ATLAS INDET-NO-085
25 January 1995

ATLAS SCT

Technical Proposal Backup Document

for the

SemiConductor Tracker (SCT)

for the

ATLAS Detector

at the

Large Hadron Collider at CERN

[Working Document - Version 1.0]

Contents

1 Physics and performance targets

1.1 Introduction

This document provides additional detailed information about the design and performance of the ATLAS Semiconductor Tracker (SCT) to backup the ATLAS Technical Proposal. The ATLAS Inner Detector consists of a barrel and forward straw system which includes a transition radiation function (the TRT), a barrel silicon microstrip tracker, a forward gas microstrip tracker (MSGC), a forward gallium arsenide microstrip tracker (GaAs), barrel and forward pixel detectors, and a vertexing layer for low luminosity running, either of silicon microstrips or pixels. The SCT consists of the barrel silicon, the forward GaAs, the barrel and forward pixels, and the vertexing layer. With regard to issues of tracking system performance, the pixel detectors will be discussed in this this document. Technical issues within the pixel system are discussed in a companion backup document [1]. An alternative to the MSGC's, based on the use of silicon strip detectors, is considered in this document. Some general issues of system performance will also consider other elements of the Inner Detector.

This backup document will evolve as results become available and the outstanding technology choices are made. It will be updated regularly with the latest version available on WWW.

1.2 Overview of Inner Detector physics specifications

A summary of the physics aims and requirements for the overall tracking can be found in Ref. [2] and was presented to the LHCC in the ATLAS status report of March 1993 [3]. The search for the standard model Higgs in the mass range between 80 GeV/c^2 and 1 TeV/c^2, for supersymmetry and for extensions of the standard model (heavy gauge bosons, strong symmetry breaking) are the primary goals at LHC and impose the basic requirements on the tracker - ie accurate measurements of high p_T tracks, location of primary vertices, identification of secondary vertices for heavy flavour tagging and B-physics. Adequate pattern recognition capabilities, at both low and the full design luminosity of LHC is an obvious requirement.

The detailed performance specifications of the ATLAS Inner Detector are gathered together in Ref. [4]. The following requirements need to be met, at minimum cost and without introducing more than 30% of a radiation length in the tracking volume -

- The reconstruction of isolated high-p_T leptons in the rapidity interval $\mid \eta \mid < 2.5$, with transverse momenta $p_T > 5$ GeV/c, with an efficiency of 95% and with a fake track rate of less than 1%. Additionally, to ensure that the leptons are isolated, there is a requirement to identify nearby low-p_T tracks, especially electron partners from photon conversions with p_T greater than 0.5 GeV/c, with an efficiency of 90%.
- A three sigma measurement of the electron charge sign up to 500 GeV/c. This requires a transverse momentum resolution of 30% at 500 GeV/c.
- Hadronic b-jets (from top quark decay) are to be tagged with an efficiency of 30% and a rejection of non-b hadronic jets of a factor of 10. This requirement relates to both the impact parameter resolution of the tracker and the two particle separation and pattern recognition capabilities within jets.
- Measurement of the z-coordinate of tracks with an accuracy of 1 mm.

At lower luminosities, other processes are of interest: top-quark physics, the supersymmetric Higgs sector (which requires τ-reconstruction), and CP-violating B-decays. These impose the following additional requirements on the Inner Detector -

- The tagging of b's with an enhanced rejection factor of 50 against non-b hadronic jets, whilst maintaining a tagging efficiency of 30% for b-jets (from top-quark decay).
- The reconstruction and identification of electrons with a p_T greater than 1 GeV/c with an efficiency of 70%.
- The reconstruction of all tracks from specific B decay vertices, above 0.5 GeV/c, with an efficiency of 95%. The reconstruction of K_s^0 decays within a radius of 50 cm with an efficiency of 90%.

- The measurement of θ with a resolution of 2 mrad.

In addition the Inner Detector should provide a level-2 track trigger which selects tracks with p_T greater than 20 GeV/c with an efficiency of 90% and a fake track rate of less than 10% per real track.

1.3 Precision tracking requirements

The overall concept of the ATLAS Inner Detector is to utilize both continuous tracking (TRT straws) and discrete precision tracking (solid-state detectors) to cover a rapidity interval $\mid \eta \mid < 2.5$. Physically, the detector is divided into barrel and forward regions. The requirements and the concept are essentially the same over the whole detector although the technology varies. In Figure 12 we show a layout, to be discussed in more detail later, for the baseline ATLAS Inner Detector.

The TRT provides a minimum of 36 closely spaced points per track at the outermost radius of the tracking volume where the occupancy and track density is smallest. It is optimised for pattern recognition, but if the systematics are well controlled, it will also provide a precision track point measurement at large radius (effectively 60 μm at a radius of 90 cm).

Additionally, the TRT provides rejection against hadrons because of its transition radiation function. At the cost of a small amount of additional material and radial space, and with dual-threshold electronics, electron identification is achieved.

The TRT is also foreseen to provide both a high-p_T track trigger and electron trigger.

The precision tracking layers inside the inner TRT radius of 60 cm are required to meet specified measurement resolutions and track reconstruction efficiencies. At the design luminosity of 10^{34} cm^{-2}s^{-1}, there will be a mean of about 20 overlapping events. Granularities of the order of 10^5 channels/layer/unit of rapidity are required to keep occupancies below 1%.

The radiation expected over the 10 year lifetime of the experiment precludes the use of silicon strip detectors (or MSGCs) below a radius of 30 cm. However, GaAs or pixel detectors thinned to 150 μm are expected to survive down to a radius of 10 cm. This constraint, due to radiation, is a significant limitation on the achievable impact parameter resolution of the tracker at low p_T and high luminosity.

The requirements on the precision layers, silicon strips and pixels in the barrel region; MSGC, GaAs strips and pixels in the Forward region, are :

- p_T resolution

  A transverse momentum resolution of 30% at a p_T of 500 GeV/c requires a **sagitta measurement of about 20 μm at a radius of 30 cm**, with magnetic field of 2 T, a beam spot constraint of 15 μm, and with the design performance of the TRT.

  To maintain this precision without the beam-spot constraint requires additional **20 μm measurements at a radius of 10 cm**.

- Alignment with tracks

  The final alignment and monitoring of the tracker will be achieved using tracks. The precision tracker must have more than 3 layers to allow an internal alignment and reduce systematics to below the intrinsic measurement precision (10 μm). The layers must be arranged so that the extrapolation of these systematic uncertainties to the TRT is less than 30 μm. In practice this requires **>3 20 μm measurements spread over the available radial space with one near the maximum possible radius (60 cm)**.

- z and θ resolution.

  The specifications translate into a requirement of a z resolution of 0.5 mm from the pixels close to the vertex and 1-2 mm from the strips at the maximum possible radius. However, to keep this specification for K_s^0 decays, requires measurements with a **precision in z of 0.5 mm at radii of 30 and 60 cm**.

- Pattern recognition

  While the TRT has excellent track-finding ability in the $r\phi$ projection, the precision tracker is required to reconstruct helices in 3D. Two approaches to track reconstruction can be envisaged. In the first, tracks are identified in the TRT and then linked to one or more genuine space-points such as those provided by the pixels. In the second approach, track finding is attempted independently within the precision tracker and then matched to the TRT. To achieve and understand the track finding efficiencies specified will require both approaches and requires, **>3 precision tracking layers, each providing two dimensional measurements**.

- Two-track separation

  Two-track resolution of 200 μm at a radius of 30 cm is required to keep the losses of tracks within b-jets from top below 5%.

- Triggering

  The requirement for a level-2 track trigger is for **4 equally spaced precision measurements**.

- Vertexing

  The vertexing requirements are driven by the ability to tag b-jets at high luminosity and reconstruct B meson decays at low luminosity. In the former case, given the constraints on layer radii due to radiation levels, 30 μm is taken as a target impact parameter resolution, for stiff tracks. This however requires a **track point measurement at a radius of 10 cm with a precision of 15 μm**. In the high luminosity geometry, for tracks below 10 GeV/c, the impact parameter resolution will be dominated by multiple-scattering. This can be remedied at low luminosity by the **addition of a thin vertexing layer at the smallest possible radius**. Such a layer should provide an **impact parameter resolution of < 25 μm for high p_T tracks and < 40 μm above 3 GeV/c**.

- Robustness

  Access to the ATLAS detector will be difficult and it is therefore a design principle that the loss of information from any particular subdetector will not be fatal to either the pattern recognition or resolution capabilities. The specification for the **number of precision measurements is therefore set to > 5 space-points**, to allow for inefficiencies and permit a test of the fit.

1.4 Design of the SCT

The design of the SCT requires a minimization of both materials and cost while satisfying the physics performance conditions listed above. There is an optimum to be found between minimizing the number of measurement layers and the demands on detector efficiency. If we assume that the maximum efficiency per measurement plane achievable, averaged over the lifetime of the experiment, is 95%, and try to satisfy all the above requirements, we arrive at the proposed layout of six precision layers. These are approximately evenly spaced between the minimum and maximum radii of 10 cm and 60 cm. The resolution of all layers is required to be roughly the same.

In the barrel region the baseline option for the precision layers are :

- Two pixel layers at radii of 11.5 and 16.5 cm ,chosen for their radiation tolerance, pattern reconstruction capabilities, and resolution (14 μm in $r\phi$ and 87 μm in z) though this depends on the readout scheme.

- Four equally spaced silicon-strip layers, each providing an $r\phi$ and a 40 mrad stereo measurement with $20 \oplus 10$ μm precision (20 intrinsic and 10 systematic). These six layers, together with the TRT, provide the required p_T resolution, even without the beam constraint. The choice of a small angle stereo solution (rather than an orthogonal or pad layout) is preferred for pattern recognition reasons, the suppression of ghost hits, and robustness against a missing $r\phi$ measurement.

  A readout granularity of 12 cm $\times$ (75 or 112) μm depending upon the readout scheme adopted, to be fixed on the basis of test beam results and cost. Both options are consistent with the specifications on occupancy.

Finally the proposed design is robust as defined above as -

- The momentum resolution degrades by less than 50% if either the inner pixel layer or the innermost silicon strip layers do not survive due to radiation levels, or if a track has no TRT measurement.

- The momentum resolution without beam–constraint or pixels degrades by a factor of two.

- The momentum resolution without beam–constraint or pixels degrades by a factor of two.

- The four stereo strip layers have some standalone track-finding capability. If the pixels were delayed this could be further improved by the addition of extra stereo layers at radii of 25, 35 ... cm.

The SCT project being defined by technology, rather than function is not responsible for the design of the baseline precision tracking in the forward region. However, similar arguments lead to the requirement of six precision measurements per track. The innermost measurements are again provided by pixels, with the balance coming from small-angle stereo detectors. The baseline choice over most of the rapidity range are MSGC's because of their cost advantage, but GaAs stereo strip detectors are preferred in the high radiation areas.

In the case that MSGC's are replaced in ATLAS by forward silicon detectors the basic configuration is described in Section 9.

2 Operating conditions

2.1 LHC environment and running conditions

In this section we discuss expected operating conditions at the LHC for the SCT. Our expectations are based upon simulations of particle production and radiation backgrounds at the LHC, and on various measurements of radiation effects on silicon detectors. In this discussion we assume a standard luminosity profile and operational period for ATLAS. The standard profile consists of an average luminosity in the first three years of operation of 10^{33} cm^{-2}s^{-1}, followed by the "full" luminosity of 10^{34} cm^{-2}s^{-1}for the following seven years. Each of these years is assumed to have a standard 10^7 seconds of beam time. The primary proton beams are assumed to have a momentum of 7.7 TeV/c each. The proton–proton total inelastic cross section at this center of mass energy is taken to be 70 mbarn. It is assumed that the occupancies and radiation exposures are dominated by the particle fluxes created in the primary p–p collisions and from secondaries directly associated with these particles.

