2GeV the effect becomes negligible.
The degradation is worst in those eta regions where the resolution depends most on the contribution from the TRT ($1.5 < |\eta| < 2.3$) because that provides many hits at large radii in this region. At the edge of the detector acceptance the resolution is approximately half as good as in the central region.

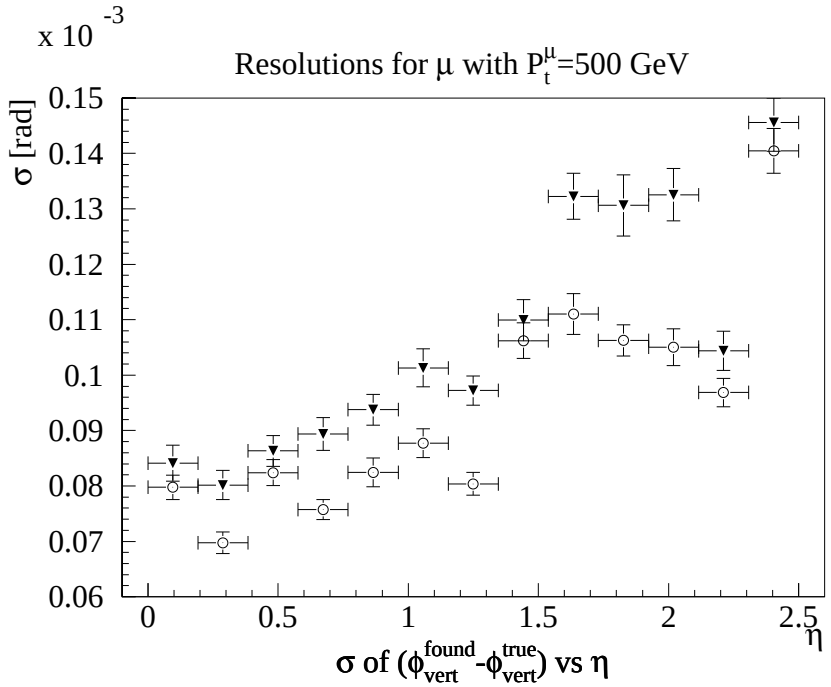


Figure 7.14: Resolutions of the azimuthal angle at the vertex for muons of 500 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

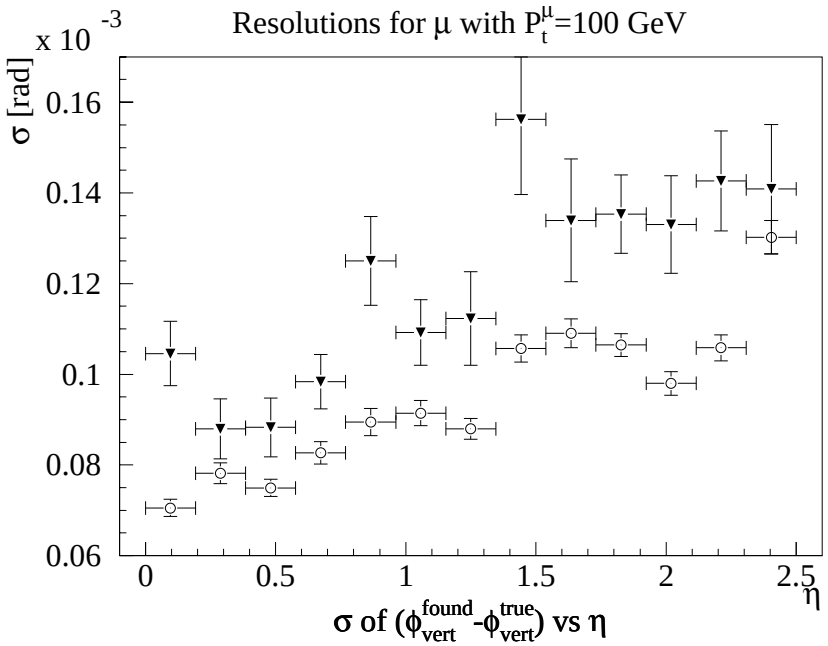


Figure 7.15: Resolutions of the azimuthal angle at the vertex for muons of 100 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

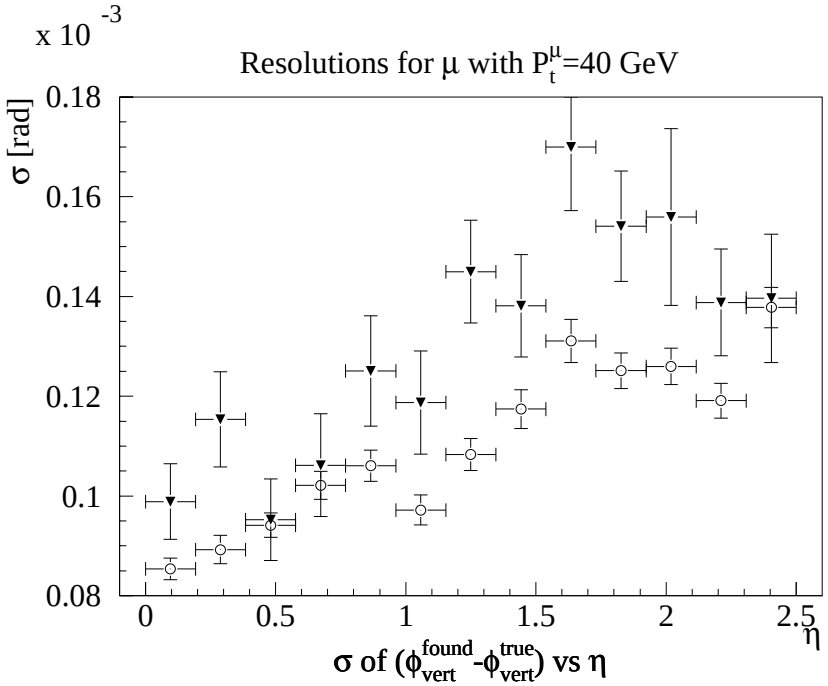


Figure 7.16: Resolutions of the azimuthal angle at the vertex for muons of 40 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

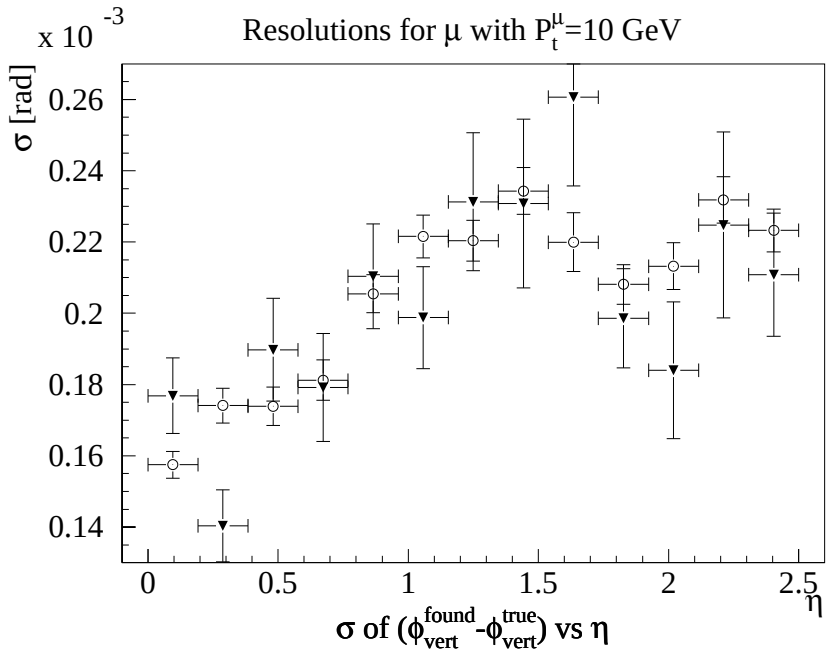


Figure 7.17: Resolutions of the azimuthal angle at the vertex for muons of 10 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

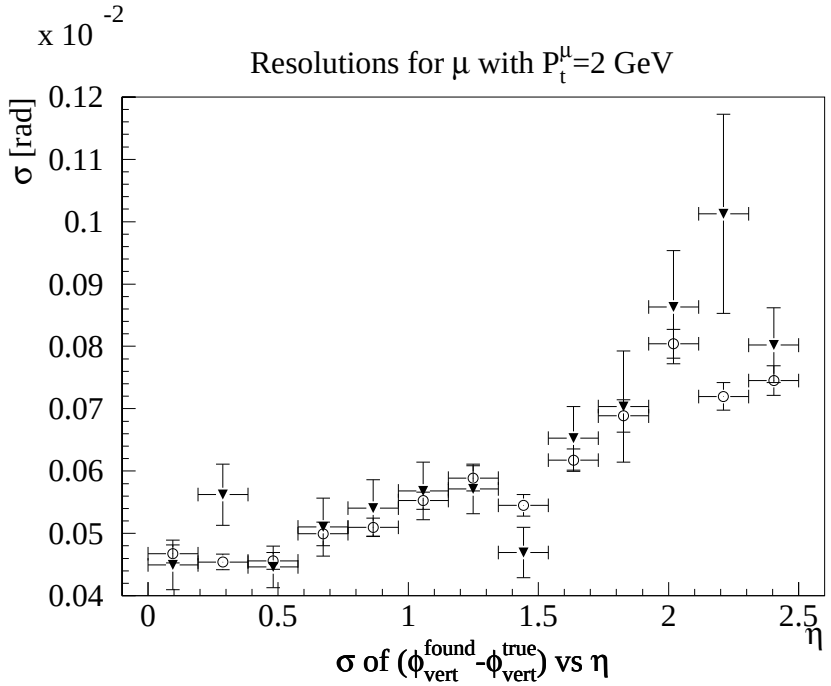


Figure 7.18: Resolutions of the azimuthal angle at the vertex for muons of 2 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

Polar Angle Resolution

The polar angle resolution is independent of transverse momentum as long as it is not dominated by multiple scattering. The 2 GeV samples are clearly multiple scattering limited. Figure 7.19 shows the resolution averaged for transverse momenta of 500, 100, 40 and 10 GeV. When plotted separately for these energies the resolutions do not differ significantly. The improvement of the resolution in the region of $0.9 < |\eta| < 1.7$ is due to the long leaverarm between the measurements from the pixel and MSGC detectors. The luminosity does not affect this resolution since it is not dependent on the TRT measurements. It decreases dramatically as the measurement of the innermost pixel layer is lost reducing the available leaver arm between two measurement. It is recovered as the MSGC wheel and GaAs disks at large Z start to contribute measurements to the tracks. These effects are not present in the 2 GeV sample as is shown in figure 7.20. This is caused by the dominance of multiple scattering.

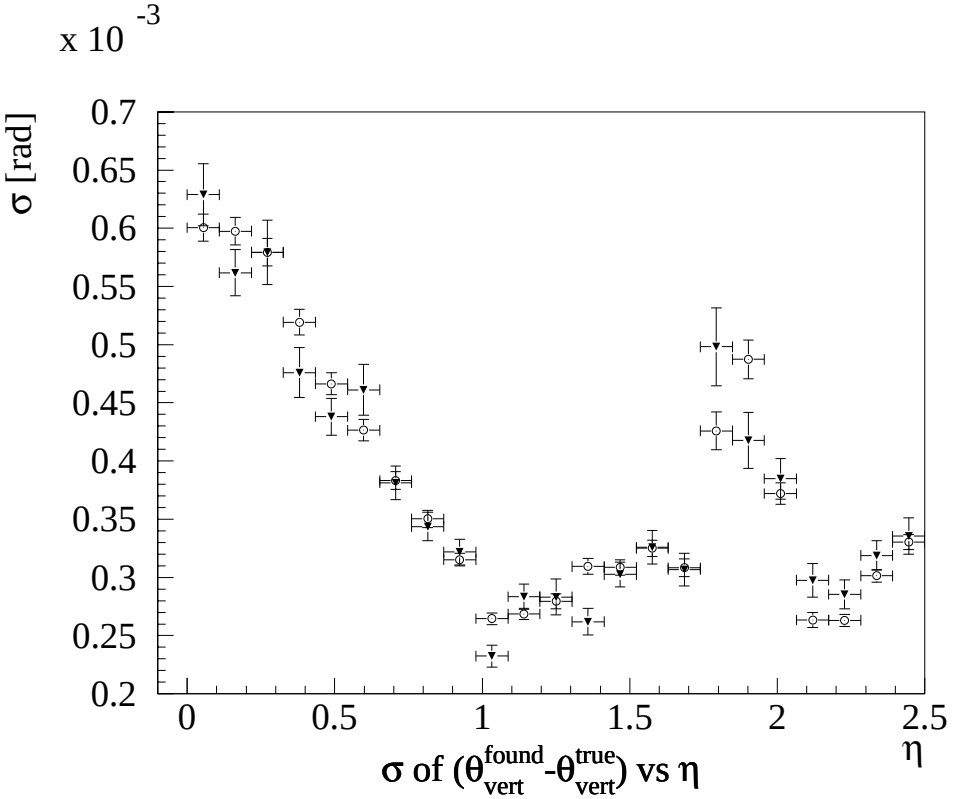


Figure 7.19: Average resolutions of the polar angle at the vertex for muons of 500, 100, 40 and 10 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

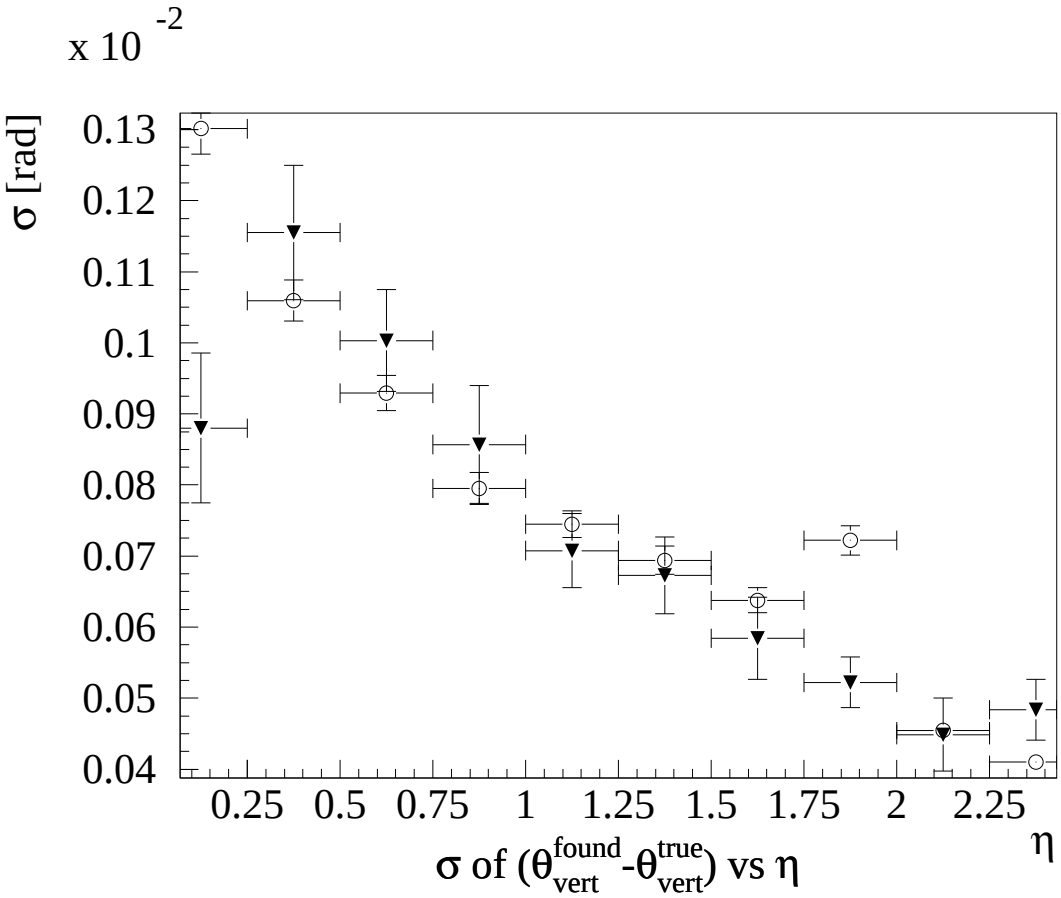


Figure 7.20: Resolution of the polar angle at the vertex for muons of 2 GeV transverse momentum. The solid triangles represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

Z-vertex resolution

The resolution of the vertex Z coordinate has no transverse momentum dependence and is shown in figure 7.21. It is also not influenced by the luminosity since it does not depend on TRT measurements. In the range of $|\eta| < 1.1$ it is extremely flat at a value of 150 μm and then decreases by more than a factor of 10 to over 160 μm .