2.2 Particle distributions

Multiplicities of primary particles have been calculated using several codes. Additionally the FLUKA [5] code has been used to calculate the albedo neutron fluence originating in the calorimeters and scattering back into the Inner Detector volume. Figure 1 shows the distribution of particles calculated using DTUJET [6]. Over most of the rapidity region the multiplicity distribution is flat, falling off beyond $|\eta| = 4$ and also dipping slightly in the central region. Figure 2 shows the average charged particle occupancy as a function of z for various radii. On average, 18 events per beam crossing occur in ATLAS at full luminosity. By combining the specific detector element sizes and the number of beam crossings integrated by a detector, the hit occupancies due to charged particles can be extracted. For example, with strips 12 cm long on a 100 μm pitch and with electronics sensitive to individual bunch crossings, a value of 0.03% occupancy on this plot corresponds to a hit occupancy of approximately 0.07% per beam-crossing.

2.3 Radiation fluences

Using Fluka, and the standard conditions defined above, calculations of the total particle fluences in the baseline detector design have been carried out [7]. These calculations include the presence of the integrated forward calorimeter. Figure 3 shows the expected particle fluence per year at full luminosity, as a function of the radius in the Inner Detector barrel region.

In the case of radiation damage to silicon detectors, it is convenient to normalise the particle fluences to an equivalent 1 MeV neutron fluence. Figure 4 shows the corresponding plots in this normalised "Damage Fluence"

The z distribution of the charged particle and neutron fluences and an equivalent damage fluence have been calculated. These are shown in Figures 5, 6 and 7. The equivalent damage is seen to be relatively insensitive to the z position.

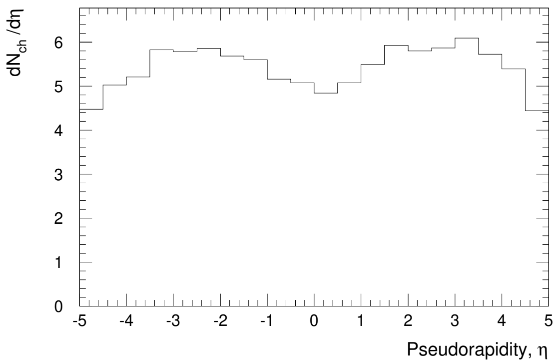


Figure 1: Charged Particle Multiplicity Distribution

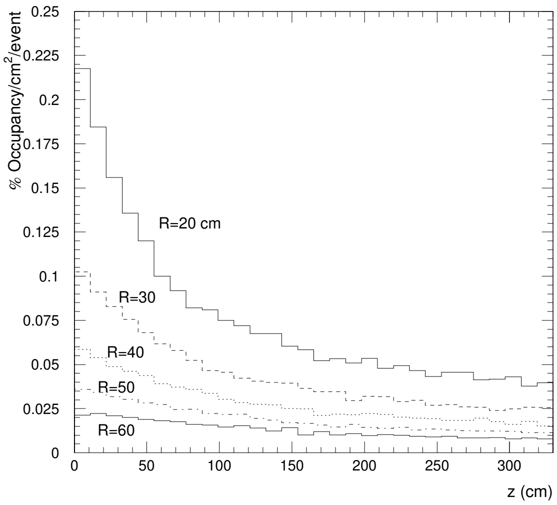


Figure 2: Charged Particle Occupancy

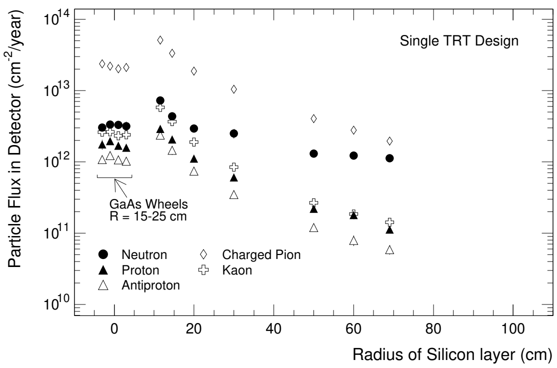


Figure 3: Particle Fluences in Barrel Region

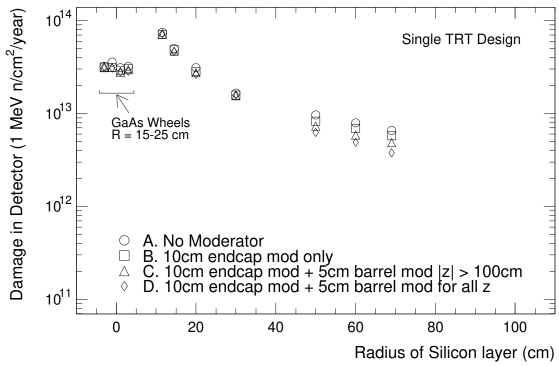


Figure 4: Equivalent 1 MeV neutron fluences in the barrel region.

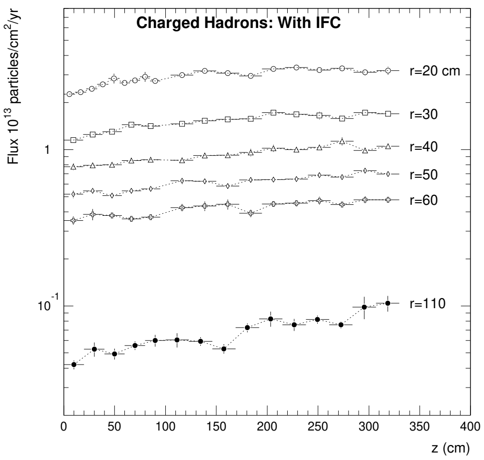


Figure 5: Z dependence of charged particle fluence.

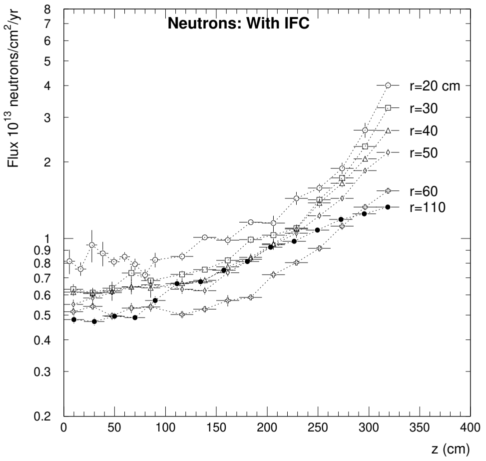


Figure 6: Z dependence of neutron fluence.

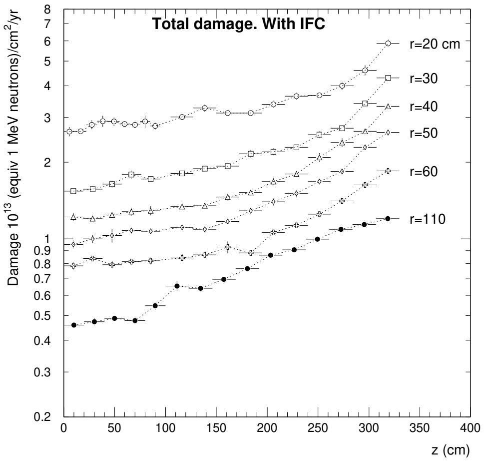


Figure 7: Z dependence of equivalent neutron damage fluence.

2.4 Radiation effects

Radiation damage to silicon detectors is categorised primarily into surface damage, which is depends on the detailed processing steps and detector design, and bulk damage which relates to generic properties of the silicon crystal itself. The surface damage may be manufacturer dependent and is therefore difficult to characterise in general. An empirical approach has been successful at reducing surface effects to levels significantly below bulk damage. which we assume will be possible for this application. For ATLAS, the primary concern will be bulk damage effects.

The most important manifestations of bulk radiation damage to silicon detectors are listed below.

- Damage to the crystal lattice creating states near the middle of the band gap in silicon resulting in **increased leakage currents**. These currents fluctuate due to the statistics of carrier generation and cause an increase in shot noise seen by the front end electronics. This extra noise can degrade tracking resolution and cause increased occupancy. Due, however, to the rather short integration times used at the LHC (25 ns beam crossing) and the reduced operating temperature, these shot noise effects remain manageable. In addition, the DC component of the leakage results in power dissipation in the silicon crystals. This current depends exponentially on temperature (the Block equation) for both damaged and undamaged devices. The potential exists for catastrophic thermal runaway if the silicon detectors themselves are not properly cooled.
- The creation of certain charged defects in the silicon crystal can result in changes to effective positive charge carrier concentration. This leads to a change in the **full depletion bias** and a type inversion for detectors which are produced of n–type material. For highly irradiated detectors the full depletion voltage can exceed the operational breakdown threshold making the device unusable. The type inversion effect will cause the location of the pn junction to move during the run. As a result, the high field regions will change position, leading to potentially unpredictable effects. If detectors have to be run partially depleted there can be effects on signal size and time development, and resolution. Coupled to these changes is the phenomenon of "reverse annealing" where the positive carrier concentration continues to increase with time after the irradiation [75],[76]. This can be controlled by cooling the detectors to near 0°C. Reverse annealing has been observed down to temperatures of 10°C, with a time constant approaching years. Lower temperatures are assumed to have a correspondingly longer reverse annealing timescale and a smaller effect over the 10 year operational lifetime.
- The creation of trapping sites in the silicon can lead to a **decrease in the signal charge** collected during the integration time.

Measurements of all these effects have been collated and re–analysed in a systematic way for ATLAS. Phenomenological models have been developed to parametrise the observed effects. The parametrisations have been applied to the calculated fluences to obtain estimates of the expected behaviour of silicon detectors at various positions in Inner Detector over the 10 year operation, assuming the standard luminosity profile [8].

For the reasons described above, when evaluating the projected performance of the detectors it is important to specify the temperature profile which the silicon detectors will be exposed to, for the entire operational period. Figures 8 and 9 show the change in leakage current and depletion voltage for each of the 10 years of running as function of radius. These plots assume the standard running scenario of 3 years at 10^{33} cm^{-2}s^{-1}and 7 years at 10^{34} cm^{-2}s^{-1}, the "Single TRT design, configuration B [7]", and the presence of the integrated forward calorimeter, as above. Detector temperature is 0° C. A 300 μm thick detector and an element area of 12 cm x 112.5 μm are used (150 μm and 300 μm x 50 μm for the 2 pixel layers). The current increase is directly proportional to the volume of each element. The depletion voltage is proportional to the square of the detector thickness. The damage functions used for the normalization to 1 MeV neutron equivalent damage are Lazo et al. and Van Ginneken [9] for energies below and above 20 MeV respectively, Aarnio/Huhtinen [10] for the pions, and Van Ginneken for all the others. For the neutrons, the neutron damage constants were used and all the charged hadrons were weighted with the proton damage constants. These values can be found in the literature [11].

*

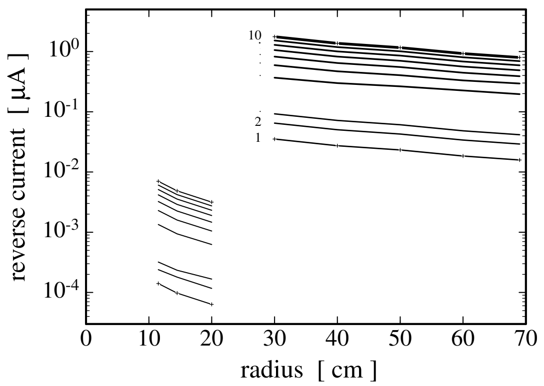


Figure 8: The leakage current per detector readout element in the SCT operating at 0° C. The calculations are for Pixels below 20cm and 12cm strips at large radii. The curves show the values at the end of each year of operation.

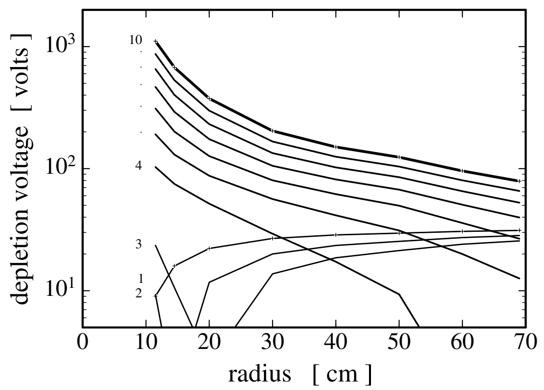


Figure 9: The voltage required to fully deplete 300 μm silicon operated at 0° C at the end of each year of operation.

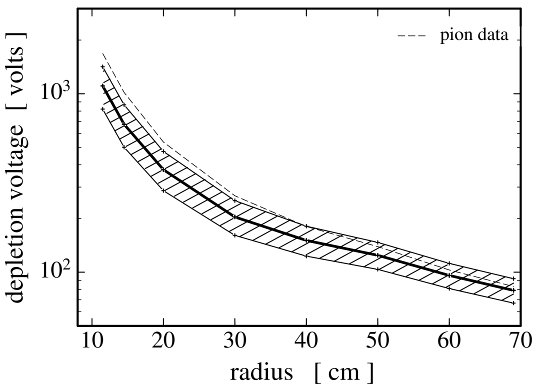


Figure 10: The voltage required to fully deplete 300 μm silicon after ten years of operation at 0° C. The cross-hatched region indicates the spread in the damage constants; the dotted line is the result using the preliminary pion damage constants .

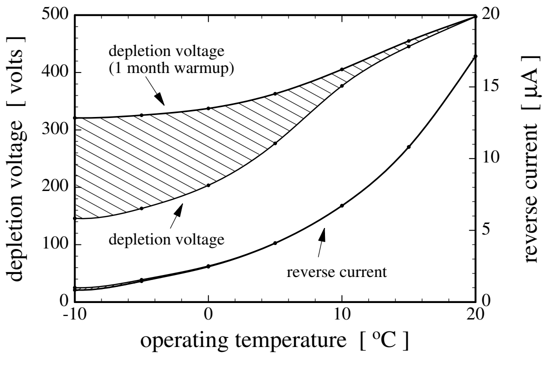


Figure 11: The dependence on operating temperature of the depletion voltage and leakage current after 10 years of operation for the detectors at a radius of 30cm.

3 Description of Layout

3.1 Overview

In this section we will discuss general issues relating to the layout of the SCT. Many of the details are covered in subsequent sections on module design (Section 6) and on mechanics (Section 7).

The full ATLAS Inner Detector occupies a cylindrical volume of radius 115 cm and half length 345 cm. It is composed of three independent sub-assemblies: the Barrel and two Forward detectors. The SCT is geographically a part of the Barrel, where it provides the precision tracking including a vertexing layer, pixel barrels, silicon strip barrels, and pixel endcaps. In the forward region, the only SCT detectors are the GaAs disks which compliment the precision tracking provided by the MSGC disks but at lower radius. The Barrel SCT is supported from the TRT, while the GaAs disks hang from the MSGC disks. The Barrel services are taken out through vertical gaps at the end of the Barrel, while those for the GaAs pass between the disks of the forward TRT modules.

The main parameters of the different SCT detector elements are summarised in Table 1. A longitudinal section of the complete Inner Detector is shown in Figure 12 while dimensions and coverage are indicated also in Figure 13

3.2 Position

3.2.1 Radial

In the ATLAS Inner Detector design all discrete tracking layers are positioned inside the TRT. This serves both to reduce the total cost of the (expensive) semiconductor detectors, move the TRT to a lower flux region to reduce occupancy, and increase the overall mechanical simplicity.

To enable good standalone track finding capability and track parameter measurements, allowing for inevitable detector inefficiencies, the ATLAS design specifies six precision layers, each providing space points with low fake probability. One layer is close to the TRT to permit precise extrapolation between the SCT and TRT and to maximise standalone momentum resolution. Inner layers are used in reconstruction of decay vertices and are consequently put at the minimum radius practical given radiation damage expectations. With the above requirements, and the restriction of six layers, track finding is most effectively achieved with a fairly uniform radial distribution of layers. The alternating layout of the stereo-strips (v-ϕ,u-ϕ,v-ϕ, u-ϕ) is chosen to reduce the rate of fake tracks formed from the combinations of ghost hits produced by close tracks and to improve the charge symmetry of the detector.

Pixels are well suited to the vertexing requirements owing to their good radiation tolerance, spatial resolution, and pattern recognition capability. The radius of the inner, high luminosity, pixel layer is the smallest at which a pixel detector can be expected to survive for ten years at LHC, while the separation provides an adequate lever arm for impact parameter measurements, extrapolation to the silicon layers, two particle separation.

For low luminosity operation, a vertexing barrel layer is specified inside the Pixel barrel. Based upon studies presently underway, this layer will either be constructed of pixels or of double sided silicon strips. The latter emphasizing low mass while the former emphasizing pattern recognition. Naturally this vertexing layer is placed at the minimum practical radius.

3.2.2 Longitudinal

The length of silicon strip and TRT barrels are equal. This simplifies the mechanical support of the SCT. The actual length is chosen to ensure that high momentum particles coming from the interaction point cross the detector planes at roughly normal incidence and to reduce the effect of the services gap on the transition radiation performance of the TRT and track finding efficiency in the overlap region.

The length of the Pixel barrel (and vertexing layer) is reduced to decrease the radiation lengths at high rapidity and because of the degradation in z resolution for tracks far from normal incidence. The barrel lengths are also chosen to account for the distribution of interaction points along the beam-line ($\sigma(z) \approx 5\,\mathrm{cm}$). The Pixel endcaps complete the rapidity coverage beyond the Pixel barrels up to $\eta = 2.5$.

In the forward region, the MSGCs are expected to survive down to a radius of 40 cm. Within this radius, radiation hard GaAs disks are positioned to ensure that the target of six space point measurements is met up to $|\eta| = 2.5$.

Vertexing barrels									
position (cm)	extent (cm)		detector	diode size	$\sigma(r\phi)$ (μm)	$\sigma(z)$ (μm)	module size (cm)	number	X_0 (%)
$r = 4$	± 35		pixel	50 μm $\times$300 μm	$10 \oplus 10$	87	1.6-2.0 $\times$6.4		1.0
or									
$r = 6$	± 26		ϕ-strips	25μm $\times$26 cm	10		3.3×26		0.5
			z-strips	50μm $\times$3.3 cm		20			
High luminosity barrels									
position (cm)	extent (cm)		detector	diode size	$\sigma(r\phi)$ (μm)	$\sigma(z)$ (μm)	module size (cm)	number	X_0 (%)
$r = 11.5$	± 35.0		pixel	50 μm $\times$300 μm	$10 \oplus 10$	87	1.6-2.0 $\times$6.4		1.1
$r = 16.5$	± 45.0		"	"	"	"	"		"
$r = 30$	± 82.1		v-ϕ 40 mrad stereo	75 or 112.5 μm $\times$12 cm	15	770	6×12	476	1.6
$r = 40$	"		u-ϕ 40 mrad stereo	"	"	"	"	630	"
$r = 50$	"		v-ϕ 40 mrad stereo	"	"	"	"	784	"
$r = 60$	"		u-ϕ 40 mrad stereo	"	"	"	"	952	"
Pixel endcaps									
position (cm)	extent (cm)		detector	diode size	$\sigma(r\phi)$ (μm)	$\sigma(r)$ (μm)	module size (cm)	number	X_0 (%)
$z = \pm 49.9$	11.5	21.3	pixel	50 μm $\times$300 μm	$10 \oplus 10$	87	1.6-2.0 $\times$6.4		1.0
$z = \pm 55.4$	"	"	"	"	"	"	"		"
$z = \pm 79.9$	"	"	"	"	"	"	"		"
$z = \pm 85.0$	"	"	"	"	"	"	"		"
GaAs disks									
position (cm)	extent (cm)		detector	diode size	$\sigma(r\phi)$ (μm)	$\sigma(r)$ (μm)	module size (cm)	number	X_0 (%)
$z = \pm 155.7$	20	35	u-v $\pm 0.5°$ stereo	50 μm $\times$7.6 cm	10	1200	2.7×7.6		2.5
$z = \pm 182.5$	29	44	"	"	"	"			"

Table 1: Parameters of SCT layers. The resolutions are target resolutions, the diode sizes required to achieve these depend on the readout scheme, and those given correspond to direct readout.

3.3 Resolution

The resolution of each track point measurement and the layer radii are the fundamental parameters determining performance of the tracker. In specifying the track point resolutions a number of factors must be weighed. A real system will have systematic effects such as misalignment. Material will lead to multiple scattering at lower momenta. Power dissipation, cost, channel count, and material, all tend to increase with better resolution. For these reasons it is important to not over specify the track point measurements. Conversely, poor track point resolution will undercut the performance of the tracker. The resolution versus cost argument is applicable in a particular range of segmentation but not universally since the basic silicon wafer costs are almost independent of segmentation. Certain techniques have been proposed to obtain improved track point resolution with less segmentation. However these may not perform as well for two particle separation.

Considering these various factors, ATLAS specifies that each measurement in the SCT silicon strip layers should have an accuracy of 20 μm (statistical) and 10 μm (systematic). Quantitative justifications can be found in Section 10. The strip pitch required to achieve this depends upon the final choice of readout scheme. The combination of two measurements (u-ϕ or v-ϕ) leads to an $r\phi$ resolution of $14 \oplus 7$ μm (depending on the description assumed for the systematic errors). The effective resolution in z (when fitting tracks which include other measurements of rϕ) is 500 μm. The stereo angle has been chosen to provide adequate measurements of track parameters in the r–z plane for mass reconstruction in B decays, while keeping the fraction of ghost space-points at an acceptable level. The specified value is 40 milliradians.

For the vertexing layer and the pixels, which are all at smaller radii, we specify improved track point resolution, 10 μm in ϕ and 87 or 20 μm for z for pixels and strips respectively.

3.4 Modular design

The mechanical support for each silicon strip layer of the SCT is discussed in Section 7. The fundamental sub-unit of the silicon strip SCT is a *module*. Each layer is tiled with a set of modules. A module consists of:

- Active detectors - either one or two pairs of semiconductor wafers.
- Front–end electronics.
- Read-out controller (ROC).
- Fan-ins and connectors.
- Opto-electronics for transmitting data off detector.
- Interface to permit cooling.

A modular design has a number of advantages. These include ease of assembly, repair, and testing, electrical isolation, and unity of design. A modular design may incur some overhead in material but this is far outweighed by the other advantages. Also, with a modular design, parallel production lines can exist at different facilities.

Pixel modules are described elsewhere [1]. In what follows, the modules for the silicon strip detectors are described, more detail can be found in Section 6.