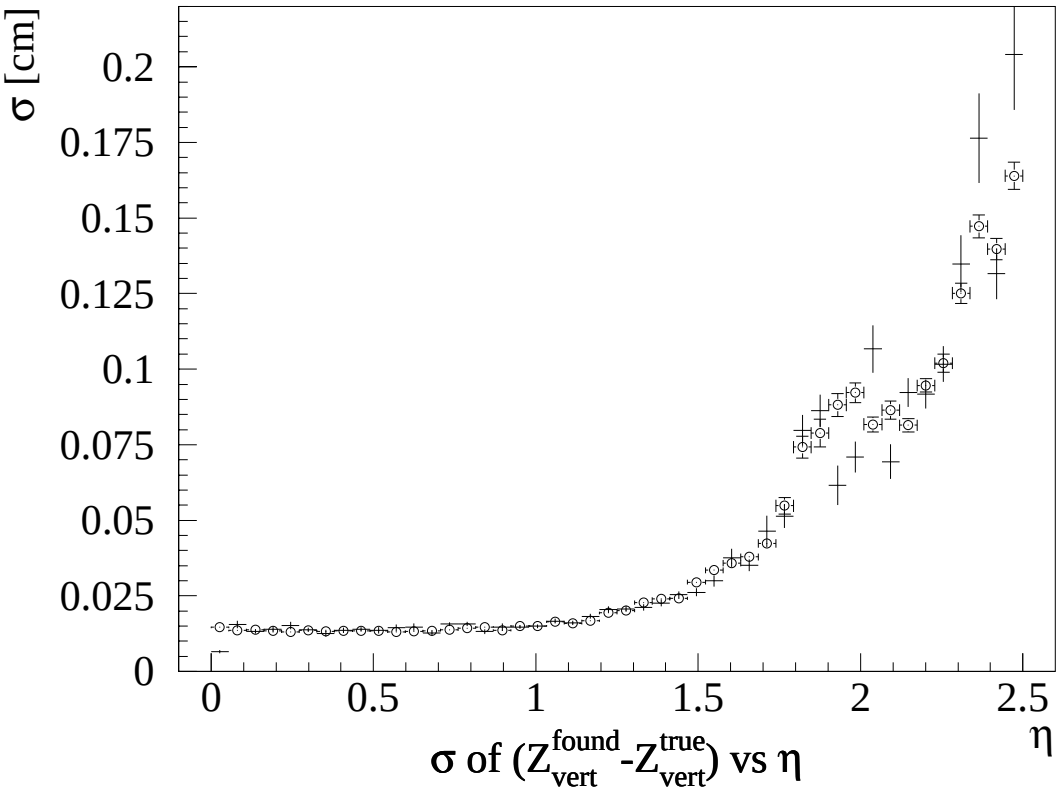


Figure 7.21: Average resolutions of the vertex Z coordinate for muons of 2, 10, 40, 100 and 500 GeV transverse momentum. The plain crosses represent a Luminosity of $1.0 \cdot \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. The open circles are for single tracks.

7.5.2 Efficiency

The track finding efficiency is shown in figure 7.22 for a 'loose seed' selection and in figure 7.23 for a 'tight seed' selection as a function of luminosity and momentum.

The results from the two seed selections differ significantly. The 'loose seed' sample has a much lower efficiency (at high luminosity) than the 'tight seed' sample. It also seems to gain in efficiency for 10 and 2 GeV tracks.

To explain the difference part of the analysis sequence has to be described in some more detail. Even though the true seed particle is always passed to the analysis as the first seed, the best reconstructed track does not have to be the track truly belonging to this seed. Before a track and it's corresponding true value is stored, the KINE values of all tracks in the road are compared to the found track[7] and the best-fit KINE track is stored as the true track. The second and third best track are stored without a separate true value. This procedure is done for two seeds per event. The stored true-track values are later on assumed to represent the seed track.
If the first two seed tracks do not closely match the momentum of the signal track, the event is excluded from the 'tight seed' sample. This procedure preferentially excludes events with high multiplicities around the signal track, as these tend to be those in which pile-up tracks can be selected as the first two truth tracks. This is not a problem for the determination of momentum resolution, but it is problematic for the efficiency calculation.

In the 'loose seed' sample reconstructed pile-up tracks have improved the efficiency at low momenta. This only happens if the track is above the background momentum cut. The 2 and 10 GeV samples have a background cuts of 1.5GeV and there are a large number of pile-up tracks above this threshold. The 40, 100 and 500 GeV samples have background a cut of 20 GeV and are unaffected by this. In the tight seed sample the efficiency is artificially high because many 'problematic' events do not have the signal track seed as the first or second seed. In order to determine the tracking efficiency more precisely more information about the true track content of all reconstructed tracks is needed. This should preferably include a list of all hits and the particles producing them for all reconstructed tracks.

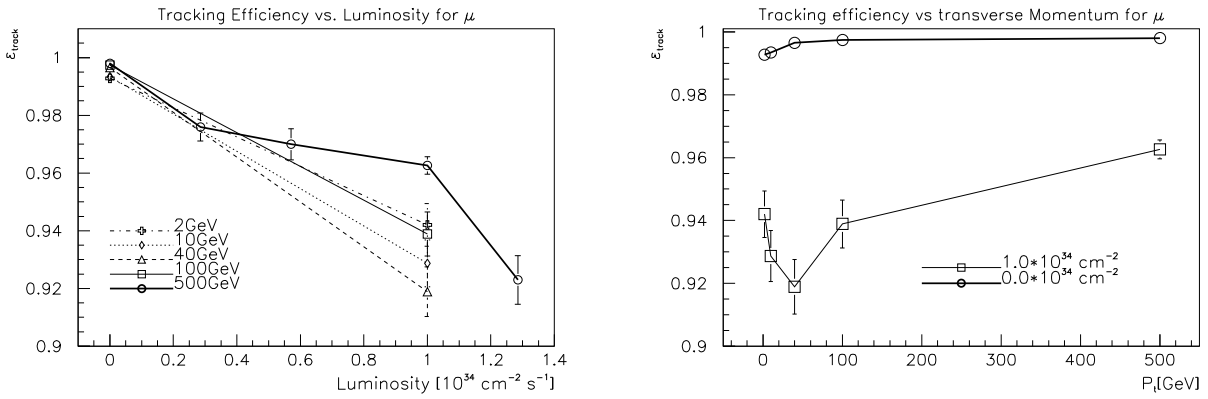


Figure 7.22: Tracking efficiency for muons in the 'loose seed' sample.

[7]The best-fit KINE track is chosen to be the one with the trajectory intersecting the calorimeter front face closest to the best reconstructed track.

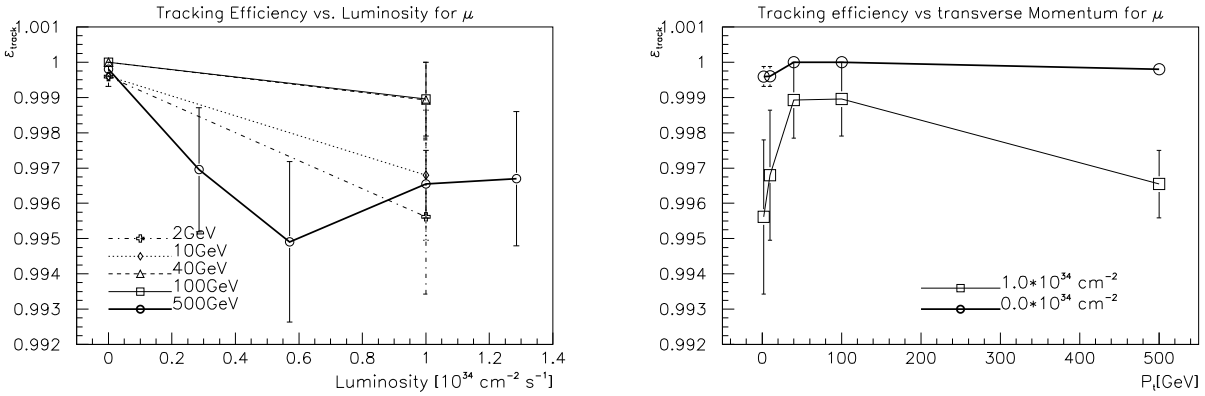


Figure 7.23: Tracking efficiency for muons in the 'tight seed' sample.

7.6 Conclusions

An initial approach to the implementation of a "road driven" pattern recognition in the ATLAS TRT and its integration into the framework of *iPatRec* has been successfully performed. This work was aimed at the improvement of the reconstruction of high p_t muons over the SCT stand alone performance. The targets were:

1. improvement of the momentum and impact parameter resolution.

2. reduction of the ghost track rate at high luminosities.

3. conservation of a high track finding efficiency at high luminosities.

Point 1 has been fully achieved. The transverse momentum resolution is close to its analytically possible limit.
It has been shown, that the resolution of those parameters depending on measurements from the TRT (p_t and ϕ) show a degradation with increasing luminosity that is consistent with the expectations from the reduction in measurement accuracy of the TRT due to hit shadowing through pile-up hits. This degradation vanishes when the resolution becomes dominated by multiple scattering which is the case for both parameters in the 2 and 10 GeV samples and the effects starts to be visible for the momentum resolution in the 40 GeV sample. The other track parameters (θ, Z_{vertex} and A0) are unaffected by the luminosity. The Z_{vertex} resolution shows no multiple scattering effects even at 2 GeV, whereas the θ resolution is clearly multiple scattering limited for the 2 GeV sample.
At high luminosities ghost track rates in the non TRT case were high and have been reduced significantly. A quantitative analysis of this problem is still pending and fake track rates are not given in this work, because a more consistent definition of a fake track for a high track density environment is needed. However it can be said, that fake track rates are at most of the order of 1%.
The track finding efficiency at highest luminosities varies between 92% and 97% for the more conservative definition explained in section 7.5.2. The problems in the definition of the track finding efficiency are closely related to those of the fake track rate and more work is needed to establish a more general definition of this variable.
Currently the algorithms used have been targeted at high p_t tracks only. The use of the TRT for low momentum tracks may have to be considerably different from the current

implementation. In order to extend the current version to the reconstruction of other charged tracks such as pions and electrons more work is needed and initial trials using the TRT in conjunction with a "bremsstrahlung-fit" for electrons show promising results. Another pending task is the adaptation of the algorithm to the latest geometry which is connected with a major portation effort to adopt the program to the latest version of DICE.

Chapter 8

Gluino-Pairs in the ATLAS Inner Detector

The tests of the pattern recognition algorithm presented in the previous chapter were based on single tracks and did not include problems related to the topology of complex physics events, such as hot spots from very collimated jets which are particularly problematic for projective detectors like the TRT. In this sense the work presented in this chapter is a logical continuation of the previous chapter. Additionaly open questions concerning the reconstruction of $\tilde{g}$-pair events are adressed and solved with the use of the pattern recognition algorithm. Furthermore the question of backgrounds internal to the signal events, a subject up to now not studied in any depth is discussed in detail. In particular the problem of semi leptonic heavy quark decays and the use of impact parameter cuts for their supression is presented. The main goals of this chapter are:

- To determine the fraction of same sign lepton pairs from the production of $\tilde{g}$'s and $\tilde{q}$'s that come from those decay chains that allow the identification of the $\tilde{g}$ as a Majorana particle.
- To find ways to distinguish these 'signal' leptons from leptons of other decay chains of $\tilde{g}$'s and $\tilde{q}$'s.
- To test the pattern recognition algorithms described in chapter 7 with realistic, physically relevant events of high multiplicitiy and obtain a more reliable estimate for the tracking efficiency, ghost track rate and charge misidentification probability. It will be shown later that heavy $\tilde{g}$ and $\tilde{q}$ production can lead to multiplicities comparable to pile-up from minimum bias events even at the highest luminosities.

The parameters for the MSSM that were chosen are shown in table 8.1. They are identical to those in [7] and it is not the aim of this chapter to reproduce the results concerning signal visibility and background rejection and their dependance on the choice of MSSM parameters as given in [7].

8.1 Motivation

In all super symmetric models the super partners of gauge bosons are assumed to be majorana fermions. This means that they are at the same time particle and antiparticle. This property is the reason for the fact that for each allowed decay mode the corresponding charge conjugate decay mode is also possible.
In all super symmetric models a new multiplicative quantum number called R-parity is

Parameter	Value	Parameter	Value
$m_{\tilde{q}_{(u,d,s,c),L/R}}$	1.2 TeV	$m_{\tilde{q}_{t,L}}$	1.2 TeV
$m_{\tilde{q}_{t,R}}$	1.2 TeV	$m_{\tilde{q}_{b,L}}$	1.2 TeV
$m_{\tilde{q}_{b,R}}$	1.2 TeV	$m_{\tilde{g}}$	600 GeV
m_A	0.5 TeV	μ	440 GeV
$\tan(\beta)$	2.0	A_t	-100 GeV
m_t	1.7 GeV		

Table 8.1: Input parameters for the MSSM

introduced. It is 1 for all particles and -1 for all super symmetric particles. In a large number of SUSY models R is a conserved quantum number in all interactions. If R is conserved super symmetric particles can only be produced in pairs and the lightest super symmetric particle is necessarily stable. If the $\tilde{g}$ exists as a Majorana particle and R-parity is conserved it is most frequently produced in a $\tilde{g}\tilde{g}$-pair and both $\tilde{g}$'s can decay via the same or similar (similar in this case means that both decays lead to the same sign lepton in the final state) decay channel and create final states with two high p_t, isolated leptons of the same sign. This signature is in general a sign for the decay of a pair of majorana particles, assuming that the leptons come from decays of particles produced in the same interaction and are not cascade decay products. Figure 8.1 shows a Feynman diagram of $\tilde{g}$ pair production via gluon gluon fusion and the decay chain of interest. An event is only considered as a signal event if the leptons (e or μ) from both W's are of the same sign and if there are no other μ's or electrons in the event passing the cuts described in subsection 8.5. The signal from the leptons is accompanied by large missing p_t from 2 escaping $\tilde{\chi}_1^o$'s and 2 ν's as well as a minimum of 4 hard jets from the quarks $\overset{(-)}{q}_1$ to $\overset{(-)}{q}_4$. Even though the majority of $\tilde{g}\tilde{g}$ decays will lead to opposite sign pairs this signal is swamped by opposite sign leptons from pairs of abundantly produced heavy quarks. In a MSSM parameter range where the branching ratio of $\tilde{g} \rightarrow \tilde{\chi}_2^{+(-)}\ q_1 q_2$ is larger, i.e. where the $\tilde{\chi}_{1,2}^{+(-)}$ are a more even mixture of $\tilde{H}^{+(-)}$ and $\tilde{W}^{+(-)}$ and the mass difference $\Delta M_{\tilde{\chi}_{1,2}^{+(-)}} = M_{\tilde{\chi}_2^{+(-)}} - M_{\tilde{\chi}_1^{+(-)}}$ becomes smaller ,the decay products of the two Z's from the decay $\tilde{\chi}_2^{+(-)} \rightarrow \tilde{\chi}_1^{+(-)}$+ Z, can further contribute either 4 jets, 4 leptons or missing p_t. In this case the outstanding signature of two real Z's[1] would make this signal very hard to miss. The two main backgrounds to this signal are listed below.

$t\bar{t} \rightarrow l^+l^- + X$: $t\bar{t}$-decays are the dominant source of opposite sign high p_t leptons. Even though there are no same sign leptons in these events, the large expected cross-section for $t\bar{t}$ of approximately 1 nb would result in 10^8 $t\bar{t}$ pairs per year at $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}and the misidentification of the charge sign of one of the leptons could fake a signal event.

$t\bar{t} \rightarrow b\bar{b} + X_1 \rightarrow \overline{u}cl^+ + X_2 \rightarrow l^+l^+ + X_3$: The cascade decays of b-quarks can lead to same sign leptons, but a large fraction of these are not isolated and do not come from the primary vertex.

[1] as long as $\Delta M_{\tilde{\chi}_{1,2}^{+(-)}} > M_{Z^o}$

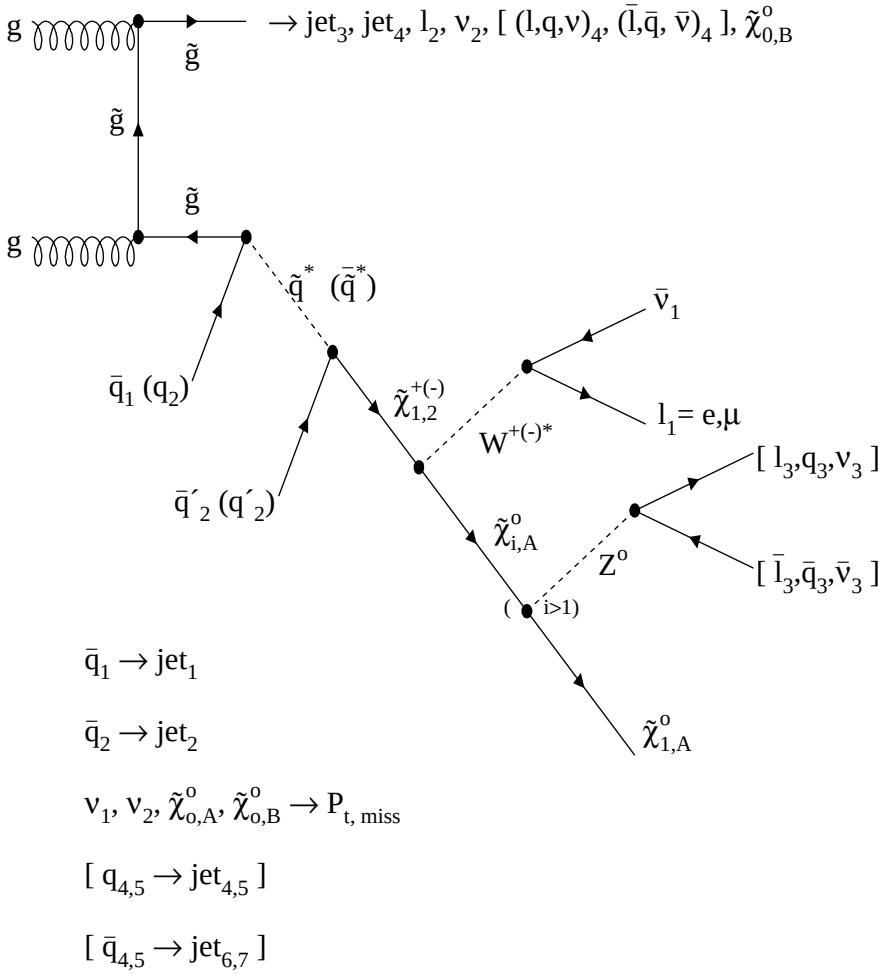


Figure 8.1: Feynman graph of $\tilde{g}$ pair production via gluon gluon fusion and of the consecutive decay chain of the $\tilde{g}$. Same sign leptons from the W's can appear if both $\tilde{g}$'s decay into $\tilde{\chi}^{+(-)}_{1,2}$'s of the same sign. The decay chain of the second $\tilde{g}$ is summarized by its final state particles only. The optional continuation of the decay chain via heavy $\tilde{\chi}^{o}_{i=2,4}$'s is only possible if the $\tilde{g}$ decayed via $\tilde{g} \to \overline{q}q'\tilde{\chi}^{+(-)}_{2}$, because the lightest chargino $\tilde{\chi}^{+(-)}_{1}$ is lighter than the second lightest neutralino $\tilde{\chi}^{o}_{2}$ for the choice of SUSY parameters described in table 8.1. The branching ratio for this process is marginal with these parameters.