The strip detectors have both ϕ strips and u or v strips. The latter are rotated by ± 40 mrad in the detector plane with respect to the ϕ strips. While the u-ϕ and v-ϕ modules are very similar, there may inevitably be small differences in the module design or in the method of support. Strips of length 12 cm are created by wire-bonding two 6×6 cm^2 silicon wafers. Two measurements are made by each module by using either double sided detectors or by gluing pairs of single sided detectors back to back (see Section 6).

The silicon wafers will be 300 μm thick. The back to back single sided module will thus be 600 μm thick. Each silicon wafer will have an inactive edge of 1.2 mm: 0.5 mm representing the cutting edge, and 0.7 mm for a multiple guard ring structure. Modules will be overlapped in z by 3 mm, with a radial spacing of 1 mm. In rϕ, it is required that the active overlap should be of the order of eight strips to permit adequate module to module alignment. This can be achieved with a geometrical overlap of about 3 mm. The only geometrical inefficiencies for high p_T tracks arise from the inactive region in the vicinity of the wire-bonds between two silicon wafers.

The tilt angle of the modules in the rϕ plane, relative to the radial direction, will be chosen to optimise the spatial resolution which can be achieved allowing for the Lorentz effect and charge sharing on the readout electrodes. A typical value is 6° but can depend upon the choice of detector substrate and technology.

The modules will be cooled to keep leakage currents down and control anti-annealing effects on the depletion voltage.

3.5 Material distributions

Detailed evaluations of material distributions for the silicon strip layers can be found in Section 6 and in [12]. In Figure 14 the material distributions of various components of the SCT are shown as a function of rapidity. These include electronics on the modules, support structures, and services running along the support structures. The distributions show a csc θ distribution due to the angle of incidence with an additional rise due to the cumulative increase of cables as a function of $|z|$. The support rings for the pixels and B-vertexing layer (pixel) are clearly visible. In Figure 15, the material distribution in the Barrel for the SCT is shown. The lower figure shows the effect of the services which pass through the gap between the Barrel and Forward regions.

3.6 Technical specifications on the strip layout

In this section we present the basic technical specifications which apply to the layout and construction of the SCT silicon strip layers.

- The basic sub–unit of the Barrel silicon tracker will be a module. A module will consist of silicon strip detectors, front–end chips, data transmission hardware, and an interface to the cooling system.
- Modules will be individually dismountable.
- The basic module dimensions are 6 cm (in ϕ) and 12 cm in z.
- The module will consist of either two pairs of back to back single sided detectors or one pair of double sided detectors.
- The following values apply to the alignment of detectors within a module:

  Placement of wafers on module: rϕ: ± 12 μm; z: ± 25 μm.
- The following numbers apply to the construction tolerances on module placement during assembly.

  Placement of module on support cylinder: r: ± 250 μm; rϕ: ± 50 μm; z: ± 100 μm- in ATLAS reference frame.

  Survey of module on support cylinder: r: ± 50 μm; rϕ: ± 25 μm; z: ± 50 μm- in ATLAS reference frame.
- Modules will be tilted by 6° (to be further optimised).
- Modules will overlap in z by 3 mm and in ϕ by 8 strips.
- Modules will be mounted on a support structure. Both the modules and the mechanical support structure are to be stable, at constant temperature, to ± 10 μm.
- Modules will be cooled during operation to an absolute average temperature 0° C. The temperature gradient will be less than 5° C.
- Electrical services will be provided on low mass cables such that the failure of one module is localised.

3.7 Expected radiation effects

Given the proposed layout, and the known effects of radiation damage, it is possible to predict the deterioration of the tracker performance over the course of the ATLAS running scenario. By performance we mean resolution, track finding capability, occupancy and so forth. There are also tentative issues such as performance with and without fully depleted detectors. These calculations require a full simulation of the tracker folded with the radiation effects. In the present version of this document no results will be presented although we forsee developing this discussion further in a future version.

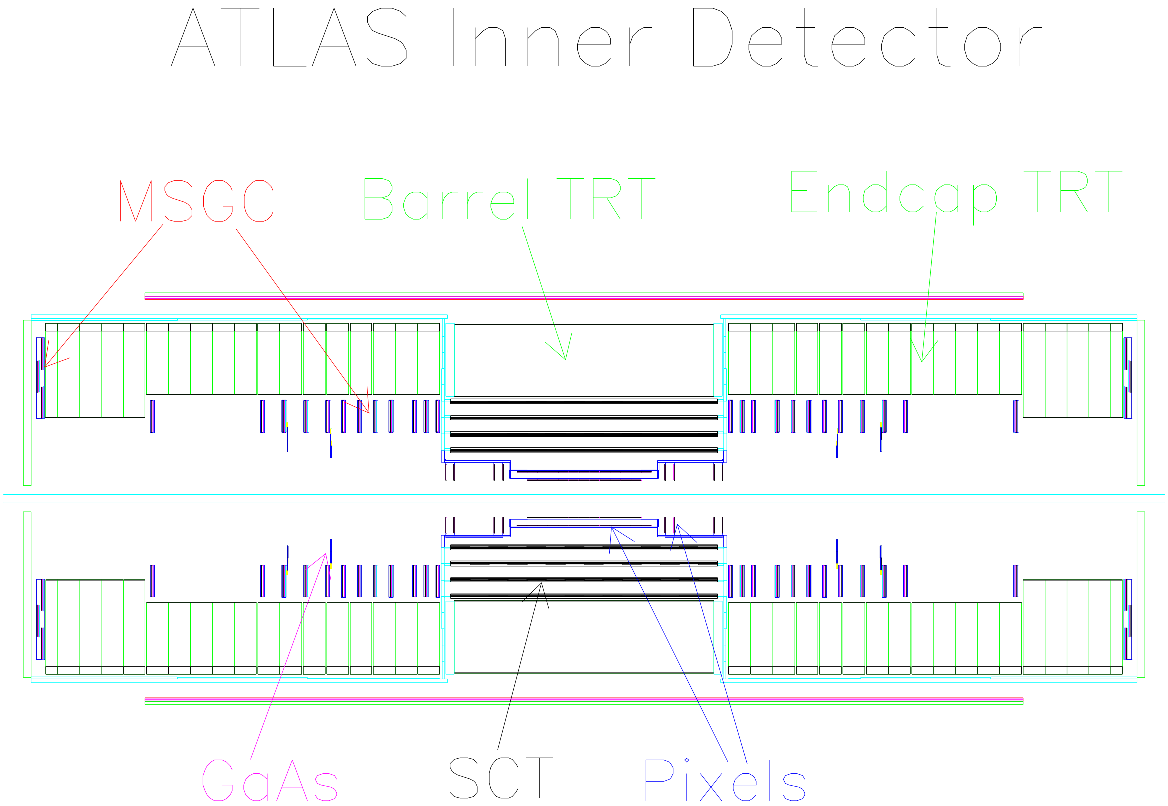


Figure 12: Longitudinal view of the ATLAS Inner Detector. The radius is 115 cm and the half-length 345 cm.

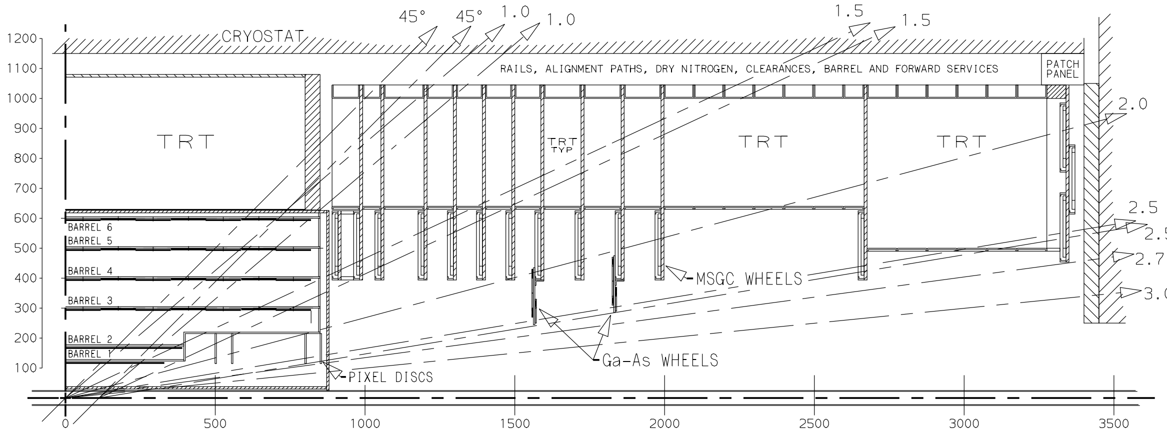


Figure 13: Quarter view of the ATLAS Inner Detector indicating dimensions and coverage contours.

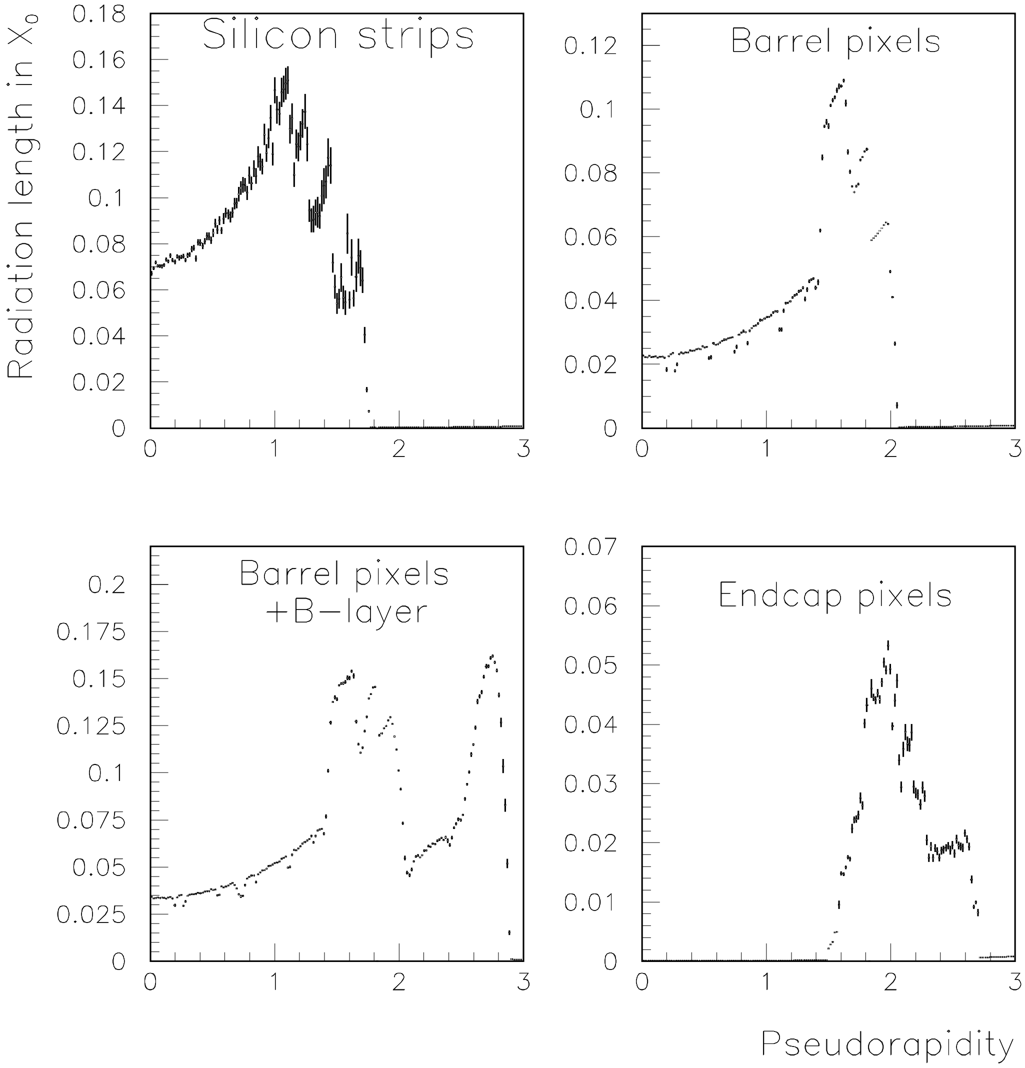


Figure 14: Material distributions in (a) silicon strips, (b) barrel pixels, (c) barrel pixels with B-vertexing layer (pixel) and (d) endcap pixels.

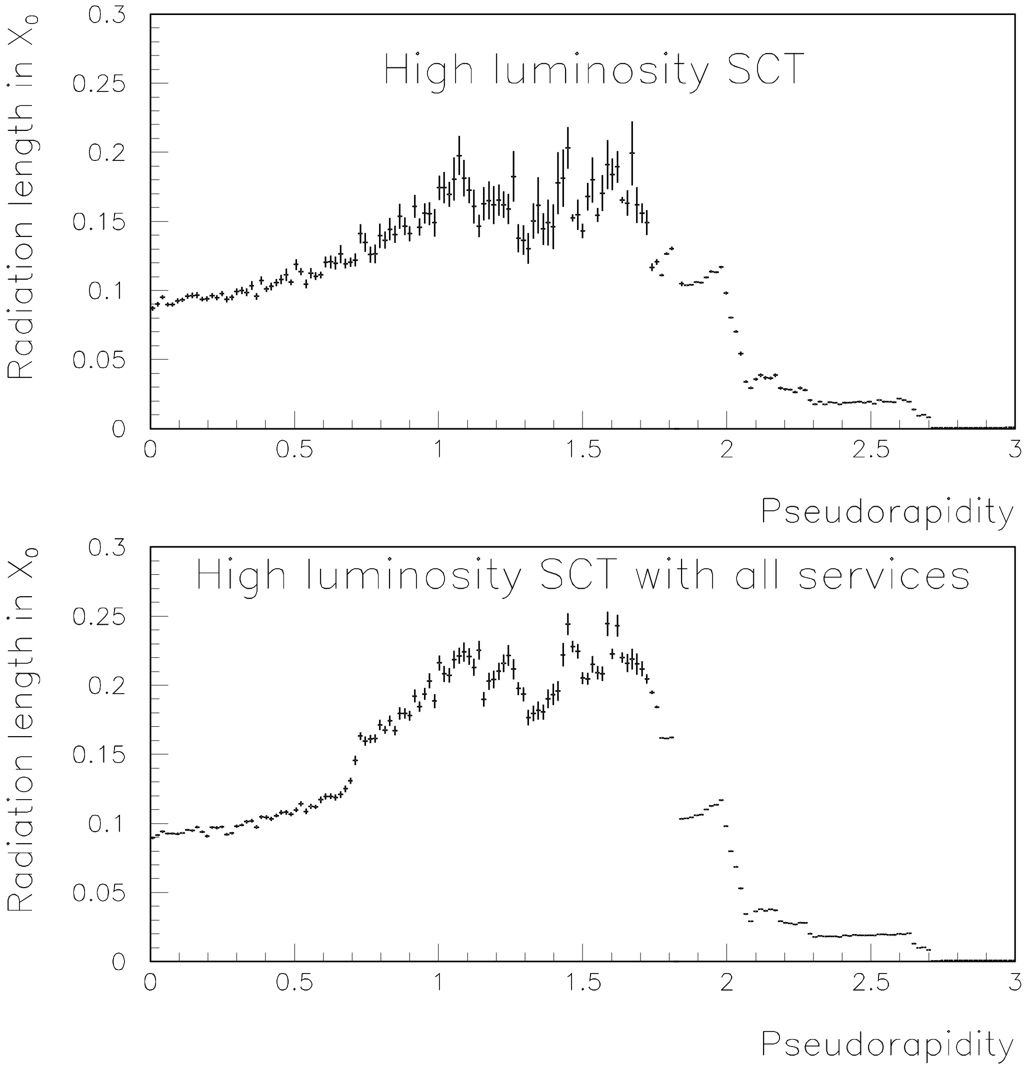


Figure 15: Material distributions in the SCT Barrel: (a) just from the barrels themselves and (b) including the services running between the Barrel and Forward regions.

4 Detectors

In this section we will discuss issues related to the choice of strip detector technology and the fabrication of detectors for the ATLAS SCT. We will concentrate on silicon in the present version of this document. Issue relevant to gallium arsenide strip detectors are reviewed in the latest report of the RD–8 collaboration [13].

4.1 Discussion of technical issues

Although development of strip detectors for ATLAS is in a fairly advanced state, there are still several principle options under consideration. Each have their advantages, drawbacks, and uncertainties as will be discussed below. A final decision will be taken on the basis of performance and cost by the end of 1995.

The principal issues are the following:

1) double sided or back to back single sided detectors

2) capacitive charge division or individual strip readout

3) capacitively or directly coupled readout

for back to back single sided detectors:

4) p-strips (collection of holes) or n-strips (collection of electrons)

5) n or p-type bulk silicon substrate

6) isolation of n-strips by by large a area p-type compensation implant or by p-type insulation strips

for capacitively coupled readout:

7) biasing by reach through or polysilicon resistors

8) strip geometry (pitch and gap width)

Many of these items are correlated, so that the choice on a particular subject may narrow the choice on other options. In addition we will depend upon technical capabilities and experience of manufacturers, so that a certain degree of freedom in meeting the requirements of the experiment should be left open. Nevertheless, general considerations leading to stable cost effective detector concepts can be made. For the present we must tolerate a few basic uncertainties as e.g. the behaviour of detectors built on p-type silicon is not well known. It is the aim of the detector prototyping program to collect the missing experimental information within the next year.

At present the following considerations apply:

1) Double sided versus back to back single sided detectors.

- Assuming a wafer thickness of $300\mu m$ (which seems necessary for adequate signal/noise ratio) use of double sided detectors saves $\approx 0.4\%$ of radiation length per layer (0.3% from silicon and 0.1% conducting silver glue).

- Using double sided detectors, heat generation by radiation induced current is smaller by the total thickness ratio at equal operating voltage.

- Front to back alignment of detectors is automatically obtained with high precision ($< 5\mu m$) with double sided detectors. This precision exceeds that which is required for ATLAS, assuming well understood survey and metrology is used during construction.

- Running double sided detectors in partial depletion may (after type inversion or on p-type substrate) lead to a loss of information from p-type strips. Partially depleted operation is possible with n-type strip back to back single sided detectors. One suffers however in S/N and in signal rise time.

- Double sided detectors have the additional feature of requiring operation with different electronics ground potential (close to the respective strip potential) on the two sides of the wafer. This is required for DC and also for AC coupled detectors in order to avoid breakthrough and other problems at high depletion voltages. The requirement should not pose fundamental problems but may bring about practical complications in power supply configuration and in requiring perhaps separate optical signal coupling for the two detector sides. This may result in an additional material burden.

2) Capacitive charge division or individual strip readout. This question goes beyond detector issues as it involves channel counts, simplicity of electronics, measurement precision, and pattern recognition.

- The assumption that individual strip readout poses less stringent requirements on signal/noise than charge division readout is not necessarily true. The noise requirements are governed by the demand of full efficiency with charge splitting and low noise hit probability leading automatically to a sufficient measurement precision in charge division readout. Charge splitting occurs for both types of readout due to diffusion and track inclination.
- Individual strip readout requires finer strip pitch in order to obtain similar measurement precision. The higher channel count therefore has to be compensated by lower individual channel cost.
- Introduction of a cluster threshold will reduce the noise hit rate. It is possible only if analog signal information is available.

3) Capacitively or directly coupled readout.

- Capacitive signal coupling protects the electronics from channel to channel and time variation of dark current and therefore also from some sources of pedestal variation (although some preamplifier designs can compensate for this by feedback or stabilization circuits). It also provides DC insulation between inputs of different channels thus preventing input currents due to amplifier input voltage offsets.
- With capacitive signal coupling the implanted strip potential and the detector outer surface (aluminium strip) potential can be chosen independently. The oxide charge buildup due to irradiation can be compensated partially by adjustment of the relative voltage.
- Direct signal coupling will lead to simpler (and therefore also less expensive) detectors. The expected price difference is however small if reach-through rather than polysilicon biasing is chosen for capacitively coupled detectors.

4) p-strips (collection of holes) or n-strips (collection of electrons) in single sided detectors.

- p-strips may not allow partially depleted operation after type inversion or with p-type substrate.
- n-strips should allow partially depleted operation, however at the price of reduced signal/noise.
- Electrons are collected faster than holes, giving a faster signal on n-strips. However, the full signal will only be reached when all carriers are collected. Prior to that, the carriers will induce a signal of opposite polarity spread over several strips.
- n-strips on n-bulk probably require double sided photolithographic processing, making the technology similar to that of true double sided detectors and thus more expensive.

5) n versus p-type silicon substrate.

- As low doped (high resistivity) n-type silicon inverts to p-type after exposure to the LHC radiation environment, it seems therefore natural to use p-type material right from the start. This way the junction position does not move from one wafer side to the other when type inversion occurs. Not having to provide for this change simplifies design and technology of the detectors. This is especially true for single sided detectors with n-strips (see also item 7 below)).
- Despite these general considerations restricting options to p-type material may be premature at this time as experience with p-type silicon is rather limited. This is also true for radiation damage studies. High priority should therefore be given to basic investigations of p-type detectors.
- Starting with p-type material may lead to a higher final operating voltage after irradiation. This will be the case if the original p-dopant concentration is greater in p-material than in n-material. Silicon of any conduction type always contains both types of dopants. The effective doping is given by the difference between the two types of dopants.

- An open question is the price of p-type material since material with minimal boron doping is desired. This may involve a more complicated crystal growing procedure. The question will be addressed in contacts with industry. Raw silicon wafer cost should be a minor fraction of fully processed detector cost.

6) Isolation of n-strips by a large area p-type compensation implant or by p-type insulation strips.

- A large area p-type compensation implant reduces the production complexity as it avoids one photolithographic step. It also allows narrowing of the gap between strips which can be an advantage in the case of capacitive charge division.
- The choice of implantation parameters (energy and dose) may be critical as one has to anticipate the radiation induced oxide charge buildup and simultaneously wants to avoid too large an overcompensation. The proper balance may not easily be achievable.
- Reach through bias structures are more easily implemented with large area compensation implants.

8) Biasing by reach-through or polysilicon resistors.

- Biasing integrated into detectors is required for capacitively coupled readout.
- Polysilicon resistor biasing has been proven to be radiation hard. It requires additional processing steps as compared to strip detectors with reach through biasing.
- Reach through biasing has been shown to work for both p and n-strips. It is also expected to be radiation hard. Experimental results exist so far only for p-strip biasing. An interesting property of reach through biasing is the nonlinear (approximatively exponential) current–voltage characteristics. This may bring about a certain protection of detectors against accidental radiation bursts as the radiation generated charge carriers are carried away faster and a smaller voltage change will appear at the electronics input.

8) Strip geometry (pitch and gap width).