8.2 SUSY Parameters and Particle Spectrum

The particle generator used for this study is ISAJET [19]. The specific implementation of the MSSM inside ISAJET is described in [18] and only the major aspects are given here. For the calculations of the decay modes of supersymmetric particles a minimally extended standard model with an SU(5) grand unification constraint is used. The masses of left and right $\tilde{q}_t$ and $\tilde{q}_b$ can be assigned separately all other $\tilde{q}$'s are assumed to be degenerate. Mixing of the left and right handed $\tilde{q}_t$'s leads to mass eigenstates called $\tilde{q}_{t,1}$ and $\tilde{q}_{t,2}$ which are not degenerate, as can be seen in table 8.2. Left and right handed $\tilde{l}$ masses and a single $\tilde{\nu}$ mass are also considered as free parameters. Further more the standard five parameters specifying the SUSY model are:

- $\mu = -2 \cdot M_1$ is the SUSY Higgs mass parameter
- $\tan(\beta)$ is the ratio of the vacuum expectation values for the two Higgs doublets.
- $m_{\tilde{g}}$ is the $\tilde{g}$ mass
- A_t is the stop quark trilinear coupling constant
- m_A is the pseudo scalar Higgs mass

The choice for all input parameters to the model has been given in table 8.1 and the resultant particle spectrum is given in table 8.2.

Particle	Mass[GeV]	Particle	Mass[GeV]	Particle	Mass[GeV]
$\tilde{q}_{u,L}$	1200.000	$\tilde{q}_{d,L}$	1200.000	$\tilde{q}_{s,L}$	1200.000
$\tilde{q}_{c,L}$	1200.000	$\tilde{q}_{b,L}$	1200.000	$\tilde{q}_{t,1}$	1188.802
$\tilde{q}_{u,R}$	1200.000	$\tilde{q}_{d,R}$	1200.000	$\tilde{q}_{s,R}$	1200.000
$\tilde{q}_{c,R}$	1200.000	$\tilde{q}_{b,R}$	1200.000	$\tilde{q}_{t,2}$	1233.716
$\tilde{\nu}_{e,L}$	600.000	$\tilde{e}_L^-$	600.000	$\tilde{\nu}_{\mu,L}$	600.000
$\tilde{\mu}_L^-$	600.000	$\tilde{\nu}_{\tau,L}$	600.000	$\tilde{\tau}_L^-$	600.000
$\tilde{e}_R^-$	600.000	$\tilde{\mu}_R^-$	600.000	$\tilde{\tau}_R^-$	600.000
$\tilde{g}$	600.000	$\tilde{\chi}_1^o$	74.877	$\tilde{\chi}_2^o$	143.986
$\tilde{\chi}_3^o$	441.450	$\tilde{\chi}_4^o$	463.584	$\tilde{\chi}_1^{+(-)}$	142.657
$\tilde{\chi}_2^{+(-)}$	460.366	h^o	100.251	H^o	506.743
A^o	500.000	H^+	505.506		

Table 8.2: MSSM particle spectrum

8.3 Cross-Section and Branching Ratios

For this study the following three processes were generated together.

- $pp \to \tilde{g}\tilde{g}$
- $pp \to \tilde{g}\tilde{q}, \quad \tilde{q}\tilde{g}$
- $pp \to \tilde{q}\tilde{q}$

It is important to include $\tilde{q}\tilde{q}$ and $\tilde{q}\tilde{g}$ production into the generated processes because $\tilde{q}$'s predominatly decay via $\tilde{g}$'s[2] and thus add to the signal. These processes are also a possible sources of same sign lepton pairs from decay chains other than those of the signal as defined in figure 8.1 and therefore also add to the background. They can not easily be distinguised from the signal via global event characteristics such as missing p_t or number of jets, because they are practically identical. Even though for this study the $\tilde{q}$-masses have been chosen much larger than $m_{\tilde{g}}$ and $\tilde{q}$ production will therefore be supressed, scenarios with more similar choices of mass will have large contributions to the signal cross-section from $\tilde{q}$ production. The cross-section for the above processes and it's subprocesses were calculated by ISAJET to be:

$$\begin{aligned} \sigma(p+p \to \tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}) &= 12.84 \quad \text{pb} \\ \sigma(p+p \to \tilde{g}\tilde{q}) &= 3.23 \quad \text{pb} \\ \sigma(p+p \to \tilde{q}\tilde{q}) &= 0.35 \quad \text{pb} \end{aligned}$$

Intrinsic to these cross-sections is the requirement $10\text{GeV} < P_t(\tilde{g}, \tilde{q}) < 2.5\text{TeV}$. This is effectively a null cut but has importance for the normalisation of the calculated cross-section. The cross-section was found to not change if the lower cut was taken to 1 GeV and the upper cut was varied between 1.2 and 2.8 TeV.
The total branching ratio for the process $\tilde{g}\tilde{g} \to$ same sign ee, $\mu\mu$ or $e\mu$ with the leptons coming from the W of interest can be calculated from tables 8.3 and 8.4. It is:

$$\begin{aligned} BR(signal) &= P(samesign) \cdot \\ &\quad BR(\tilde{g} \to \tilde{\chi}_1^{+(-)} + X)^2 \cdot \\ &\quad BR(\tilde{\chi}_1^{+(-)} \to \tilde{\chi}_1^o + e(\mu) + \nu_{e(\mu)})^2 \\ &= 0.5 \cdot 0.554^2 \cdot 0.22^2 \\ &= 0.74\% \end{aligned}$$

In the above formula $P(samesign)$ is the probability for both $\tilde{g}$'s to decay into same sign $\tilde{\chi}_1^{+(-)}$'s. The most important branching ratios of $\tilde{q}$'s in this process are shown in table 8.5. It can be seen that the majority of all $\tilde{q}$'s decay via $\tilde{g}+X$ (82%). This means that a $\tilde{q}$ produced in association with a $\tilde{g}$ is equivalent to a $\tilde{g}$-pair in 82% of all cases and a $\tilde{q}$-pair in $82\%^2 = 67\%$ of all cases. Knowing this and the relative cross-sections of $\tilde{g}\tilde{g}$ $\tilde{q}\tilde{g}$ and $\tilde{q}\tilde{q}$ production it can be found that the generated process leads to a $\tilde{g}\tilde{g}$ pair in 94.6% of all cases. The second largest channel in $\tilde{q}$ decays is that of $\tilde{q}_R$ to $\tilde{\chi}_{1,2}^{+(-)} + X$ which is a source for same sign leptons if the $\tilde{q}$ was produced in association with a $\tilde{g}$. These same sign pairs are practically indistinguishable from the signal but only happens in 8.6% of all $\tilde{q}$ decays.

[2] see table 8.5

For these above reasons the signal branching ratio of 0.74% has to be multiplied with the signal fraction of the total production cross-section $F_{\tilde{g}\tilde{g}}$. Thus one obtains a signal cross-section of :

$$\begin{aligned}\sigma_{signal} &= BR(signal) \cdot F_{\tilde{g}\tilde{g}} \cdot \sigma(p+p \to \tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}) \\ &= 0.74\% \cdot 94.6\% \cdot 12.84\,\mathrm{pb} \\ &= 0.09\,\mathrm{pb}\end{aligned}$$

BR	Final State	BR	Fin. State	BR	Final State
0.120	$\tilde{\chi}_1^+ \overline{q}_u\ q_d$	**0.120**	$\tilde{\chi}_1^- q_u\ \overline{q}_d$	**0.120**	$\tilde{\chi}_1^+ \overline{q}_c\ q_s$
0.120	$\tilde{\chi}_1^- q_c\ \overline{q}_s$	**0.037**	$\tilde{\chi}_1^+ \overline{q}_t\ q_b$	**0.037**	$\tilde{\chi}_1^- q_t\ \overline{q}_b$
$< 0.1\%$	$\tilde{\chi}_2^+ \overline{q}_u\ q_d$	$< 0.1\%$	$\tilde{\chi}_2^- q_u\ \overline{q}_d$	$< 0.1\%$	$\tilde{\chi}_2^+ \overline{q}_c\ q_s$
$< 0.1\%$	$\tilde{\chi}_2^- q_c\ \overline{q}_s$	0.001	$\tilde{\chi}_1^o g$	0.040	$\tilde{\chi}_1^o q_u\ \overline{q}_u$
0.016	$\tilde{\chi}_1^o q_d\ \overline{q}_d$	0.016	$\tilde{\chi}_1^o q_s\ \overline{q}_s$	0.040	$\tilde{\chi}_1^o q_c\ \overline{q}_c$
0.016	$\tilde{\chi}_1^o q_b\ \overline{q}_b$	0.005	$\tilde{\chi}_1^o q_t\ \overline{q}_t$	0.002	$\tilde{\chi}_2^o g$
0.063	$\tilde{\chi}_2^o q_u\ \overline{q}_u$	0.056	$\tilde{\chi}_2^o q_d\ \overline{q}_d$	0.056	$\tilde{\chi}_2^o q_s\ \overline{q}_s$
0.063	$\tilde{\chi}_2^o q_c\ \overline{q}_c$	0.056	$\tilde{\chi}_2^o q_b\ \overline{q}_b$	0.004	$\tilde{\chi}_2^o q_t\ \overline{q}_t$
0.006	$\tilde{\chi}_3^o g$	$< 0.1\%$	$\tilde{\chi}_3^o q_u\ \overline{q}_u$	$< 0.1\%$	$\tilde{\chi}_3^o q_d\ \overline{q}_d$
$< 0.1\%$	$\tilde{\chi}_3^o q_s\ \overline{q}_s$	$< 0.1\%$	$\tilde{\chi}_3^o q_c\ \overline{q}_c$	$< 0.1\%$	$\tilde{\chi}_3^o q_b\ \overline{q}_b$
0.004	$\tilde{\chi}_4^o g$	$< 0.1\%$	$\tilde{\chi}_4^o q_u\ \overline{q}_u$	$< 0.1\%$	$\tilde{\chi}_4^o q_d\ \overline{q}_d$
$< 0.1\%$	$\tilde{\chi}_4^o q_s\ \overline{q}_s$	$< 0.1\%$	$\tilde{\chi}_4^o q_c\ \overline{q}_c$	$< 0.1\%$	$\tilde{\chi}_4^o q_b\ \overline{q}_b$
0.445	$\tilde{\chi}_{i=1,4}^o$X	0.555	$\tilde{\chi}_{1,2}^{+(-)}$X		

Table 8.3: $\tilde{g}$-branching ratios. The branching ratios relevant for the signal are those into $\tilde{\chi}_1^{+(-)}$+X and are typed in bold face. They add up to 55.4%. The decay into $\tilde{\chi}_2^{+(-)} + X$ is not considered because it is very small.

Particle	BR	Final State	Particle	BR	Final State
$\tilde{\chi}_1^+$	**0.111**	$\tilde{\chi}_1^o e^+\ \nu_e$	$\tilde{\chi}_1^+$	**0.111**	$\tilde{\chi}_1^o \mu^+\ \nu_\mu$
$\tilde{\chi}_1^+$	0.111	$\tilde{\chi}_1^o \tau^+\ \nu_\tau$	$\tilde{\chi}_1^+$	0.333	$\tilde{\chi}_1^o q_u\ \overline{q}_d$
$\tilde{\chi}_1^+$	0.333	$\tilde{\chi}_1^o q_c\ \overline{q}_s$			

Table 8.4: $\tilde{\chi}_1^{+(-)}$-branching ratios. The charge conjugate branching ratios for $\tilde{\chi}_1^-$ are assumed identical to those for $\tilde{\chi}_1^+$. Since $\tilde{\chi}_1^{+(-)}$ is lighter than $\tilde{\chi}_2^o$, $\tilde{\chi}_1^{+(-)}$ exclusively decays to $\tilde{\chi}_1^o$ and it's visible final states are identical to that of a W^* with additional missing p_t. The branching ratios relevant for the signal are typed in bold face. They add up to 22%.

Particle	BR	Final State	Particle	BR	Final State
$\tilde{q}_{d,s,b,R}$	0.988	$\tilde{g}q$	$\tilde{q}_{u,c,R}$	0.953	$\tilde{g}q$
$\tilde{q}_{u,..,b,R}$	≈ 0.0	$\tilde{\chi}_{1,2}^{+(-)}\overline{q}$,	$\tilde{q}_{t,R}$	0.587	$\tilde{g}q$
$\tilde{q}_{t,R}$	0.15	$\tilde{\chi}_{1,2}^{+(-)}\overline{q}$,	$\tilde{q}_{u,..,b,L}$	0.785	$\tilde{g}q$
$\tilde{q}_{u,..,b,L}$	≈ 0.14	$\tilde{\chi}_{1,2}^{+(-)}\overline{q}$,	$\tilde{q}_{t,L}$	0.478	$\tilde{g}q$
$\tilde{q}_{t,L}$	0.186	$\tilde{\chi}_{1,2}^{+(-)}\overline{q}$,	$\tilde{q}_{any}$	**0.822**	$\tilde{g}$X

Table 8.5: $\tilde{q}$-branching ratios relevant to this analysis. The charge conjugate branching ratios are assumed identical. The average ratio of any $\tilde{q}$ into $\tilde{g}$X is typed in bold face.

8.4 Data Samples

For this study only μ's were accepted as final state leptons because the pattern recognition programs used for reconstruction had not yet been tuned for electrons. It is however assumed that the electron fraction of the signal is accessible with similar detector performance as the μ fraction. Since there are at least two leptons in the signal the total signal including electrons will be four times that with only μ's. The simulation of the data samples was done in three steps.

1. Generation of the events with ISAJET and implementation of cuts at particle level

2. Simulation of the event inside the ATLAS inner detector using DICE

3. Background addition of minimum bias events, digitisation and reconstruction using DICE/ATRECON/iPatRec. The minimum bias events were produced beforehand with PYTHIA 7.1 and separately run through the detector simulation with DICE.

All computations were performed on the Oxford NPL DEC-ALPHA clusters central server which was a Digital 2100 Server 5/250. [3]
In order to reduce the CPU time and storage requirements for the simulation all cuts described below were implemented on the particle level during the event generation with ISAJET and only events passing all cuts were output to the detector simulation and reconstruction and into ntuples for statistical analysis. This was necessary since the event size of approximately 60 kB at the output of the generator is inconveniently large if cuts with a rejection factor of 10 have to be implemented after generation and if a sample size of $\geq$ 1000 events is wanted. The event sample written after all particle level cuts was used as input to the reconstruction and is referred to as the "hard"-sample where one event after cuts required 84 sec. Furthermore the cuts could not be implemented inside the full detector reconstruction because the event size after detector simulation (at the input to the reconstruction) grows to 1.8 MB[4] requiring 180 sec/event for simulation and 90 sec/event for digitisation and reconstruction. In principle also events before cuts could have been written to the ntuples of the "hard" sample, but an event by event correspondence between these ntuples and the results of the reconstruction phase was considered more important.
To measure the effects of the particle level cuts a second sample of events with a reduced

[3]Each CPU of this machine is rated at 277.1 SpecInt(92) and 410.4 SpecFp(92). The machine is equipped with 384 MB memory per CPU and 48 GB of internal disk storage.

[4]This does not include the minimum bias pile-up

A-type[5] cut was generated. Events passing the reduced A cut were written to ntuples only. This sample is referred to as the "soft"-sample and each event required 2.5 seconds after cuts. All data sample are described in table 8.6

sample	# events after cut of type						events in		$\int \mathcal{L}$
	none	A_1	A_2	B	C	D	ntuple	ZEBRA	$[10^5\ \mathrm{pb}^{-1}]$
"soft"	950037	-	20000	9021	477	2589	20000	-	0.739
"hard"	877553	5573	-	2534	117	1000	1000	1000	0.683

Table 8.6: Data sample description; the details of the cuts can be found in table 8.7. Implicit to all samples is the ISAJET intrinsic "protocut". The column "events in ZEBRA" specifies how many events have been written for later detector simulation and reconstruction. One LHC-year at $10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ corresponds to $1 \cdot 10^5\ \mathrm{pb}^{-1}$.

8.5 Cuts

Table 8.7 shows the cuts imposed on the dataset before detector simulation, background addition and reconstruction.