- Strip geometries have to be chosen such that the capacitive load to the amplifier remains low enough to obtain a sufficiently high signal/noise ratio. In the case of individual strip readout this requires rather narrow strips and large gaps. Quantitative information from computer simulations relating geometry and capacitance can be found in reference [14]. A potential problem with small strip geometry is the role of the conducting layers in between the strips which can be formed either by electron accumulation or inversion layers (on the p-strip side) or the insulation implant (on the n-side region). It may influence depletion voltage, electric field distribution (increased danger of breakdown), charge collection speed and signal induction to non–neighbouring strips during charge collection. These questions still have to be studied.
- For capacitive charge division readout, reduction of capacitive load is achieved rather by increasing of the number of charge division strips than by widening the gap in between strips. This way the influence of the badly controlled gap area on detector performance is reduced and more predictable detector behaviour is expected. Quantitative results are again provided in reference [14].

4.2 Vendor Specification

A critical aspect of the detector development process is to understand the capabilities and interests of commercial vendors of silicon strip detectors. It is also important to gather information on detector costs since at some point in the design process certain compromises will be required. To this end we have prepared a technical description of the detector needs for ATLAS to be distributed to vendors. The material distributed is not a formal call for tenders. We want to know who is interested in further discussion and basic ideas on pricing. The latter is clearly non-binding and confidential if requested. In this process we chose to narrow the consideration to the following basic configurations:

ATLAS A: Single sided, capacitively coupled, n strip readout.

ATLAS B: Single sided, directly coupled, n or p strip readout.

ATLAS C: Double sided, capacitively coupled, n and p strip readout.
Further details on these specifications are given in Appendix B.

4.3 Technical Implementation

Past development efforts for detectors meeting the requirements of ATLAS have been done together with industry and in the semiconductor processing laboratory of one of the participating institutes of ATLAS. It is difficult to describe here all the work which has been performed. Much of the work involved general investigations of properties of single detector elements, leading to a large fraction of the knowledge which has been compiled in the first part of this section. Here, some detector concepts will be proposed which have been conceived with the aim of simple and therefore cost effective production, meeting nevertheless the demands of the experiment. Three types of detectors will be proposed which correspond to the specifications given to industrial manufacturers. In addition possible variations in the design and in production methods will be discussed.

Common to all types of detectors is the need to run reliably at fairly high voltage, thus requiring breakdown protection structures at the detector rim. Breakdown protection at the detector rim is accomplished by assuring a gradual drop of the potential at the outer insulator surface from the detector bias voltage towards the bulk potential near the detector rim[15]. Such a task can be accomplished by the introduction of additional processing steps such as e.g. covering the insulator with a very high resistivity conducting layer. In the examples proposed below the protection structures have been designed without complications of technology to keep production costs low. Not included in the description of technical solutions is passivation of the detectors which might be necessary for protection during the assembly procedure. Assuming that single sided detectors are only passivated on the strip side this production step makes no difference in comparing the options.

Double sided strip detectors with capacitive charge division and capacitively coupled readout
This type of detector is presented first, since it is the most complicated one and provides the optimum information to be obtained from a strip detector, combined with smallest radiation length. Its use probably requires full depletion. Savings in complexity of other solutions can be judged by comparison with this solution.

The cross section of a detector is shown in Figure 16. Here it is assumed for simplicity that strips on the two wafer surfaces run parallel, while in reality they have a finite angle with respect to each other.

The cut along the strips is shown on the left. Implanted strips (p on top, n on bottom) are not electrically connected externally and separated from the aluminium strips by a dielectric which may be either simple silicon oxide or a more complicated thin layer structure as e.g. silicon nitride on top of silicon dioxide. Reach through biasing from the bias strips positioned at close distance ($\approx 5\mu m$) from the strip ends keeps the strips within 10 to 20 VDC from the respective bias voltages of the two sides before irradiation. The detector guard rings assure a gradual drop of the bias voltage towards the rim of the detector. The potential of these guard rings is again provided by a reach through mechanism. An important feature of this guard ring structure is the definition of the potential on the outer insulator surface. This is accomplished by connecting the metal part with the implanted part of the guard ring.

Notice the difference between p- (top) and n- (bottom) side. A p-type implant around the strips or a flat p-type compensation implantation is needed at the n-side. Its purpose is the (over) compensation of the always positive oxide charge. Without it one would encounter a shorting of all n-type implants.

The concept presented in figure 16 should work for both n- and p-type silicon substrate. For n-type substrate the lower guard ring structure initially has no function as the rim and undepleted bulk will assume roughly the positive bias voltage. The voltage drop will occur at the top (junction) side. Once type inversion is reached the junction moves to the bottom and the lower guard ring structure becomes active.

Despite the high functionality of the device the production technology stays fairly simple. The following major production steps are required:

1) thermal oxidation

2) boron implantation through oxide of one wafer side (p- compensation implant)
3) contact hole (oxide) opening on p-side
4) contact hole (oxide) opening on n-side
5) boron implantation on p-side (p+ strips and guard rings)
6) phosphorous implantation on n-side (n+ strips and guard rings)
7) tempering (activation of dopants)
8) aluminisation p-side, structuring
9) aluminisation n-side, structuring
10) sintering (metal-semiconductor contact)

The labour and cost intensive steps are those involving photolithography (3, 4, 5, 6, 8 and 9) and those involving implantations (2, 5, and 6). The whole detector production involves

- 6 photolithographic steps and
- 3 implantation steps

It is not completely trivial to avoid all defects in the insulating oxide. Problems may arise in the oxidation process or in photolithography. Both problems can be avoided by using instead of a single oxide insulator a two layer structure consisting e.g. of silicon nitride deposited on top of thermal oxide and using a two step etching technique for contact hole opening.

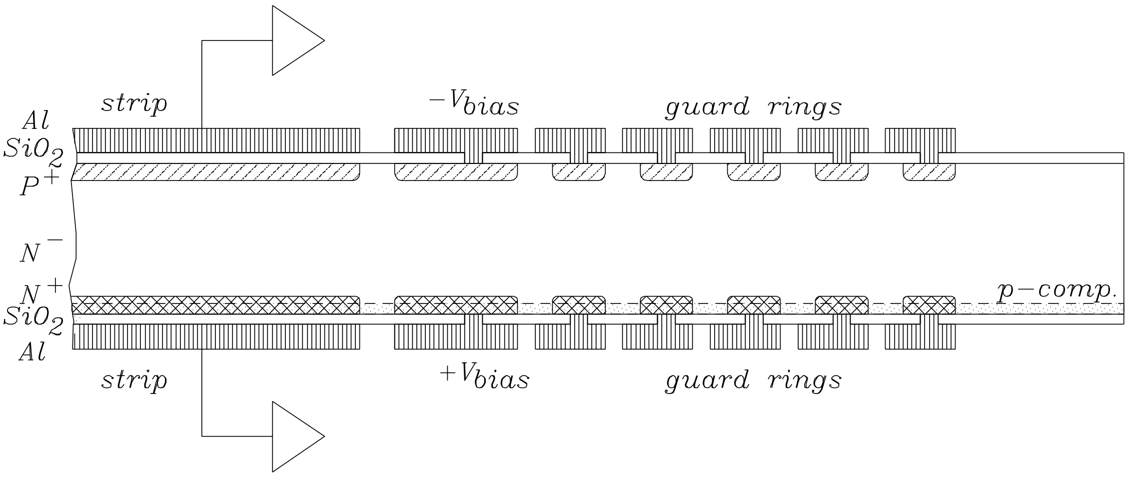
Using a p-type silicon substrate one may even simplify the concept further. High voltage breakdown protection structures are not needed on the p-strip side and the p-biasing structure can be contacted through the bulk through a forward biased diode on the rim of the n-side (figure 17). This way contact holes on the p-strip side are superfluous and step number 3 can be omitted. One thus arrives at rather:

- 5 photolithographic steps and
- 3 implantation steps

Single sided strip detectors with capacitive charge division and capacitively coupled readout
Single sided strip detectors can be derived from the previously presented double sided one by simplification of the structure. a variety of possibilities are available:

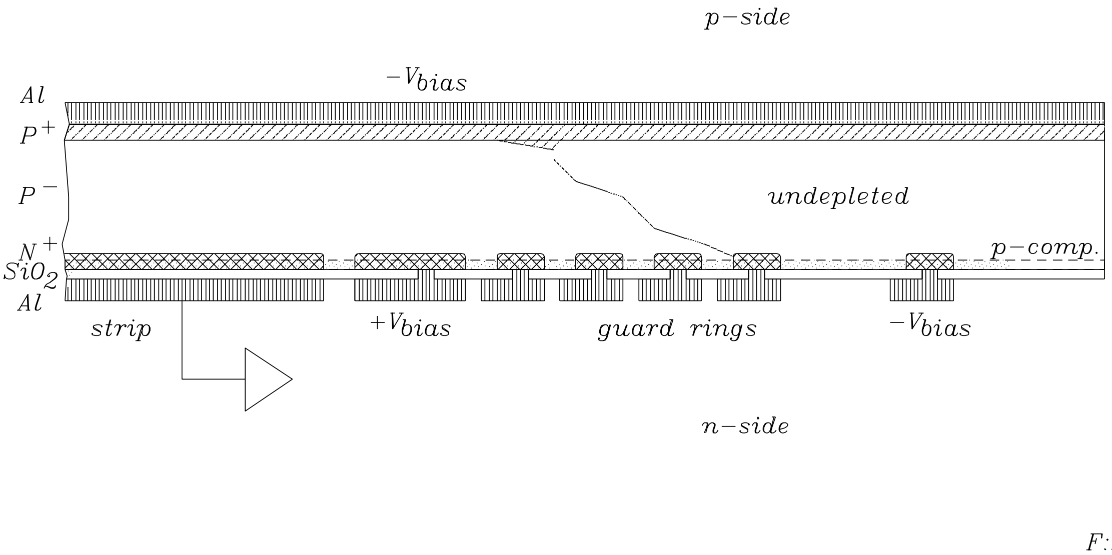
a) p-strips on n-type silicon
b) n-strips on p-type silicon
c) n-strips on n-type silicon
d) p-strips on p-type silicon

Neglecting for the moment the performance arguments, the technical implementation of these four options will be considered. Changing from double sided to single sided detectors is accomplished by replacing the strip pattern in figure 16 on one side by a large area implantation. This large area implantation in some cases may extend over the full wafer size, so that it can be done without photolithography.



F:1

Figure 16: Cross section of a double sided strip detector with capacitive charge division and capacitively coupled readout



F:2

Figure 17: Cross section of a single sided strip detector with capacitive charge division and capacitively coupled readout, built on p-type silicon

a) p-strips on n-type silicon

- Assuming that the bulk material transforms uniformly from n-type to p-type under irradiation, the junction would still move from the top to the bottom side in figure 16 so that the large area n-region should not be allowed to extend to the cutting edge and a guard ring structure on the bottom side would still be needed[16]. Producing such a detector would be almost as complicated as fabricating a true double sided detector if this further lithography was needed.

- Detectors produced in a much simpler fashion, having a simple unstructured n-layer on the bottom have been observed to survive type inversion. The reason for this survival is not completely clear yet and further investigations on mechanisms and reliability are needed. It may be that the surface of the detector does not invert at least up to fairly high irradiation doses, thus being responsible for the survival of the detectors after type inversion. Further investigation including e.g. the effect of microcracks developing near the cutting edge due to mechanical or thermal stress after irradiation are needed before considering this option as safe. If the surface does not invert, the high voltage drop remains on the top side outside of the breakdown protection structure. Adoption of this detector concept would bring considerable simplification as the n-side would only require one simple large area phosphorous implantation (or diffusion) and a large area aluminisation. One then has in total 3 photolithographic steps and 2 implantations.

b) n-strips on p-type silicon

- Here the top p-side can simply be replaced by large area p-implant (or diffusion) followed by large area metallization. One then has in total 3 photolithographic steps and 3 implantations.

c) n-strips on n-type silicon

- Here problems occur before type inversion if a single large area p-implant is used on the top side, as breakdown at the cutting edge might be expected. Photolithographic structuring of the p-side is therefore probably needed.

- Changing from double sided to single sided strip detectors would bring the simplification of running all electronics at the same potential.

d) p-strips on p-type silicon

- This option again does not lead to large simplification of the production process as photolithography is probably required on both sides of the wafer.

Considering the technical implications only the second option of version a) (p-strips on n-type wafer, relying on non-inversion of the surface) and option b) (n-strips on p-type wafer) seem attractive. While option b) is slightly more complicated – it requires an additional large area p-type implantation – it offers important advantages:

- It does not rely on the ill understood non-inversion of the surface region and therefore is not prone to possible reliability problems connected with it.

- Electrons are collected thus providing faster signal shape.

- Partial depletion operation should be possible.

Single sided strip detectors with directly coupled individual strip readout These detectors can be derived from the previous types by simply removing the insulator (oxide) between strip implant and metal. Taking the previous considerations into account we only have to consider two options:

a) p-strips on n-type silicon, relying on non-inversion of the surface

b) n-strips on p-type silicon

The design should however not be taken over straightforwardly as the following additional considerations apply:

- The strip to strip capacitance has to be kept very small so as to limit the capacitive load giving rise to the noise of the amplifier. (With capacitive charge division this load is reduced by having several strip-strip capacitances in series).
- In order to simplify technology with respect to capacitively coupled detectors contact holes and strip implantations should be done in a unique step. This brings about the danger of aluminium spiking into the silicon bulk at the edges of the implantation regions. Aluminium therefore is not supposed to cover the oxide edges. A barrier metallization or some other technique has to be used in order to prevent spiking.

Keeping the strip to strip capacitance low can be accomplished by narrowing the strip width and by increasing the gap between strips. The region between strips however does not necessarily behave like an insulator but may considerably increase strip to strip capacitance. This, in the case of p-strips, can be due to the formation of a conducting electron layer. In the case of n-strips the measures taken for strip-strip insulation may be responsible.

A further potential problem with narrowing the strip width is an increase of electric field strength respectively breakthrough. Capacitance and high field effects are related with each other as described in reference[17].

a) p-strips on n-type silicon, relying on non-inversion of the surface A cross section perpendicular to the strips is shown in figure 18. The potential on the outer surface of the oxide above the gap will be approximately equal to the strip potential. For a wide enough gap and low enough oxide charge no conducting layer will form in the gap region. As the oxide charge will be compensated by the electric field inside the oxide, the potential of the Si-SiO_2 interface will differ from the strip potential by slightly more than the flat band voltage. This voltage difference can reach appreciable values (especially after irradiation) and may lead to a significant increase of the needed detector bias voltage as well as to a danger of breakdown at the strip edges.

- Producing the detector requires 2 photolithographic steps (contact hole and strip opening and aluminium patterning) and 2 implantations (p-strips and large area n-type).

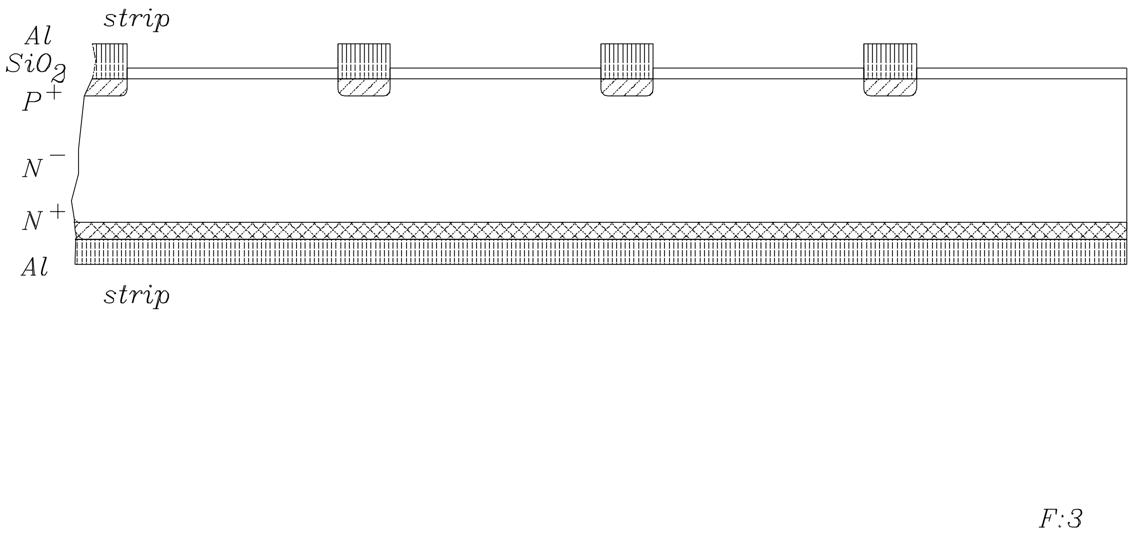


Figure 18: Cross section perpendicular to the strip direction of a single sided strip detector with individual strip dc coupled readout built on n-type silicon

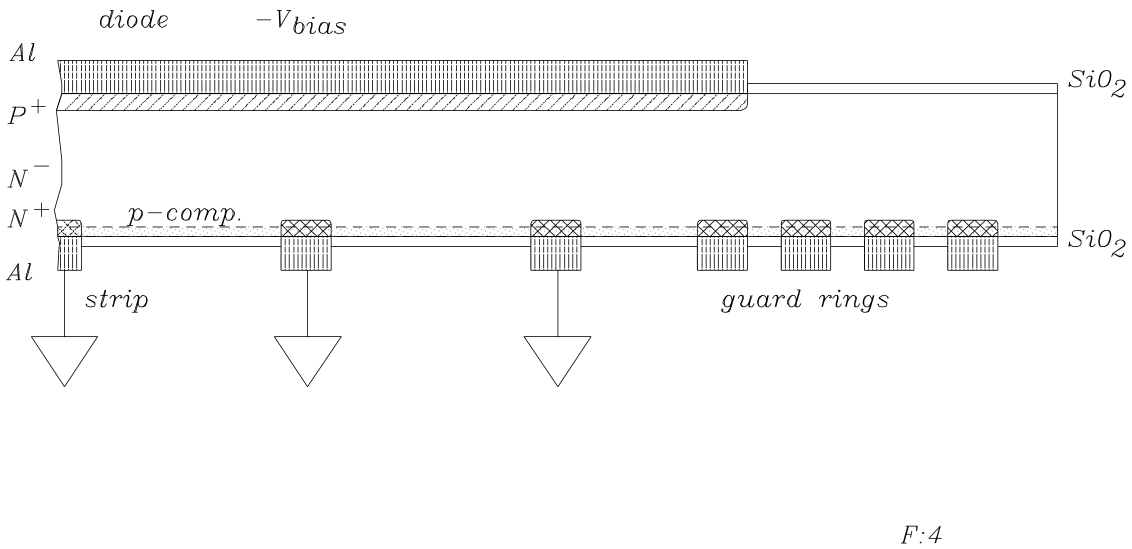


Figure 19: Cross section perpendicular to the strip direction of a single sided strip detector with individual strip dc coupled readout built on n-type silicon. n-strip insulation by large area compensation implant.

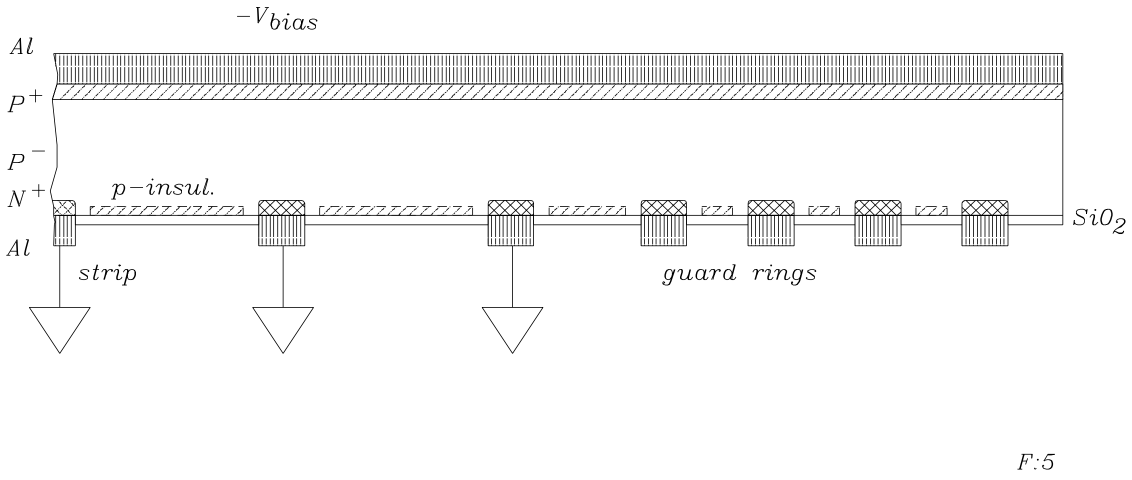


Figure 20: Cross section perpendicular to the strip direction of a single sided strip detector with individual strip dc coupled readout built on p-type silicon. n-strip insulation by wide p-strip implant.