Cut	description
Proto	$10\,\mathrm{GeV} < P_t^{\tilde{g},\tilde{q}} < 2.5\,\mathrm{TeV}$
A_1	≥ 2 same sign μ with: ${p_t}^{\mu} > 25\,\mathrm{GeV}$ and $\lvert\eta^{\mu}\rvert < 2.6$
B	$A_1 \wedge \geq 2$ same sign isolated μ's: $E_t^{isolation} < 5\,\mathrm{GeV}$, $R < 0.3$
C	B $\wedge > 2$ same sign prompt μ's with $A0 < 2.5 \cdot \sigma_{A0}({p_t}^{\mu}, \eta^{\mu})$
D	B $\wedge = 2$ same sign prompt μ's with $A0 < 2.5 \cdot \sigma_{A0}({p_t}^{\mu}, \eta^{\mu})$ **signal**
A_2	like A_1 but ${p_t}^{\mu} > 10\,\mathrm{GeV}$

Table 8.7: Cuts before reconstruction. $R = \sqrt[2]{\Delta\phi^2 + \Delta\eta^2}$ is the radius of the isolation cone. The impact parameter resolution was simulated using the parameterisation described in subsection 7.5.1. The simulation of particle lifetime is explained in subsection 8.5.1.

8.5.1 Lifetime Simulation

Even though no background processes were explicitly simulated, the background from $\tilde{g}$-$\tilde{g}$- and $\tilde{g}$-$\tilde{q}$-production in the form of μ's which do not come from the W decay of interest is large. The main source of these μ's are t, b and c quarks in $\mathrm{jets}_{1..4}$ or the decays of $\tilde{g}$ directly to $\tilde{\chi}_1^o q\overline{q}$ as well as τ's and c quarks from the decay of $\tilde{\chi}_1^{+(-)}$. Most of these μ's are associated to the decay of semi stable particles with lifetimes in the range of $c \cdot \tau \hat{=} \mathcal{O}(100)\mu$m . Because ISAJET does not propagate particles in space but only deals with particle momentum coordinates a crude simulation of particle lifetime was added. The decay history of each μ passing the cuts was searched for hadrons containing b or c quarks. Decay length for three classes of particles were assumed and an impact parameter was assigned to these particles according to an exponential decay distribution with a mean as shown in table 8.8. Cascade decays like $b \to c \to d$ were assigned two independent decay times. Particles with multiple b or c hadrons were assigned only one lifetime, since they were considered to have gone through exited states with no significant lifetime. Particles with decay length larger than

[5] see table 8.7

Class	$c \cdot \tau[\mu\text{m}]$
Charged C-Hadrons	317
Neutral C-Hadrons	124
all B-Hadrons	388

Table 8.8: Lifetimes assumed for particles classes. The lifetime is assumed to be the weak lifetime of the b- or c-quark. With the exception of the distinction between charged and neutral C-Hadrons the quark lifetime was assumed to be independent of the type of hadron it is contained in.

1 cm were not considered and they are assumed to not significantly contribute μ's to this sample. They would also be easily identifiable in the detector. It was assumed that the impact parameter be independent of the decaying particles p_t because the opposite effects from the collimation of the decay particles around the initial particle direction and the increased lifetime through the Lorenz boost balance to first order. The impact parameter was then smeared with a circular, Gaussian beam spread of 12.5 μm sigma and the detectors impact parameter resolution as a function of particle p_t and η as described in subsection 7.5.1. A cut on the significance of the impact parameter of 2.5 σ_{A0} was used to suppress non prompt μ's. The variable choosen to describe the performance of the impactparameter cut is the "escape fraction", abbreviated as f_e. f_e is defined for samples of muons with decay length as the ratio of events before and after cuts. The ratio $f_e^{rel} = f_e(C^o \text{ or } C^{\pm})/f_e(B)$ is called "relative escape fraction" and measures the effectiveness of the cut on samples with C^o or $C^{\pm}$ hadrons relative to that for a sample with B hadrons. In order to monitor the understanding and correct implementation of the particle lifetime simulation and detector resolution, expectations for f_e and f_e^{rel} were analytically calculated and compared to the measurements in table 8.10. f_e and f_e^{rel} were calculated as follows:

- It is assumed that the impact parameter cut can be described by one effective cut length $A0_{cut}^{eff}$ for each sample, even though it has a p_t and η dependence. $A0_{cut}^{eff}$ is equivalent to a lifetime cut since $\tau_{cut}^{eff} = A0_{cut}^{eff}/c$. Because the p_t and η distributions for μ's from C^o, $C^{\pm}$ and B hadrons are comparable, but not identical $A0_{cut}^{eff}$ should be evaluated for all three samples separately.

- f_e and f_e^{rel} was measured on samples of μ's from non cascade decays of C^o, $C^{\pm}$ and B hadrons; the values of which are shown in table 8.9.

- Under the above assumptions and in the case of non cascade decays the relation between f_e and $A0_{cut}^{eff}$ is given by:

$$f_e = \left(1 - \exp\left(\frac{-A0_{cut}^{eff}}{c \cdot \tau}\right)\right) \tag{8.1}$$

  where τ is the mean lifetime of the particle which is known. The values for $A0_{cut}^{eff}$ derived from table 8.9 are:
  $A0_{cut}^{eff}(B) = 90^{+5.6}_{-6.3}\,\mu\text{m}$, $A0_{cut}^{eff}(C^o) = 89 \pm 2.4\,\mu\text{m}$ $A0_{cut}^{eff}(C^{\pm}) = 75^{+29}_{-32}\,\mu\text{m}$.
  All of these values are compatible with each other.

- If the decay is a cascade decay of type $X \to Y \to Z$ with the two lifetimes $\tau_{X\to Y}$ and $\tau_{Y\to Z}$ its distribution becomes an exponential with a lifetime of $\tau_{X\to Z} = \tau_{X\to Y} + \tau_{Y\to Z}$. In formula 8.1 τ is replaced with $\tau_{X\to Z}$ and the $A0_{cut}^{eff}$ relevant for the cascade decay is assumed to be $A0_{cut}^{eff}(X \to Y)$.

- f_e^{rel} for two decays with total lifetimes τ_a and τ_b is:

$$f_e^{rel}(\tau_a, \tau_b, \tau_{cut}^{eff}(a), \tau_{cut}^{eff}(b)) = \frac{1 - \exp(-\frac{\tau_{cut}^{eff}(a)}{\tau_a})}{1 - \exp(-\frac{\tau_{cut}^{eff}(b)}{\tau_b})} \tag{8.2}$$

- To estimate f_e for samples which each contain more than one decay type (i.e. cascade and none cascade decays), the sample is split into pure cascade and non cascade subsamples and f_e is calculated for each subsample. Then the average of f_e is calculated, weighing each subsample with the fraction of events it contains.

Following the above steps the impact parameter cut was found to behave precisely as expected as can be seen in table 8.10.

f_e			f_e^{rel}	
B	C^o	$C^{\pm}$	C^o/B	$C^{\pm}$/B
0.207±0.003	0.51±0.01	0.21±0.02	2.46±0.06	1.01±0.09

Table 8.9: Fraction of non prompt μ's passing the impact parameter cut. Only μ's from **non cascade decays** of B, $C^{\pm}$ and C^o are shown. Columns (1,2,3) show the escape fractions (see text in section 8.5.1). Columns 4 and 5 show the relative escape fractions where $f_e^{rel}(X) = f_e(X)/f_e(B)$.

Cut	f_e			f_e^{rel}	
	B	C^o	$C^{\pm}$	C^o/B	$C^{\pm}$/B
α	0.1968±0.0032	0.350±0.017	0.159±0.015	1.776±0.091	0.81±0.08
β	0.1157±0.0024	0.0709±0.0012	0.088±0.010	0.613±0.016	0.761±0.09
γ	0.1095±0.0078	0.1050±0.0076	0.005±0.002	0.959±0.097	0.05±0.02
α_{exp}	0.191	0.342	0.168	1.79	0.88

Table 8.10: Fraction of non prompt μ's passing various cuts. μ's from **all decays** of B, $C^{\pm}$ and C^o are shown. The table is similar to table 8.9. The first column specifies the cuts: α=impact parameter cut, β=impact parameter and isolation cut, γ=cut D (see table 8.7) in the "hard" sample, α_{exp}=expectation for impact parameter cut. The fractions of cascade decays in the subsamples (needed to calculate the expected values) is shown in table 8.11. Columns (2,3,4) are the escape fractions. Columns 5 and 6 are the relative escape fractions. The $C^{\pm}$ sample for row γ is statistics limited as it only contains 5 events.

	parent particle class		
decay	B	C^o	$C^{\pm}$
B→$C^o + X \to \mu + X$'	31.9%	49.8%	-
B→$C^{\pm} + X \to \mu + X$'	2.1%	-	47.5%

Table 8.11: Percentage of cascade decays in the the decays of B, C^o and $C^{\pm}$ Hadrons.

8.6 Event Characteristics

Some of the global event characteristics are given in this section. An analysis of the achieved signal purity after all particle level cuts is given and the problems of this signal channel are discussed.

8.6.1 Multiplicities

Figure 8.2 shows charged and neutral track multiplicities of the signal events before and after cuts. The very large average number of charged tracks in the tracker acceptance of 240 tracks per event (with tails up to 500 tracks per event) is very important for the TRT tracker and it is comparable to the multiplicities expected from minimum bias pile-up.

8.6.2 Lifetime Content

Table 8.12 shows a breakdown of the μ parent particles into three categories of lifetime contributing particles before and after cuts. The total number of μ's coming from decays of long lived particles (B,C^o,$C^{\pm}$) in the "hard" sample which was used for later reconstruction is 433 out of 2000. The influence of the impactparameter cut has been explained in detail in section 8.5.1. The actually achieved reductions and their expected values were shown in table 8.10.

Sample	Cut	B	C^o	$C^{\pm}$
soft	A_2	0.438	0.280	0.019
soft	A_2+impact	0.149	0.169	0.0053
soft	A_2+impact+isol	0.134	0.151	0.0045
hard	D	0.110	0.105	0.002

Table 8.12: B-hadron, neutral and charged C-hadron content of the μ history for different cuts. For each cut, the values are the fraction of μ with lifetime from a specific parent particle class out of all muons passing the cut. The cuts are explained in table 8.7.

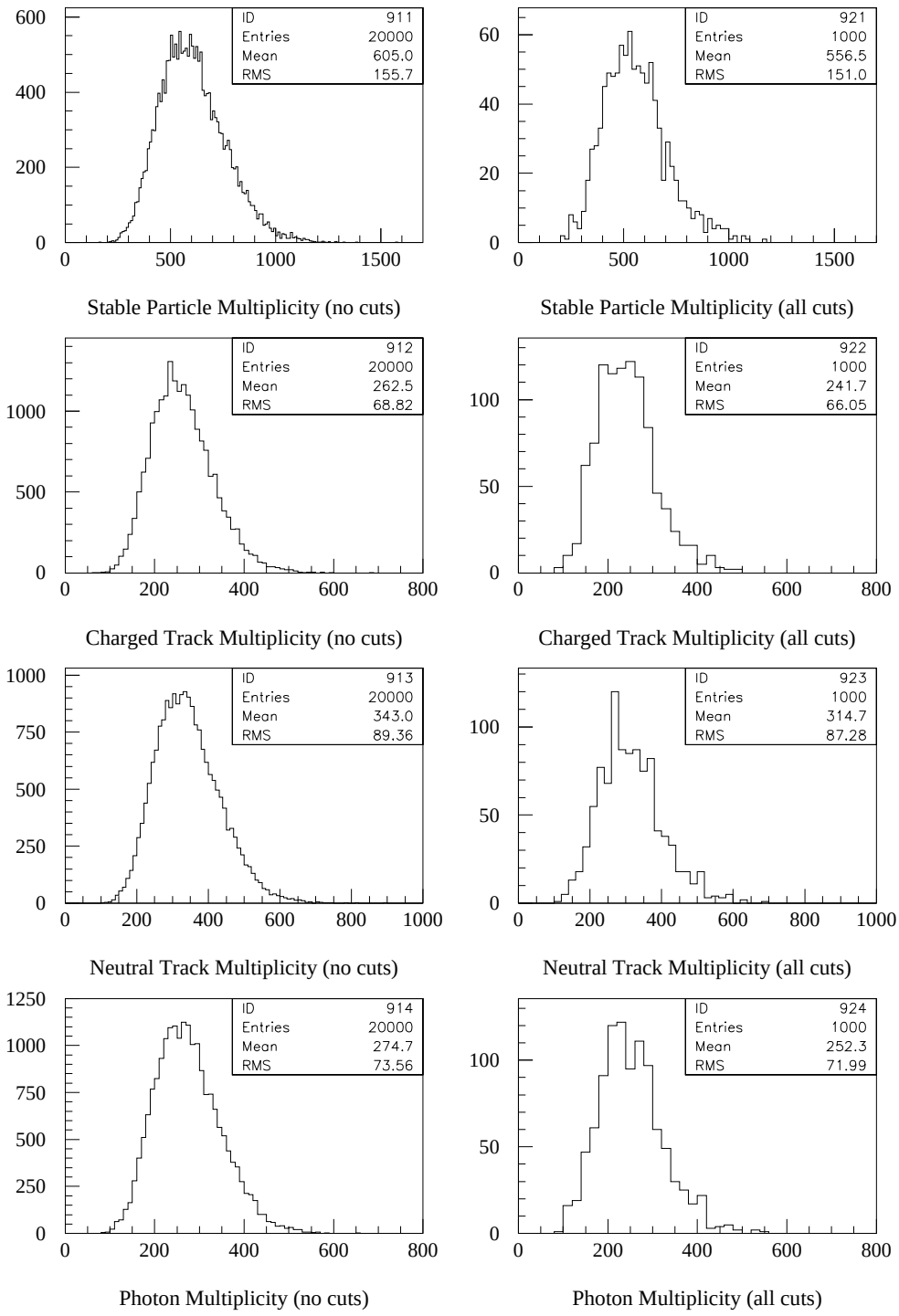


Figure 8.2: Multiplicities before and after cuts. The histograms marked 'no cuts' correspond to the "soft" sample described in table 8.6 and have passed cut A_2 from table 8.7. The histograms marked 'all cuts' correspond to the "hard" sample and have passed cut D.

8.6.3 Impact Parameter

Figure 8.3 shows the impact parameter distribution before and after cuts. The peak at low impact parameters in the "soft" sample distributions corresponds to the beam spread of 12.5 microns.

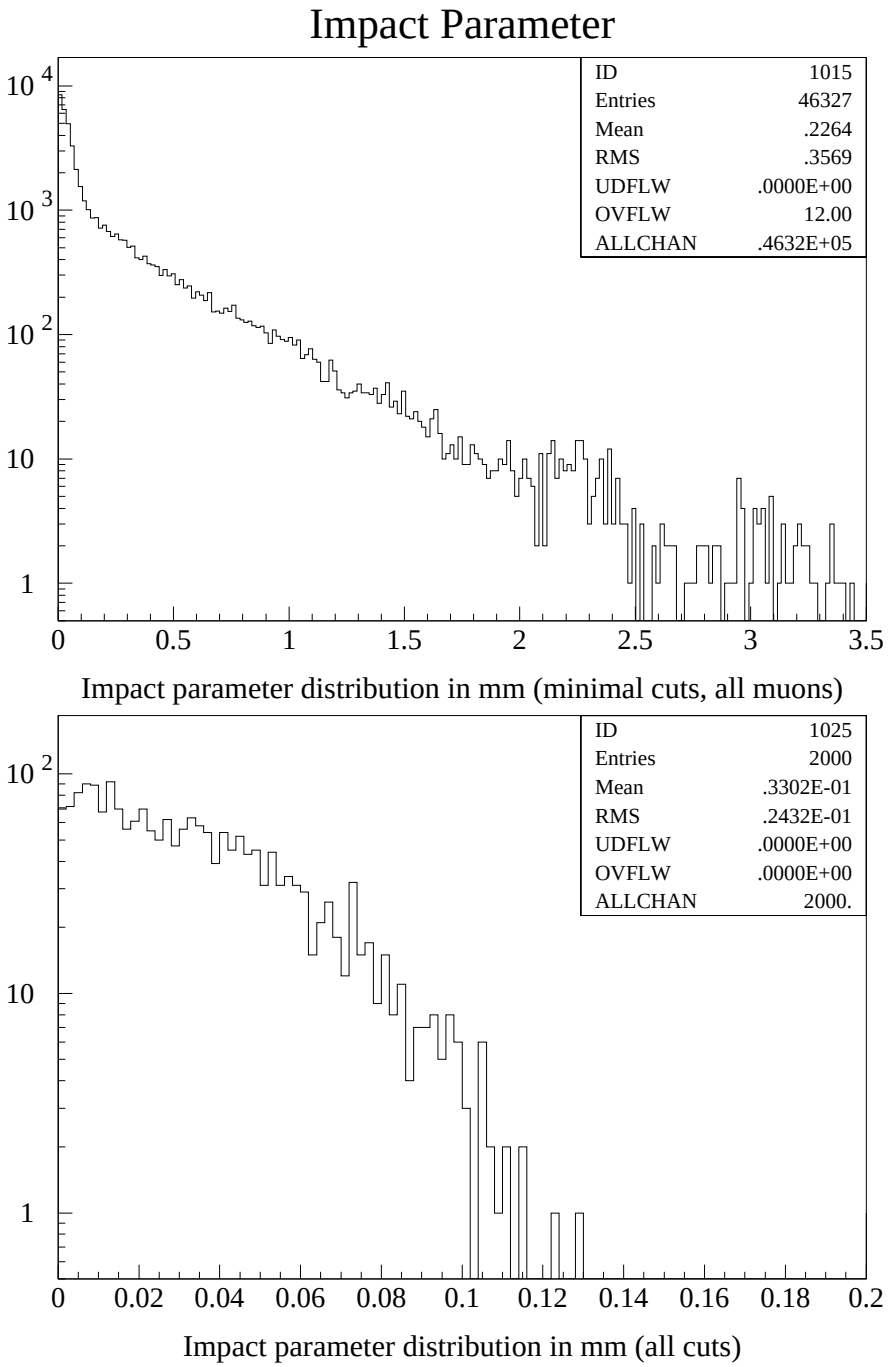


Figure 8.3: Impact parameter distributions from the "soft" and "hard" sample (see table 8.6. The first plots contains all μ's from the "soft" sample. The second plot contains the two signal μ's per event from the "hard" sample.

8.6.4 Isolation Energies

The transverse isolation energies for all and for charged particles inside an isolation cone of $\sqrt[2]{\Delta\phi^2 + \Delta\eta^2} < 0.3$ are shown in figures 8.4 and 8.5 before and after cuts. The second and third histogram are very similar, showing that there is little correlation between the isolation cut and the other cuts applied to the "hard" sample.

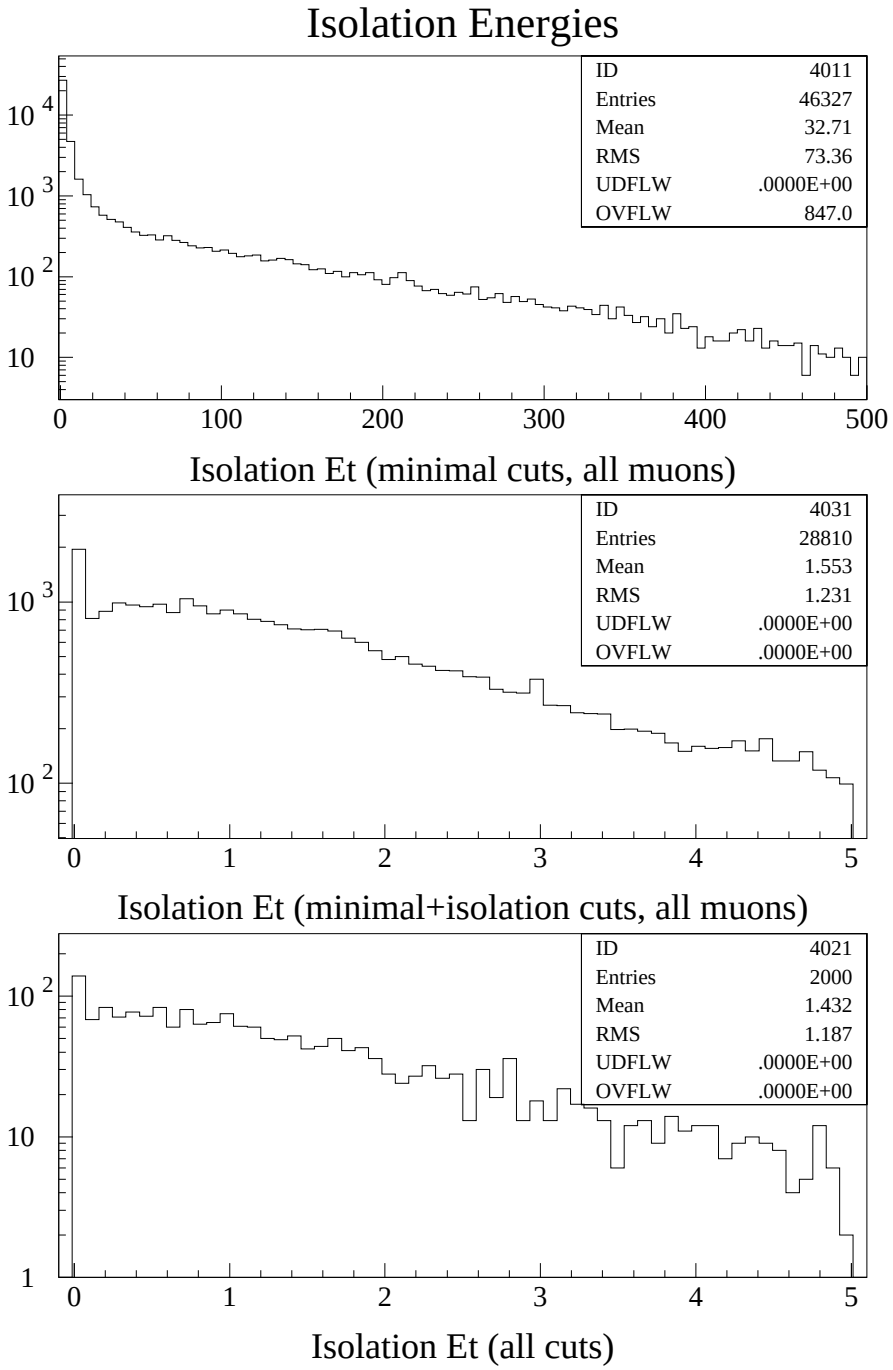


Figure 8.4: Transverse isolation energy inside a cone of $\Delta R = 0.3$. The top histogram contains all μ's from the "soft" sample. The second histogram contains all μ's from the soft sample passing the isolation cut and the third histogram contains the two signal μ's per event from the "hard" sample.

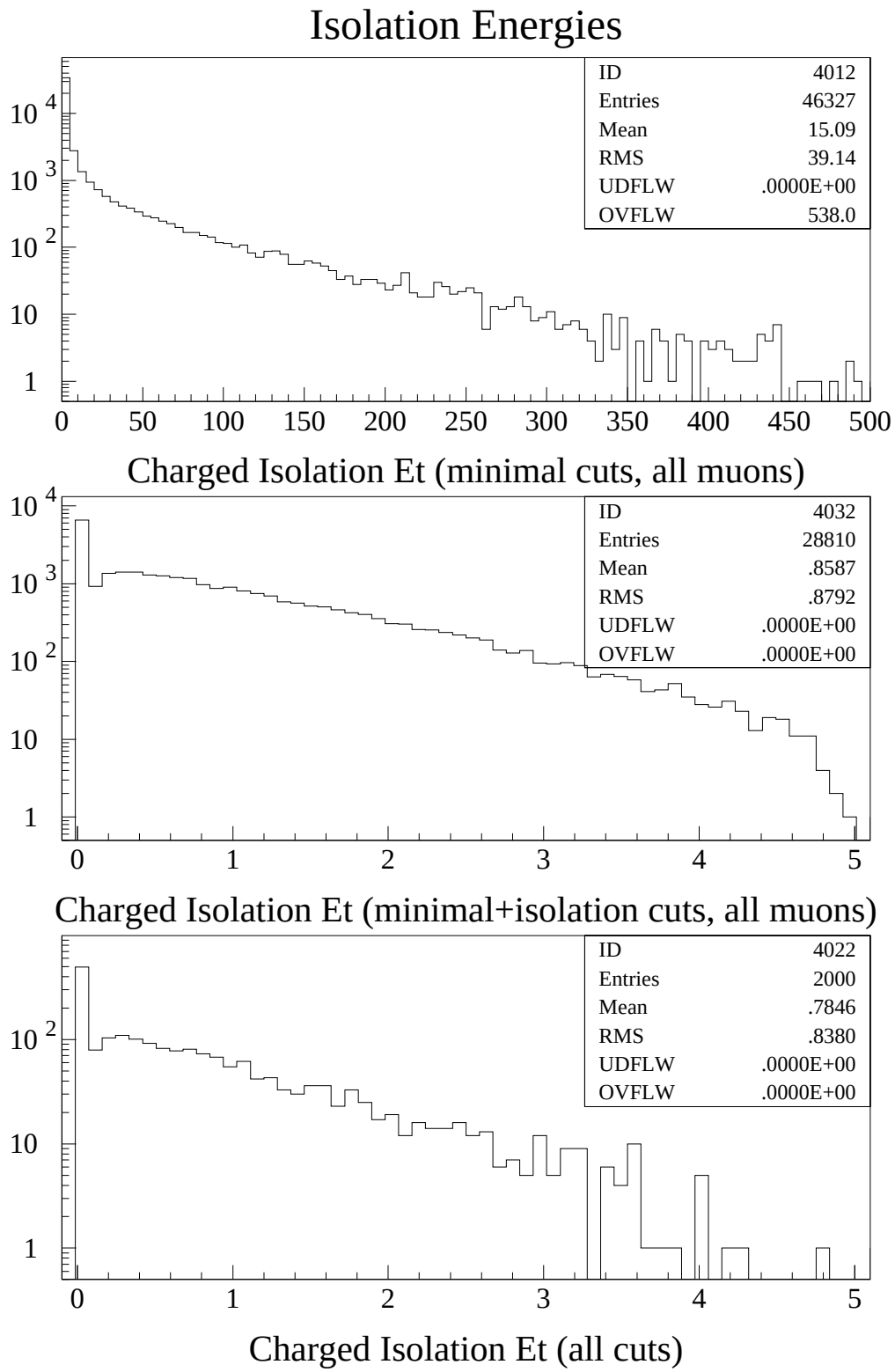


Figure 8.5: Transverse isolation energy contributed from charged tracks, inside a cone of $\Delta R = 0.3$. The top histogram contains all μ's from the "soft" sample. The second histogram contains all μ's from the soft sample passing the isolation cut and the third histogram contains the two signal μ's per event from the "hard" sample.

8.6.5 p_t Spectra

Figure 8.6 shows the transverse momentum spectrum of the μ's for different samples and cuts. It is obvious that the spektrum is rather soft and falls only marginally slower than a simple exponential function. This is expected, because the μ's originate from the decay of a virtual W.

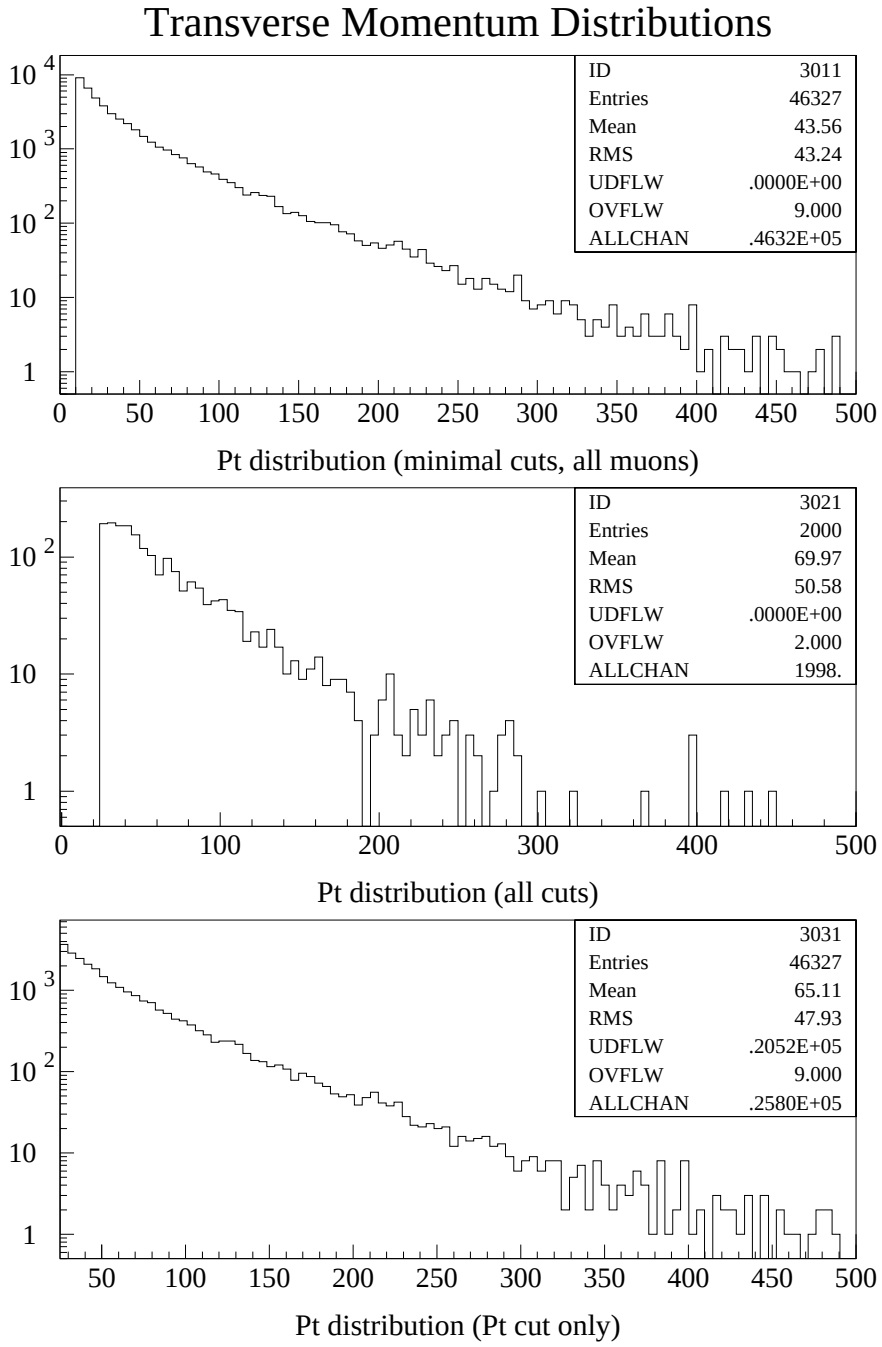


Figure 8.6: p_t distributions.
Top: All μ's from the "soft" sample. **Middle:** "hard" sample after cut D
Bottom: "soft" sample after p_t cut of 25 GeV (44% of all signal μ's do not pass the p_t cut)

8.6.6 Multiple Muons

Figure 8.7 shows how the cuts influence the total number of μ's in the event. μ's were

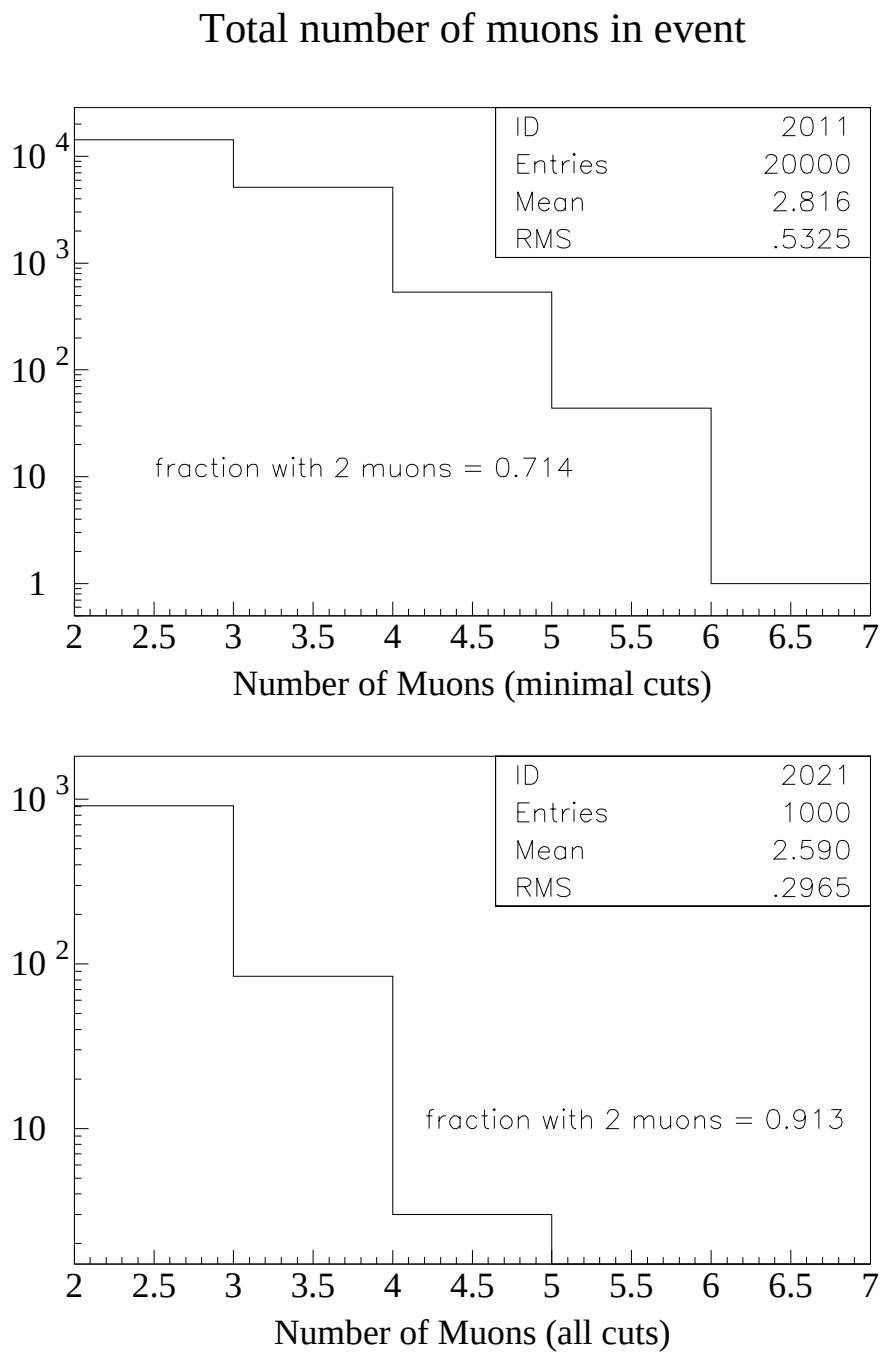


Figure 8.7: Total number of μ's for the soft sample and the hard sample after cuts D. The cuts reduce the fraction of events with multiple μ's because in cut D only exactly two μ's passing D were allowed. In events from the "hard" sample with $n > 2$ μ's n-2 μ's failed at least one of the conditions in cut D

counted if they entered the detector acceptance ($|\eta^{\mu}| < 2.6$) with $p_t{}^{\mu} > 3$GeV. If there were more than 2 μ's in the "hard" sample they must have failed one of the conditions in cut D, because only events with exactly two μ's passing all cuts were accepted. The total number of μ's is a critical number, because only for exactly two μ's the signal condition of "same sign" μ's is well defined. This implies that otherwise good events will have to be

discarded if a third μ's passes all cuts. Since the p_t spectrum shown in figure 8.6 of the μ's peaks towards low p_t a harder p_t cut is no solution to this problem. Already a cut at 25 GeV discards 44 % of all μ's.
Since the mean p_t after all cuts (70 GeV) is very similar to that after only the p_t cut (65 GeV) it can be assumed that there is little correlation between the p_t cut and the other cuts. A procedure identifying a pair of μ's as belonging to the W^*'s from these events is needed. It is not clear how this could be done best. In section 8.7 the Z-coordinate of the μ's vertex has been used to discard μ's from pile up events, but this can not be used to correlate μ's from different branches of the gluino decay chain. Angular distributions between the signal μ's are not easily calculable and the large p_t^{miss} makes it difficult to use correlations between the p_t of the μ's. Using the angular distribution between the two jets from the $\tilde{g}$ three body decay could be one possibility to enhance the signal.

8.6.7 Signal Purity

Finally the purity of the event sample is shown in figure 8.8 which shows the parent particle identification code (in ISAJET convention) of the muon for the following four cases in which $R_{\tilde{q}/\tilde{g}}$ indicates the ratio of the number of μ's from $\tilde{q}$'s. to that from $\tilde{g}$'s.

- 2000 μ's that have been output to the reconstruction (1601 from $\tilde{g}$'s, 346 from $\tilde{q}$'s and 53 from other hadrons[6], $R_{\tilde{g}/\tilde{q}} = 0.216$)

- 1602 μ's which do not have long lived parent particles (1307 from $\tilde{g}$'s, 280 from $\tilde{q}$'s and 15 from other hadrons, $R_{\tilde{g}/\tilde{q}} = 0.214$)

- 1233 μ's which do not have long lived parent particles and did have a $\tilde{\chi}_1^{+(-)}$ in their history. (1009 from $\tilde{g}$'s, 224 from $\tilde{q}$'s[7], $R_{\tilde{g}/\tilde{q}} = 0.222$).

- 808 μ's in which both μ's in the event satisfied the above conditions (658 μ's with one μ from $\tilde{g}$, 540 with both μ's from $\tilde{g}$ and 150 μ's from $\tilde{q}$'s. $R_{\tilde{g}/\tilde{q}} = 0.228$)

$R_{\tilde{g}/\tilde{q}}$ is comparable to the ratio of the number of $\tilde{q}$'s to the number of $\tilde{g}$'s initially produce (before particle level cuts) which is 0.213 calculated from the relative cross-sections.
From the 808 remaining $\mu's$ 33 did not have a gluino in their decay history, leaving 371 pure signal events, 224 of which have been produced through initial $\tilde{g}\tilde{g}$ pair production. This means, that only 37.1% of the events accepted by the particle level cuts are signal events. The two main sources of background events remaining in the sample are listed below.

1. μ's from b and c quarks (398)

2. decays of $\tilde{g}$'s and $\tilde{q}$'s not via $\tilde{\chi}_1^{+(-)}$ (369)

This result emphasises the need to find a correlation between the μ's of the signal events, other than their sign of charge.

[6] from the underlying event

[7] other hadrons are excluded by definition, because these can never decay via a charginoa

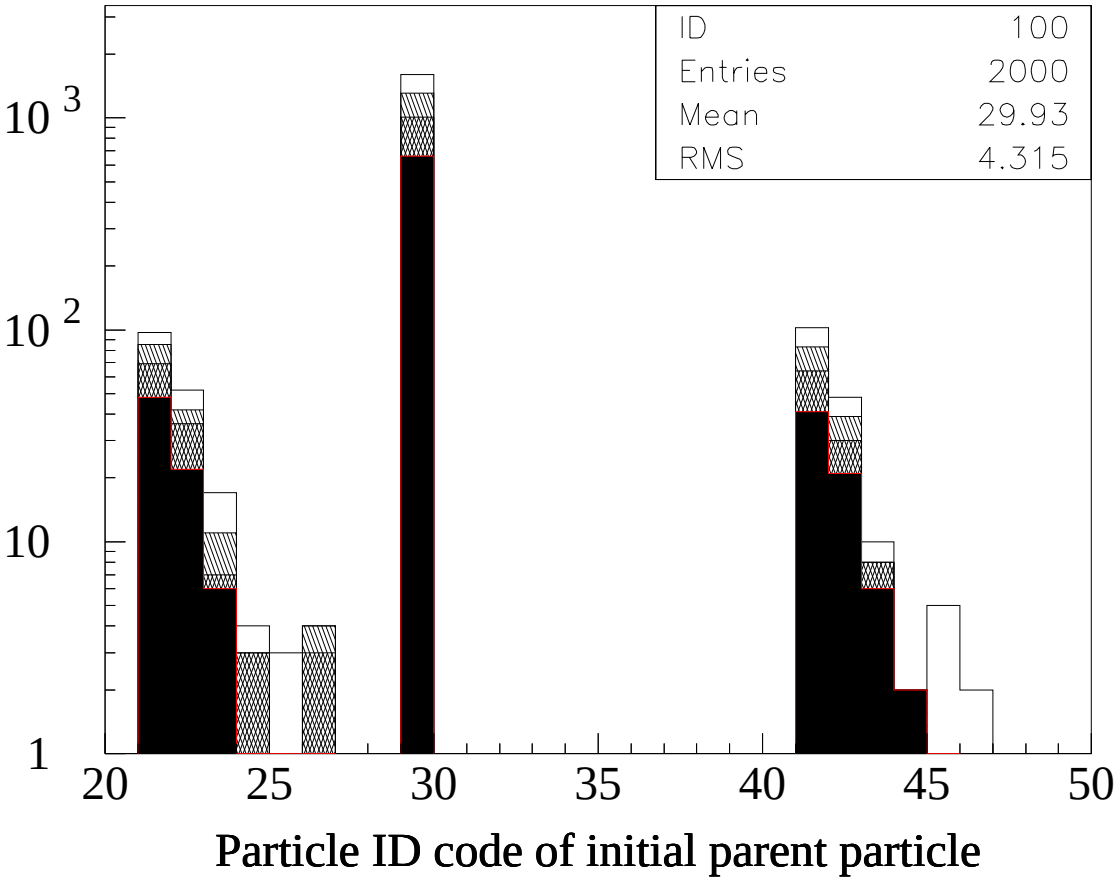


Figure 8.8: Particle identity code of the original mother of the μ's from the "hard" sample after all particle level cuts (cut D in table 8.7) for 4 requirements on the signal purity.
1. = not hatched $\hat{=}$ all muons in the "hard sample" (2000)
2. = single hatched $\hat{=}$ require no long lived particles (1602)
3. = double hatched $\hat{=}$ additionaly require $\tilde{\chi}_1^{+(-)}$ in the decay chain (1233)
4. = solid black $\hat{=}$ both μ's satisfy 3. (808). This is the signal condition.
The particle codes of interest are:
21...25=$\tilde{q}_L^i$ i=(u,d,s,c,b) and 26=$\tilde{q}_1^t$
41...45=$\tilde{q}_R^i$ i=(u,d,s,c,b) and 46=$\tilde{q}_2^t$
30=$\tilde{g}$

8.7 Pattern Recognition Results

In this section the reconstruction of the 2000 μ's from the hard sample is described. Detailed measurements of the detector performance in comparison to the expectations from events with single tracks and pileup are given for all track parameters in the form of "pull" variables which are the distributions of the track parameters around their true values normalised to the expected resolutions as determined in chapter 7.

8.7.1 Roads

The reconstruction was performed in a rather narrow initial road, assuming that the μ track prameters have been measured by the muon detectors and were interpolated into the inner detector. The roads had a curvature incorporating particles with transverse momenta between -20 and 20 GeV. At the vertex the width of the road in Z remained at $\pm$ 20 cm allowing for the complete longitudinal vertex spread to be covered. In the transverse direction it was tightened to ± 20 μm , assuming that the particles of interest are prompt. If the extrapolation accuracy of the muon system in the Z coordinate is shown to be good down to transverse momenta of 20 GeV the road width in Z can be drastically reduced and pattern recognition ambiguities could be avoided. The widths in η and ϕ at the radius of the calorimeter were assumed to be known with accuracies of 0.04 and 0.08 respectively. This is significantly more accurate than the roads described in chapter 7 because the lowest transverse momenta are now 25 GeV and not 2 GeV and thus measurements of both muons and electrons by the muons system and the electromagnetic calorimeters could be much more precise. A further motivation for the differences in the road parameters is the fact that in chapter 7 roads were maintained large in order to be able to find low momentum accompanying particles which is usefull for various other studies but was thought unnecessary for this reconstruction.

8.7.2 Post Reconstruction Cuts

In order to minimise the missidentification probability and maintain a high reconstruction efficiency two cuts were applied to the sample of reconstructed tracks and tuned.

A One and only one track per seed must have been reconstructed.

B The pull of the separation of the vertices of the two reconstructed tracks in z must be less than 2.4 and condition **A** is required for both tracks[8].

Cut A discards the cases in which there was pattern recognition confusion and ghost tracks were found. Allthough in most of the cases discarded by cut A a track with good match to the true track was found it is not possible to correctly choose with good accuracy between the alternative tracks on the basis of the χ^2 of the track fit or the number of hits on the track. A more detailed analysis considering the number of holes and shared hits, as well as the number of hits contributed from the TRT and discrete trackers separately could possibly avoid this cut and obtain a higher efficiency.
Cut B discards events in which a track from a pile up event which usually has different vertex coordinates is paired with a track from a signal μ.

[8]Because the pull of Z_v^1-Z_v^2 has a width of only 0.75 the cut at 2.4 corresponds to a cut at 3.24 σ

8.7.3 Efficiencies, Fake Tracks, Sign of Charge

The efficiencies, fake track rates, charge misidentification rates, number of tracks passing the cuts and the mean of a normalised χ^2-deviation of the reconstructed tack from the true track are shown in table 8.13. In total 2230 tracks were reconstructed from 2000 seeds. For 23 seeds no track was found. And of the 1542 uniquely reconstructed tracks 2 had the wrong sign. This corresponds to a fake track rate, single track efficiency and charge misidentification probability of:

$$R_{fake} = \frac{2230 - (2000 - 23)}{2000 - 23} = \frac{253}{1977} = 12.8\%$$

$$\epsilon_{track} = \frac{1977}{2000} = 98.9\%$$

$$P_{wrong\ sign} = \frac{2}{1542} = 0.13\%$$

	No cut	cut A	cut B
$\#\mu_{tracks}, \leq$ 1per seed	1977	1620	1542
$\#\mu_{pairs}, \leq$ 1per event	977	801	771
$\#\mathrm{WS}_{tracks}, \leq$ 1per seed	7 (165)	4	2
$\sigma({p_t}^{-1})_{pull}$	0.917	0.904	0.884
$\sigma(\phi_0)_{pull}$	0.910	0.897	0.883
$\sigma(\theta_0)_{pull}$	0.845	0.831	0.821
$\sigma(A0)_{pull}$	0.862	0.856	0.851
$\sigma(Z_v)_{pull}$	0.746	0.741	0.733
${\chi^2}_{true-found}$	0.847	0.829	0.792
ϵ_{track}	0.989	0.810	0.771
ϵ_{pair}	0.977	0.801	0.771
f_{WS}	0.35%	0.25%	0.13%

Table 8.13: Reconstruction results as a function of the post reconstruction cuts. WS stands for **W**rong **S**ign of charge on a reconstructed track. In total 2000 μ-seeds were input into the reconstruction. The meanig of the σ's of the different pull variables is explained in the following sections.

8.7.4 Transverse Momentum

Figures 8.9 shows the pull variables for the inverse transverse momentum. The normalisation was done using a two dimensinal extrapolation in p_t and η between the resolutions measured for single muons at full luminosity as shown in chapter 7, section 7.5.1. The sigma of the pull is appr. 10% too small which indicates that the reconstruction performed slighly better than expected which could be due to the isolation cut and the narrower initial roads. The normalised χ^2 is 0.97 and the distribution shows a perfect Gaussian shape.

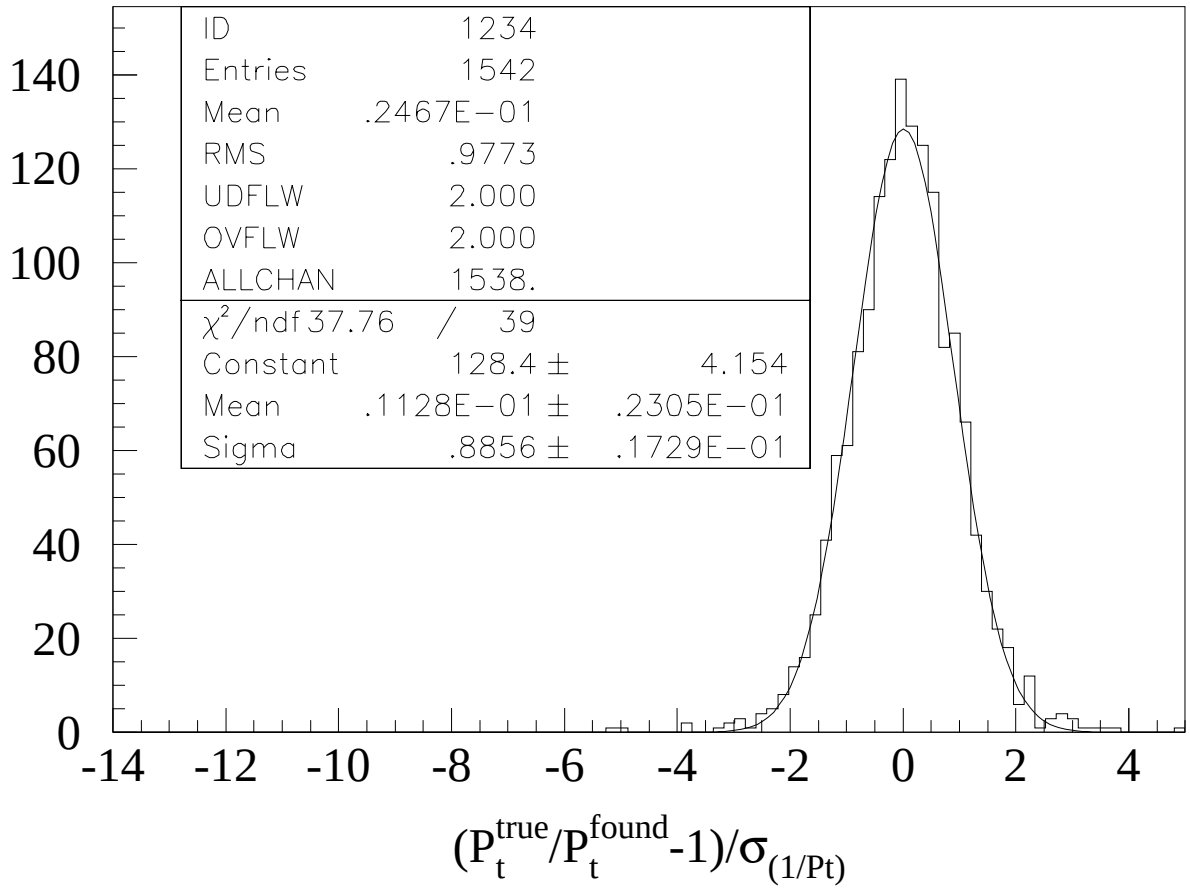


Figure 8.9: Pull variable for the inverse transverse momentum using a two dimensional extrapolation in p_t and η between the resolutions obtained from single muons with minimum bias pileup corresponding to a luminosity of $1{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$.

8.7.5 Impact Parameter

The impact parameter pull has been calculated using the two dimensional parameterisation of the impact paramert resolution as a function of p_t and η described in chapter 7, section 7.5.1 and is shown in figure 8.10. The distribution is very similar to that for the inverse transverse momentum and the shape is also perfectly gaussian with a normalised χ^2 of 1.09 and no nongaussian tails.

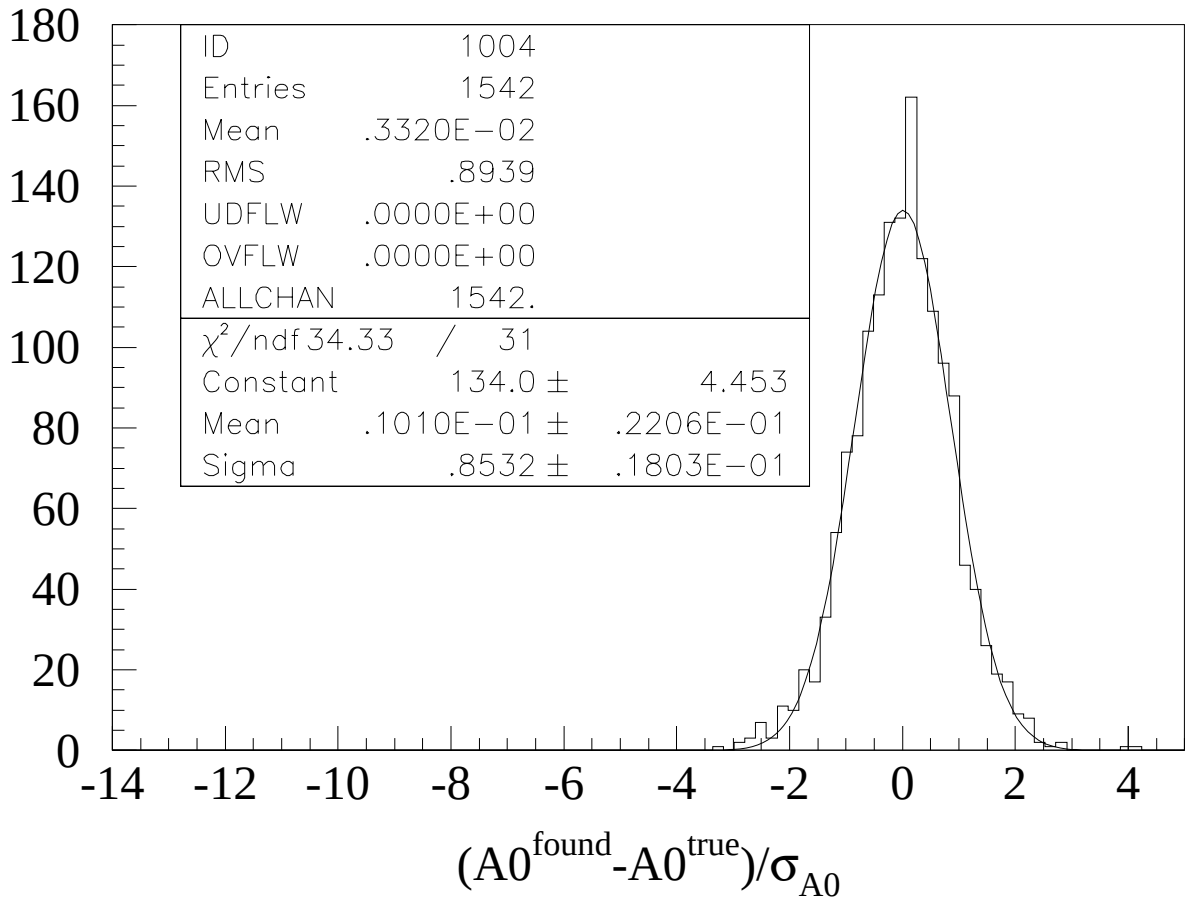


Figure 8.10: Pull variable for the impact parameter using the two dimensional parameterisation of the impact paramert resolution obtained for single muons with minimum bias pileup corresponding to a luminosity of $1{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$for normalisation.

8.7.6 Azimuthal Angle

The pull for ϕ_0 has been calculated using the same extrapolation technique as used for the inverse p_t pull. The results are equaly good. The χ^2 of 0.86 is even lower than that for the inverse p_t pull and the σ of 0.882 is practically identical to that of the inverse p_t pull. Figure 8.11 shows the distribution of the ϕ pull variable.

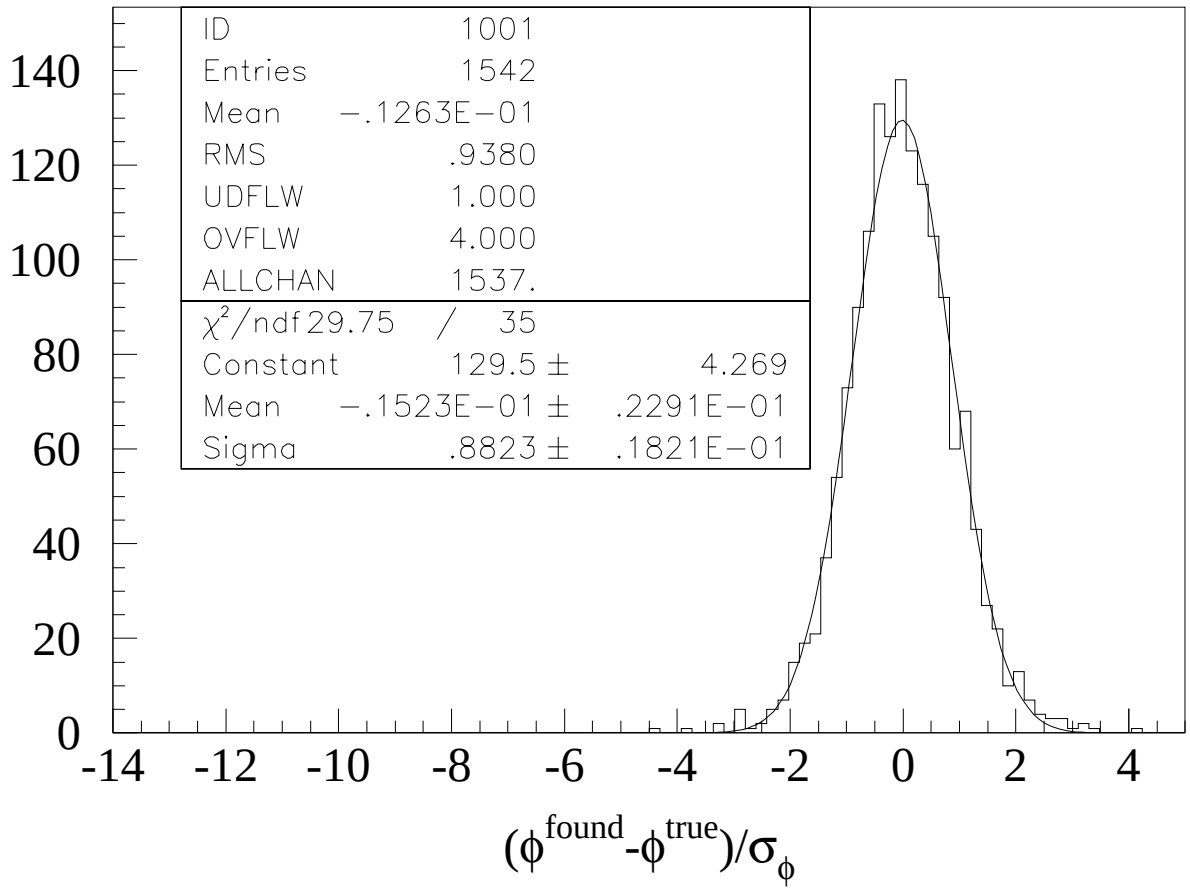


Figure 8.11: Pull variable for ϕ_o using a two dimensional extrapolation in p_t and η between the resolutions obtained from single muons with minimum bias pileup corresponding to a luminosity of $1{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$for normalisation.

8.7.7 Polar Angle

Because the resolution of the polar angle depends only weakly in the transverse momentum it has been determined as a function of η from a momentum averaged resolution containing muons of 2,10,40,100 and 500 GeV transverse momentum without minimum bias pileup. The contribution from the 2 GeV sample has biased this average towards worse resolutions because at 2 GeV the resolution is strongly multiple scattering limited. Given the weight of the 2 GeV sample in the average and the separate resolutions as shown in chapter 7 section 7.5.1 a deterioration of the resolution of approximately 20% due to this effect is expected. The σ of the θ-pull of 0.82 and the above expectation for the missestimate of the resolution make σ compatible with 1. This means, that the detector performence in the polar angle resolution is as good as for single tracks. This was expected from the negligible change of the resolution with increasing luminosity as shown in figure 7.19.

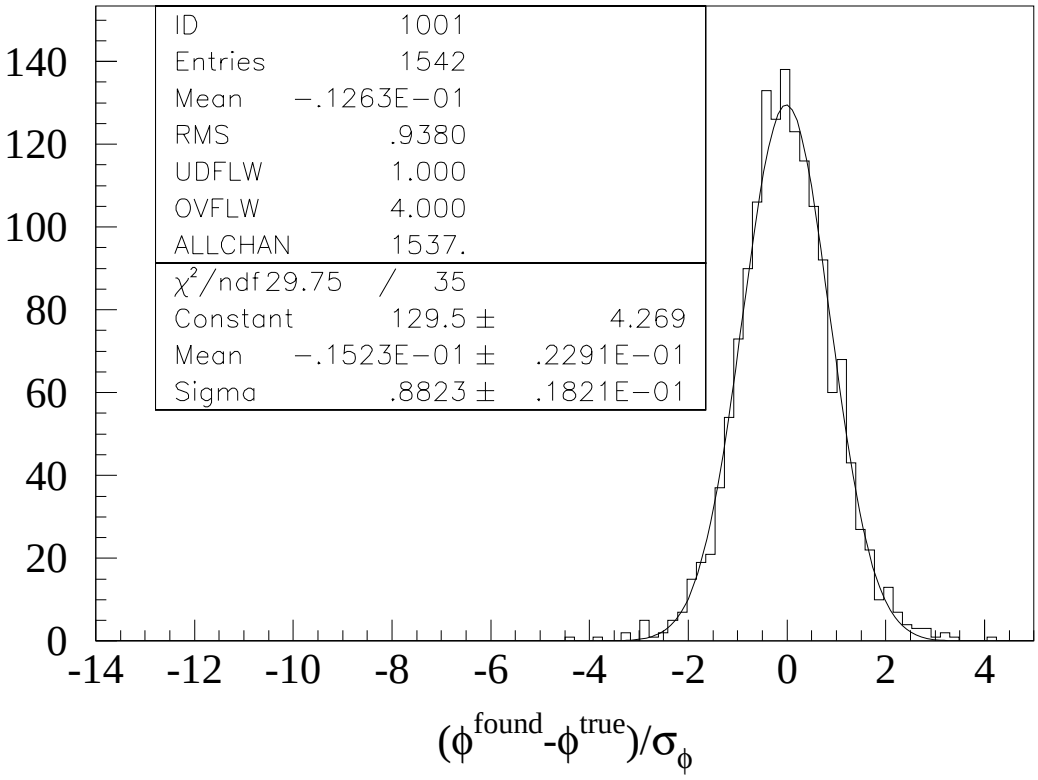


Figure 8.12: Pull variable for θ_o using a one dimensional extrapolation in η between the resolutions obtained from single muons without minimum bias pileup for normalisation. The resolution was estimated from a momentum average over samples of μ's with p_t= 2,10,40,100 and 500 GeV. Because the 2 GeV sample shows a multiple scattering dominated resolution (see figure 7.20) a 20% improvement in the resolution is expected when neglecting the 2 GeV sample, making the widht of the above pull distribution compatible with 1.

8.7.8 Z-Vertex

The Z_v pull shown in figure 8.13 differs from all other pull variables because it is not as Gaussian. This is expected, because allready the resolutions obtained from single muons showed none gaussian shapes for each η value. The resolutions are deformed from a gaussian shape towards a box shape. This is also understandable, because only few measurements contribute to the resolution (effectively only the pixel detectors contribute) and the measurements are coarse in comparison to random track fluctuations from multiple scattering. The badly normalised χ^2 of 1.53 is a manifestation of the none Gaussian shape. Furthermore the pull variable has a fitted width of only 0.73 which differs from the RMS value of 0.8. Since there is only a negligible momentum dependance of the Z_v resolution it is not clear why this variable has narrower pull than the other variables. A possible explanation could be the isolation cut applied to this sample which may reduce the number of clusters in the pixel layers from closely spaced tracks which can merge into each other and thus degrade the z resolution.

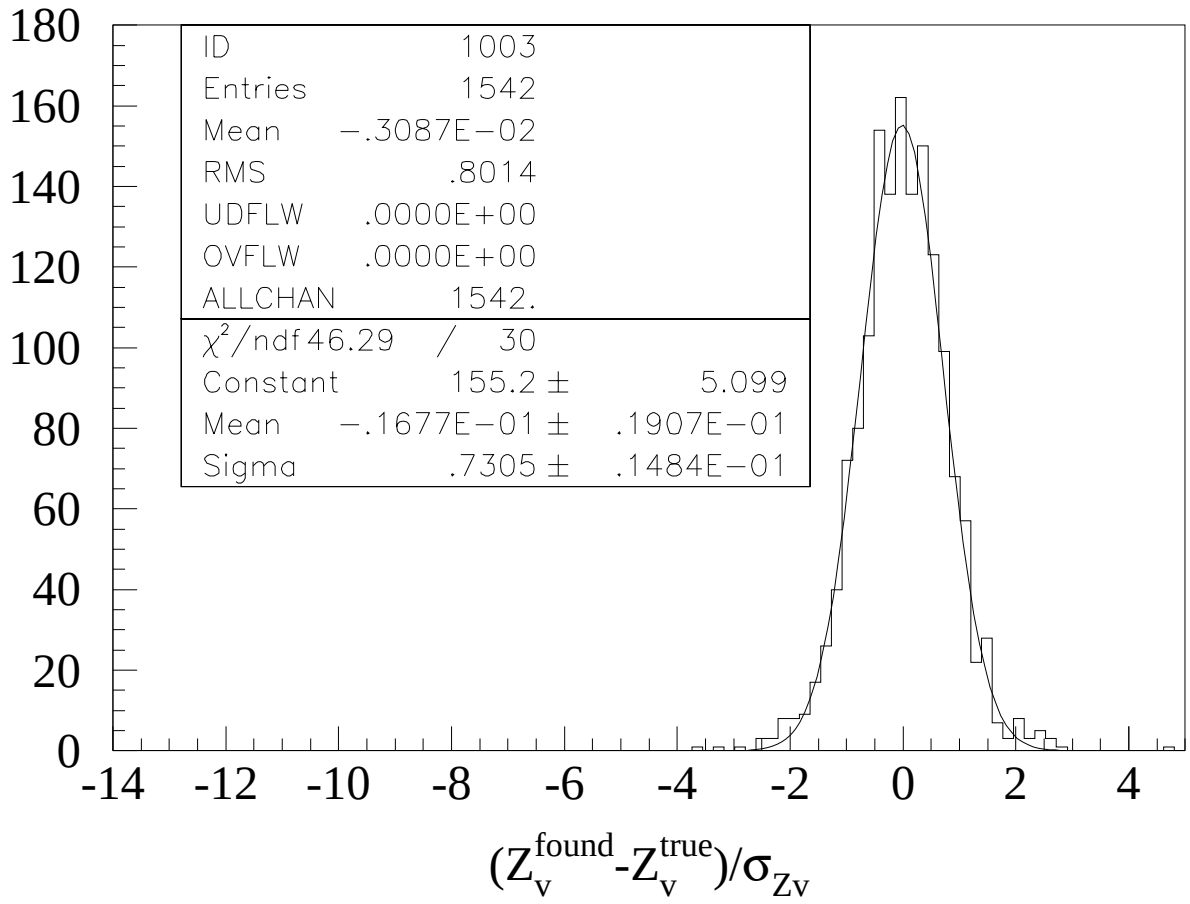


Figure 8.13: Pull variable for Z_v using a one dimensional extrapolation in η between the resolutions obtained from single muons without minimum bias pileup. The resolution was estimated from a momentum average over samples of μ's with p_t= 2,10,40,100 and 500 GeV.

8.8 Conclusions

Leptons are produced abundantly in the decays of $\tilde{g}$'s and $\tilde{q}$'s. This is emphasised by the fact that in a sample of 20000 events[9] corresponding to an integrated luminosity of $0.739 \cdot 10^5$ pb a total of 46327 μ's with $p_t{}^{\mu} > 10\,\mathrm{GeV}$ were found with only 16600 expected signal μ's.
Only 0.74% of all $\tilde{g}\tilde{g}$ pairs decay via $\tilde{\chi}_1^{+(-)}$ into same sign lepton pairs. Only this decay channel shows the majorana properties of the $\tilde{g}$. Same sign leptons from other branches of the $\tilde{g}$ decay chain or from other decay chains are not a signal for the decay of a majorana particle.
$\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ events have been produced and filtered by particle level cuts, yielding a sample of 1000 same sign muon events[10], corresponding to an integrated luminosity of $0.683\ 10^5\ \mathrm{pb}^{-1}$ which is equivalent to 70% of one year of full LHC luminosity. After impactparameter, p_t and isolation cuts have been implemented on the particle level and two and only two μ's per event have been required to pass the cuts the resultant sample of 1000 μ-pair events contains only 404 events in which both μ's have been produced through the signal decay chain. The total cross-section for the signal process of 0.09 pb could have given 15347μ pairs[11]. The large discrepancy between the possible and achieved signal size is partially due to the mass difference $m_{\tilde{g}} - m_{\tilde{\chi}_1^o} < m_W$ which forced the signal to proceed via a virtual W giving rather soft leptons out of which 44% do not pass the p_t cut of 25 GeV. 29% of the events are lost due to the presence of a third μ passing the cuts. The rest of the signal reduction is due to the isolation cut.
It seems necessary to ask for further correlations between the μ's in order to not quadratically loose signal with each cut and to more efficiently discriminate the signal agins the background which is intrinsic to the signal events.
1000 μ-pairs surviving the particle level cuts were simulated with the full ATLAS inner detector simulation and were reconstructed using the algorithms described in chapter 7. Before any post reconstruction cuts were applied the single track efficiency was found to be 98.9% with a fake track rate of 12.8% and a charge misidentification probability (F_{WS}) of 0.35%. Optimising the sample for minimal F_{WS} cuts on the number of tracks found in a given road and the separation of the Z-vertices of the two reconstructed μ's were implemented achieving $F_{WS} = 0.13\%$ at a μ pair efficiency of 77.1%. The detector resolutions for the 5 track parameters p_t, ϕ_0, θ_0, A0 and Z_{vert} obtained in chapter 7 have been used to calculate a normalised deviation of the reconstructed from the true track parameters. All parameters except Z_{vert}[12] where found to be perfectly Gaussian with σ's in the range from 0.77 to 0.88. The mean of a combined χ^2 deviation of the 5 track parameters from their true values was found to be 0.85 before and 0.79 after post reconstruction cuts.

[9] "soft sample" see table 8.6

[10] "hard sample" see table 8.6

[11] note that this is a factor 4 smaller than the previously quoted cross-section times the integrated luminosity because the cross-section given before also includes decays with electrons

[12] Z_{vert} is expected to deviate from a Gaussian shape towards a box distribution

Chapter 9

Summary

The next generation of pp-collider experiments proposed for the LHC will have, as one of their main components, a large inner tracking system. Due to the very high interaction rate and particle multiplicities at the LHC the most common solutions for large volume tracking, drift- or time projection chambers, can not be used. This fact, besides many others, lead to the choice of semiconductor tracking detectors for a substantial fraction of the ATLAS inner tracking detector. This predominantly silicon based detector system will not be comparable to silicon detectors used in high energy physics today. It will cover an area close to 80 m^2, consisting of nearly 100 million readout channels, dissipating several 10 kW of electric power. The system will need to have its readout electronics local to the detector element and require cooling of the detectors to temperatures below 0^oC. Furthermore it will be subjected to extreme radiation levels. One of the biggest differences to existing systems will be the fact that detector design and production of the ATLAS silicon detector system has to be oriented towards the principles of industrial serial production whilst making use of very delicate and expensive materials and at the same time having to maintain extremely small mechanical tolerances. The new aspects of development are demanded not only by the cost of the detector, but also by the amount of labour involved in its construction. The access to the inner detector cavity will be far less frequent than what is common for detectors existing now, requiring more stringent standards of quality control to be enforced.

Following the requirements of many physics processes, the extreme environmental conditions in the ATLAS inner detector cavity and engineering considerations related to the construction and operation of a very large scale high precision silicon detector, two prototype modules of an ATLAS barrel silicon tracker have been designed and built, together with their support structures and cooling system.
A construction process, suitable for a low cost large quantity production has been developed, avoiding labour intensive microscopic alignment steps by the use of edge alignment on precision cut silicon detectors. The tools and jigs for the complete process have been designed following the above strategy.
The module concept has been optimised towards maximal detector performance and highest modularity. The design presented here utilises a removable conductive rubber contact between the detectors and the electronic boards thus allowing the separate recuperation of both components in case of failures.
The design was aimed at an optimal decoupling of the thermal and mechanical variations occuring in the electronic boards from those in the silicon detectors. The use of separate cooling connections for detectors and electronics, as well as the non rigid connection between these components were targeted towards this decoupling. It allows the use of cheap

and standard materials for the electronic boards and enables the operation of the silicon detectors at temperatures below those of the electronics.
The mounting process of a module onto its support structure was designed to be possible without the use of robotic tools. The mechanical fixation is based on a self positioning mechanical connection between the support and each detector. The support structure provides an extra degree of modularity, subdividing the large cylinders of the ATLAS inner detector barrels into more easily manageable sections. These sections have to have full electrical, mechanical and thermal functionality without depending on their neighbour sections.

The prototypes, which were intended for thermal and mechanical measurements have been tested using ESPI[1] and IR-thermography in conjunctions with a system of miniaturised resistance thermometers incorporated into the structure of the prototype. The tests involved operation of the prototype under normal and extreme conditions such as cooling failures and excessive power consumption. During the tests all distortions and temperature distributions of the prototype remained within their specifications. The distortions of the modules measured in the $r\phi$ (z) direction were smaller than 3 (10) μm. The degree of mechanical decoupling between the detectors and the electronic boards was quantified by measuring the ratio of the maximal out of plane distortion in the detectors to those in the boards and was found to be $22\mu\text{m}/100\mu\text{m}$. The thermal decoupling is evident in the ratio of temperature changes in the detectors and the boards for the case of increasing the power in the boards to 6.25 times its nominal value which was found to be $1^oC/5^oC$.
ESPI has been used for the first time in the design of large area high energy physics tracking detectors. It has been developed to a state in which it is now routinely used for tests of ATLAS tracking structures. The spatial resolution achieved with a phase stepped in-plane setup was shown to be 13 nm over an area of 1311 cm^2 measured with an object distance of 3.5 m.
Image analysis and manipulation methods for noise reduction and phase calculation in ESPI interferograms have been implemented and were successfully used. ESPI offers a unique way to measure complete distortion fields at very high resolution without mechanically or thermally influencing the test object. A total power of only $\mathcal{O}(50\ \mu\text{W cm}^{-2})$ is usually induced by the laser in the test object. ESPI has been proven to be a versatile tool in the development of tracking detectors.
In conjunction with the ESPI measurements IR images were used to understand the temperature changes driving the observed distortions and to verify the prototypes thermal performance. The IR images and ESPI interferograms were also used as input parameters for a finite element simulation which was tuned to correctly describe the measurements of a test object.
A method of calibrating the indirect temperature measurement from the IR images with the temperature sensors visible in the image has also been developed. It uses an ambient model that includes a semi transparent window, all sources of radiation to first order reflections in the object and the window and calculates the transmission of the ambient atmosphere as a function of temperature, relative humidity and object distance.
Concluding from the experience gathered in the production of the prototypes and the results of the tests it can be said that out of plane distortions are very easily introduced into the thin detector modules by mismatches of CTE or temperature gradients. The only way of eliminating the source of these distortions is a strict implementation of a

[1]Electronic Speckle Pattern Interferometry

symmetric geometry and a symmetric heat flow. Since building a perfectly symmetric module is not possible, great care should be taken to control the remaining out of plane distortions for example by avoiding long free lever arms. In view of the above distortions it is also crucial to ensure a good and homogeneous contact between the cooling pipes and the modules because variations in the contact conductivity will lead to uncontrolled thermal gradients which are a possible source of distortions. The principle of symmetry has now been fully incorporated into the latest design of an ATLAS silicon tracker module.

In the second part of this work an algorithm using the silicon detectors in conjunction with the other subdetectors as a pattern recognition tool has been developed and tested. The application of the algorithm to the reconstruction of a demanding physics channel serves as the final test.
The central idea of the PR-algorithm was to combine the different capabilities of the discrete detectors and the continuous tracking detectors. The PR is initiated by external "seeds" indicating regions of interest found by other subdetectors such as the calorimeters or the muon system. The algorithm uses a combinatorial approach to choose hits from the discrete detectors because there are only a few hits and only a combinatorial approach can in principle assure that the true combination is at least tested in every case. For the TRT which will add up to 60 hits to a track, a combinatorial approach is not possible and not necessary since tracks will leave a clear signature in a histogram of the ϕ coordinates measured in the TRT if a correction for the track curvature is applied. The algorithm allows TRT information to be used at early stages in the decision processes governing the selection of hits in the discrete trackers.
The algorithm was initially tested by measuring the track parameter resolutions, efficiencies and fake track rates obtained for muon tracks as a function of luminosity up to $1.3{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$, muon transverse momentum and polar angle. It was found that the resolutions have Gaussian shape with very little non Gaussian tails. The resolutions are very close to their analytical limits and their evolution with luminosity is as expected from the occupancies of the TRT. The overall performance at full luminosity exceeds the specifications for all track parameters except at very large η.
The outstanding features of the algorithm are its stability against pile up and the good resolutions. Even at a luminosity of $1.3{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$which is 30% above the LHC design luminosity the track parameter resolutions were consistent with their expected analytical limits and the resolutions did not develop any significant non Gaussian tails. These are the key requirements for the reconstruction of events at the highest luminosities with additional need for accurate track measurements. One class of events with extreme multiplicities is the production of $\tilde{g}$'s and $\tilde{q}$'s. The $\tilde{g}$-pair production in addition requires very accurate charge sign measurement if the pair is supposed to be identified by a pair of same sign leptons indicating the decay of a heavy Majorana particle.

For the set of SUSY parameters used in this study a cross section of 12.84 pb was calculated for the production of $\tilde{g}$-pairs and a branching ratio of 0.7% for the decay of the pair into same sign muons or electrons yielding approximately 9000 signal events per LHC year.
The multiplicities of charged tracks in the tracking detector acceptance coming from the $\tilde{g}$-pair decay, has a mean of 262 with tails extending to 700. The photon multiplicities are similar with a mean of 274 and a maximum of 600. This is particularly interesting for final states including electrons due to the problem of fake electron candidates from early conversions.

The main source of background for same sign leptons in these events are the semi leptonic decays of heavy quarks. In order to suppress this background with an impact parameter cut a decay length simulation has been added to the event generator[2]. The other cuts applied at particle level to reduce the heavy quark background were an isolation cut and a cut on the number of final muons.
In order to simulate the influence of minimum bias pile-up, events passing the particle level cuts were superimposed with pile-up events equivalent to a luminosity of $1.0{\cdot}10^{34}\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ before they were reconstructed in the detector. A track reconstruction efficiency of 98.9% at a charge misidentification probability of 0.35% was found without any cuts after the reconstruction. To optimise the reconstructed sample, cuts on the z-separation of the two reconstructed vertices of each event and the total number of reconstructed tracks were applied. With these cuts the charge sign misidentification probability could be lowered to 0.13% at a track finding efficiency of 77%. All track parameter resolutions are correctly described by the parameterisations obtained from single tracks plus pile up.
It can be concluded that the pattern recognition and track fitting algorithm presented in this work is well adapted to the reconstruction of $\tilde{g}$-pair events. Misidentified opposite sign lepton pairs from $t\bar{t}$ events have up to now been considered to be the dominant background. With the good charge sign identification obtained here this is no longer true. The dominant background for $\tilde{g}$-pair events arises from the signal event itself in the form of a large number of leptons from the decays of heavy quarks in the many jets of $\tilde{g}$-pair events. Using only the criterion of the sign of charge of a detected lepton pair to identify the decays of $\tilde{g}$-pairs does not sufficiently discriminate this background against the signal. This is still true if an isolation cut is made and leptons from secondary vertices are vetoed. Therefore further correlations between the two signal leptons, such as their angular distributions should be studied.

Due to their extreme multiplicities and the large number of jets gluino pair decays depend more than most other physics channels to be be studied with ATLAS on the high granularity and precision of the silicon detectors. One of the essential prerequisites to providing this precision with an ATLAS silicon detector is the mechanical and thermal stability of the individual detector modules. Alongside tests of these properties on a detector module developed for this work, measurement techniques for the complicated task of measuring such distortions and temperature fields have been customised to fulfill the special requirements of testing large area silicon tracker modules and it has been possible to show that the module studied in this work is well suited for its tasks.

[2]ISAJET was used to simulate the production and decay of the $\tilde{g}$'s and $\tilde{q}$'s.

Bibliography

[1] F.Halzen, A.D. Martin, Quarks and Leptons, Wiley (1994),
D.H.Perkins, Introduction to High Energy Physics, Addison Wesley (1987)

[2] CDF Collaboration, Evidence for Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s}$ = 1.8 GeV, FERMILAB-PUB-94/097-E and Phys. Rev. Lett. **73** (1994) 225.

[3] ATLAS Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN, The ATLAS Collaboration, CERN/LHCC/94-43 (December 1994),

[4] see above p.82

[5] see above p.91

[6] see above p.95

[7] see above p.258

[8] see above p.245

[9] Figure taken from the ATLAS WWW pages engineering drawing section.

[10] T.Åkesson et al., Particle Idendification Performance of a Straw Transition Radiation Tracker Prototype, NIM A372:70-84, 1996.

[11] Private Conversation with Kenway Smith, University of Glasgow.

[12] Radiation Damage of Si-Detectors and Operational Projections for LHC Experiments, H.Feick, E.Fretwurst, G.Lindström, T.Schulz, CERN-SITP-TR-112, 10.01.1995

[13] The Performance Specification of the ATLAS Inner Detector, D.Froidevaux, M.A. Parker, ATLAS INDET-NOTE 46

[14] Z Resolution and Placement Precision Requirements for B-Physics of the ATLAS Silicon Tracker, Volker Büscher, Tony Weidberg, ATLAS INDET-NOTE 54

[15] Users Manual for AGEMA Thermovision 900 Series, AGEMA Infrared Systems AB, Danderyd, 1989.

[16] Private conversation with Dr. David Dibley, AGEMA Infrared Systems AB, Cambridge

[17] IPATREC: inner detector pattern recognition and track finding, Roger Clift, Alan Poppleton, ATLAS Internal Note, SOFT-NO-009, 14 June 1994

[18] Simulating Supersymmetry with ISAJET 7.0 / ISASUSY 1.0, H. Baer et al., in Proceedings of Physics at current Accelerators and the Supercollider, Ed. J.L. Hewett, A.R. White, D. Zeppenfeld, Argonne, 1993 - (ANL HEP CP 93-92)

[19] ISAJET 7.0, A Monte Carlo Event Generator for $p-p$ and $\overline{p}-p$ Reactions. Frank E. Page, Serban D. Protopopescu. Documentation file of ISAJET 7.0

[20] Theory and Problems of Fluid Mechanics and hydraulics, Ronald V. Giles et al., Schaum Outline Series, McGraw-Hill, New York, NY, 1989.

[21] Users manual for the Innova 90 Series, Coherent Lasers Inc., Santa Clara, 1984.

[22] Holographic and Speckle Interferometry, R.Jones and C.Wykes, Cambridge Studies in Modern Optics 6, Cambridge University Press, Cambridge 1983, 1989

[23] Fourier Fringe Analysis – A Review of Techniques, D.R. Burton, in Proceedings of "Fringe Analysis '94", p.33-40, published by: FASIG, The Fringe Analysis Special Interest Group 1994.

[24] "Installation et utilisation du compateur photoelectrique et interferential du Bureau International des Poids er Mesures",P. Carré, Meterologia **2**, 13 (1966)

[25] "Phase-Measurement Interferometry techniques", Katherine Creath, Progress in Optics XXVI, p.351-391, 1988

Acknowledgements

Over the four years in which the bulk of the work presented in this thesis was done I had the chance to work in two different universities. The EIV group of Claus Gößling at the University of Dortmund, were this work started and to a great extend was written up, was host to the design and construction phase of the detectors described in chapter 4. Over many month an extremely close collaboration with Joachim Fuss, the mechanical engineer of the EIV group, with additional feedback from the ATLAS SCT community lead to the design of the detectors. I thank him very much for his ideas and enthusiasm. Thanks also to Carl Haber, Richard Apsimon, Geoff Tappern, Mike Tyndel and many others within the ATLAS SCT group for taking the efforts and suggestions of a student serious. The design was realised in most parts in Dortmund with the very patient help from the local mechanical workshops who were excellent in their field and knew how to cope with the sometimes rather "different" views of a physicist in a workshop. The electronic boards have been designed by Bohdan Lisowski whom I would like to thank for this effort. They were manufactured in the CERN surface treatment laboratories. I owe many thanks to Claus Gössling for his support during all of my work, in particular for the introduction into the ATLAS community, the openness for discussions and many attempts to cure excessive use of optimism in making time scales. To all members of the EIV group my thanks go for a nearly infinite number of usefull suggestions to mechanical problems. Here in particular Thomas Weisse and his rigorous ways of approaching problems were of great help. I am also very happy to have found a great deal of understanding within the EIV group for the sudden disappearance of a large fraction of the groups power supplies, a PC, a DMM, a NIM crate and 80% of all relay scanners as well as a good part of one of the laboratories tools when I took all of this with me to Oxford. And to Teo Villet my thanks go for his unceasing efford to prove to all of us, not just that there are other ways of making a living than beeing a physicist but that most of them will make it, at least financially, a better one.

To Christian Lange I owe thanks for proving that handmade, artistic, high energy physics setups can work even if they do not look like it and for his diploma thesis in which he proved the feasibility of the connection technique described in this thesis by implementing it in a silicon based muon detection system.

Coming to Oxford in January of 1994 the welcome, both on private as well as professional grounds was so warm that I immediately felt at home. This atmosphere of friendliness, tolerance, respect and not at last academic enthusiasm was the foundation of the success of all work I have done there. To all friends that I have found there I would like to say that I will never forget this time.

It was the NAPL[3] which hosted me as an EEC fellow and supplied excellent working conditions. My great respect and thankfulness goes to David Howell who showed me a way into the wonderful world of lasers, optics, soldering irons, cooling systems, jig-plate, and not at least the Whitbread Race.

Constructing the numerous support and auxiliary systems necessary for the measurements in Oxford the mechanical and electrical workshops as well as the drawing offices were of utmost importance. My thanks go in particular to Collin Graham because he always knew where to find everything, to Mike Williams for his great willingness to accept all my workshop jobs and finish them until the day before, to Roger Giles and his crew for valuable

[3] **N**uclear **A**nd **P**article Physics **L**aboratory

advice and help with the day to day electronics problems and to Alan Holmes for his finite element calculations and his lessons on when and why to have a tea break.
Work in an office is nothing without the people in the office. The people in my office were simply the best. Many thanks for being my friends to Anish Grewal, Adrian Fox-Murphy, Daniel Buira-Clark and Richard Hawkings. Work without guidance from a more experienced point of view will always be slow and often be in vain. Tony Weidberg as my supervisor was my source of experience and guidance in the particle physics world. As my and my families friend he was father of Fiona, our sons first English girl friend and husband to Sheila who welcomed all of us many times.
The work presented in chapter 7 would not have been possible without the substantial and experienced contributions from Jim Loken. For this as well as for his cheerfulness I owe him many thanks. The code developed for chapter 7 is based on the software package *iPatRec* written by Allan Poppleton and Roger Clifft, with input from many other ATLAS collaborators. I would like to thank the authors for their willingness to explain the interior of *iPatRec* to Jim and myself.
Thanks also to Pete Gronbech (Oxford Particle Physics Computing), Andrew Samsun (RAL computing support) and Claus Wacker et al. (University of Dortmund, HEP-cluster) for making their platforms as pleasant an environment to work in as possible.
Special thanks also to BJ for letting me know that famous physicists are famous humans and to Peter White for introducing a stranger into super symmetry.
Thanks also to all the people who have read my thesis and given me good advice, here in particular to Renate Wunstorf who made more suggestions per page than anybody else I have ever heard of.

My greatest thanks I owe to Ute and Tobi. They supported me through this period like nobody else could have done. They were more than just tolerant with my working hours and without the balance between family and work my personal balance would have been long lost. The admiration I got for my lunch packets should have really been given to Ute.