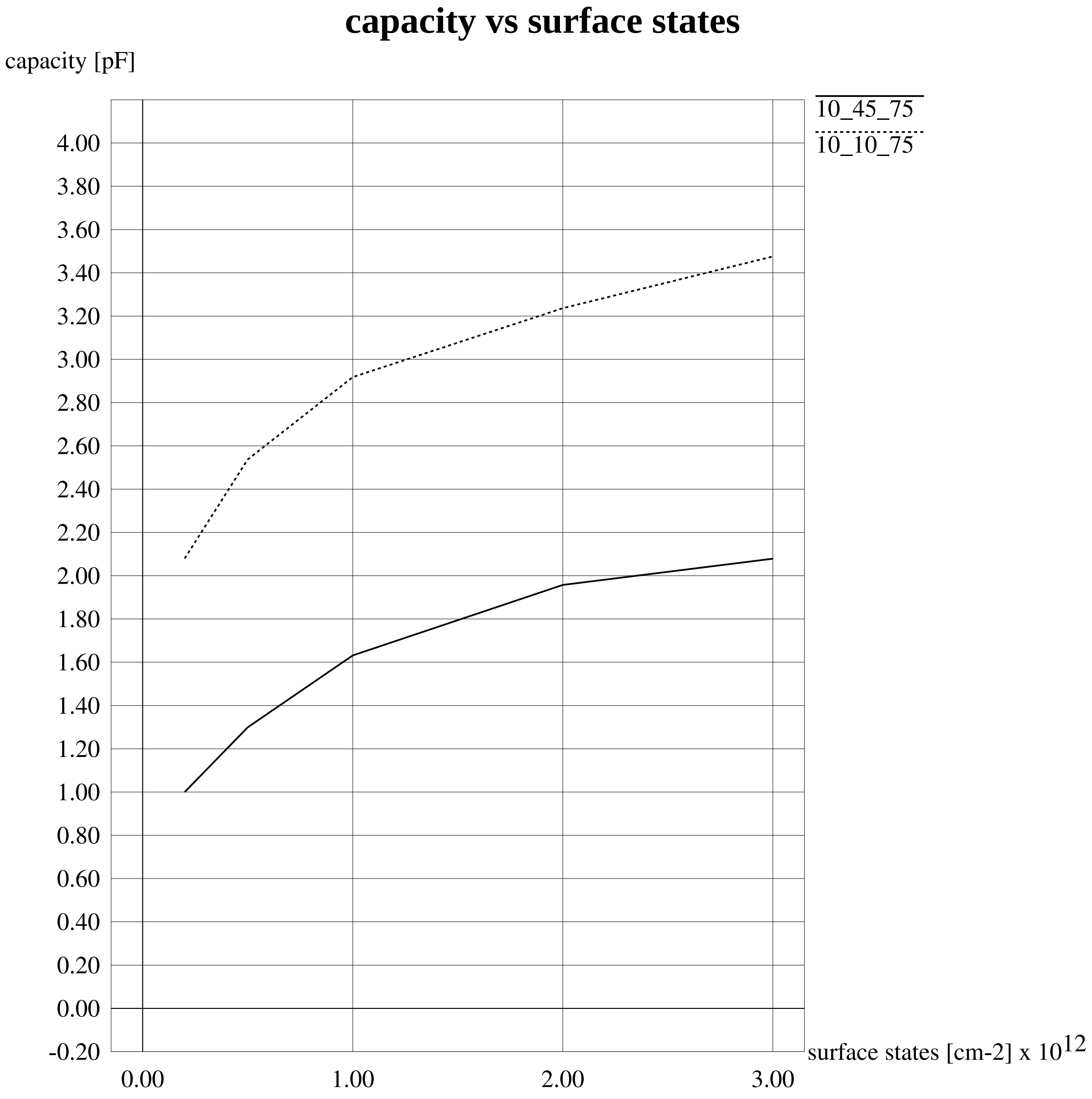


Figure 21: Total strip capacitance as function of oxide charge density for $10\mu m$ narrow n-type strips of $75\mu m$ pitch. Strip-strip insulation is accomplished with p-type insulation strips. Compared is the result for narrow ($10\mu m$) and wide ($45\mu m$) p-strips. One notices that for narrow insulation strips the capacitance is roughly a factor two higher

b) n-strips on p-type silicon n-strip insulation can be obtained either by a large area compensation as described above or by a p-strip implant totally surrounding the n-strips. The two options are shown in figures 19 and 20. Their specific properties are discussed below:

- A large area compensation implant (figure 19) has to be strong enough to tolerate the saturation oxide charge buildup due to irradiation. A wide conducting layer will therefore be present in between readout strips. Its conductivity and width will diminish with oxide charge buildup and therefore also the strip-strip capacitance will drop.
- p-type insulation strips have to be kept rather wide since the n-strips will effectively be widened by the surrounding electron inversion layer which will almost reach the insulation strips. This is shown in Figure 21 where the total strip capacitance as function of the oxide surface charge density is shown once for narrow p-type insulation strips, once for wide insulation strips. A factor of two variation in capacitance is observed. As in the previous case one therefore has a wide conducting layer in between the readout strips. Increasing oxide charge due to irradiation will strengthen the electron inversion layer and also lead to a higher strip-strip capacitance with irradiation.

A potential problem of the large gaps in between readout strips (in both versions) are long charge collection times in case that the hits are near the midpoint between two readout strips. The large capacitance between neighbouring insulation strips compared to the strip-strip capacitance can spread the signal over several strips during charge collection.

5 Electronics

This section discusses the design of the on–detector front–end electronics for the Inner Detector, and the off–detector readout electronics. At each level of the detector readout, several options are under consideration, and are being developed and compared with the ATLAS readout specifications.

5.1 Requirements of the ATLAS SCT electronics

Any chosen front–end and readout architecture must satisfy minimal performance requirements at both the component and system level. These requirements apply to the front–end preamplifier performance and subsequent binary or analogue signal pipeline, to the digitisation and sparsification of relevant data, and to the subsequent readout. Requirements relate to noise, speed, power consumption, and robustness.

A recent study [18] specified these minimal requirements in the case of the Atlas SCT; the report concentrated on the front–end preamplifier and analogue pipeline electronics. More recently, key requirements of the data digitisation and transmission have been defined.

Front–end performance. Important requirements listed in the study of ref. [18] are summarised in Table 2.

The total electronics noise after 10 years of operation should not seriously degrade the occupancy and consequent pattern recognition. A target single–channel noise of 1300–1500 electrons RMS was specified, following all signal processing, time slot to time slot pipeline variations, the serial noise following radiation–induced detector leakage currents, and parasitic noise or baseline variations. Because of the strip capacitances foreseen for the ATLAS SCT detector, and the short collection times imposed by the bunch crossing, this figure is difficult to achieve.

Time slot to time slot, or channel to channel variations in the pedestal and gain of the preamplifier, pipeline and read amplifier, have an important influence on subsequent sparsification and digitisation steps for all analogue (digitisation on–chip or off–chip) architectures. The comparator stability is also crucial for binary architectures. For this reason, an rms spread of $<50\%$ of the rms noise is required for the following effects:

- channel to channel variations (unless corrected at the sparsification comparator step),
- time slot to time slot variations,
- uncorrected baseline and other correlated variations of the pedestal or gain.

A key issue is the acceptable power consumption of the front–end electronics chip. After accounting for power losses, an upper limit of 40 mW cm^{-2} been estimated [19]. This translates to < 4.8 mW per channel for a 12 cm strip of 100 μm pitch.

To account for geometrical effects and Landau fluctuations, a maximum signal of 8–10 mips is specified. To reduce the pile–up to acceptable levels, a total pulse length of 15 bunch crossings is specified (3 bunch crossings for a binary system not using subsequent signal shaping). The recovery time from saturating pulses is required to not significantly deteriorate the effective occupancy. A value of 50 nsec has been specified for the separation of successive signals in a given channel; a shorter double–pulse resolution is possible, but this is already a constraint for other parts of the Inner Detector. For similar reasons 2 dead crossings following a Level–1 trigger are allowed.

Trigger rates. For stochastic Level–1 trigger rates of up to 100 kHz, and a 2 μsec pipeline, derandomising buffers are specified to maintain $< 0.1\%$ data loss.

Digitisation of analogue data. In analogue solutions, the above performance requirements on noise and on–chip power consumption must be met for both the sparsification and analog–to–digital conversion.

Analogue readout architectures include a sparsification circuit to suppress empty channels before readout, and in most readout scenarios this must be on–chip. The sparsification circuit is defined to include the discriminator function which selects those channels to digitise and provides a scheme of keeping track of the channel ID of digitised channels for later identification during data transmission. The sparsification may occur either before or after the digitisation step.

An ATLAS specification for a sparse readout circuit has been developed. It is discussed in Section 5.3.6. Several sparsification and digitisation architectures are being prototyped and evaluated with respect to the ATLAS specifications. Recently a baseline circuit concept was selected for the ATLAS SCT based upon a Wilkenson type ADC.

Naturally, the channel-to-channel and stability requirements of the preamplifier and pipeline steps will also affect the performance of the sparsification. Work in progress towards meeting these requirements is described in Section 5.3.

Data readout. For each of the readout options, a data transfer protocol has been defined, and a simplified read-out controller is under construction. This work is described in Section 5.6.

Control and Monitoring. The requirements for control and monitoring are described in Section 5.7. Preliminary specifications for a protocol and a possible implementation are presented. A time-schedule and development milestones are put forward.

Off–detector requirements For reasons of cost a maximum amount of functionality will be integrated onto the Front–end ASICS. In Section 5.8 the electronic system requirements and overall architecture are presented. Details are provided for the Clock and Control module, the Receiver modules, power supplies and distribution using minimal mass cables

5.2 Overview of front–end architectural options for the ATLAS SCT electronics.

In each of the proposed ATLAS front–end architecture options, detector signals are amplified, shaped, and stored in on-detector pipelines until a read–out request is prompted by a Level–1 (L1)trigger. At that point either binary, digital, or analogue data is transferred from the front–end chips to the off-detector data and trigger buffers.

The three architectures are listed below, and schematised in Figure 22

- **Binary:** A fast bipolar preamplifier-shaper-comparator is followed by a CMOS chip consisting of a single bit binary pipeline and a sparse readout that allows for a L1 trigger latency of at least 2 μsec.
- **Digital:** A preamplifier (either CMOS or bipolar) is followed by a 2 μsec analogue pipeline, a sparsification circuit, and an on-detector ADC. Digital information is transferred off-detector via optical fibre following a L1 trigger.
- **Analogue:** A preamplifier (either CMOS or bipolar) is followed by a 2 μsec analogue pipeline, and a fast analogue MUX transfers all data off–detector in analogue form via optical readout following an L1 trigger.

Several specific implementations of the preamplifier, analogue pipeline and signal processing design are being evaluated.

- a fast CMOS or bipolar preamplifier followed by a 2 μsec analogue pipeline on which the full analogue information (voltage or charge) is stored.
- a slow (50–75 nsec) CMOS preamplifier followed by an analogue pipeline on which sequential voltage samples are stored every 25nsec. Subsequent signal analysis using either 3 or 4 voltage samples (deconvolution) is used to identify the bunch crossing.
- a fast bipolar preamp-shaper-discriminator chip.

A number of readout schemes apply depending upon the various architectural options. In the case of binary readout, data from several chips are sequentially read into a read–out controller (ROC), and transferred by optical link to an off-detector receiver module. In the case of digital data, the pulse height information and addresses are transferred (either serially or in parallel) to a ROC controller for subsequent transmission over a digital optical link. In the case of analogue data use is made of optical modulation techniques to transfer data off the detector with little additional power consumption.

5.3 On–detector electronics for the ATLAS SCT detector.

As described in Section 5.2, several architectural options are being investigated for the ATLAS SCT on–detector electronics. No option has at this stage met the performance specifications imposed by ATLAS. In this section we describe the design and status of each functional unit of a possible ATLAS front–end chip (preamplifier, pipeline, digitisation and sparsification etc.) and describe future developments aimed at reaching the ATLAS performance specifications. In the case of a binary readout scheme, we describe the full readout chain.

5.3.1 The Binary readout architecture.

The binary scheme records only the presence of a hit for a given beam crossing. The position resolution is essentially given by the geometrical patterning precision of the silicon detector and does not rely on analog amplitude interpolation. Efficiency and occupancy are determined by the signal/noise ratio and the threshold matching.

The analogue front–end chain consists of a preamplifier, a shaper and a comparator. Comparator hits are stored in a time–sliced (or data driven) binary buffer. Upon receipt of a trigger signal, the hit pattern in the appropriate time slice is used to provide a sparsified readout. This represents the simplest form of a sparse readout system and allows substantial simplifications in the practical implementation, since

- analogue dynamic range requirements are reduced (linearity is required only up to the threshold), with substantial savings in power and chip area,
- readout bit rate is minimised (reducing the number and speed of readout lines),

As in all local sparsification schemes, threshold control is crucial. The reduced dynamic range requirement is beneficial , since it allows a sufficiently high gain up to the comparator input to mitigate systematic comparator threshold variations. Gain uniformity and baseline variations can be maintained at adequate levels (<5%), as has been demonstrated in several binary front–end systems operated in the laboratory, beam tests, and an ongoing experiment.

A full, monolithically integrated binary strip detector readout is operating successfully in ZEUS [20]. This design was a test bed for many of the concepts incorporated into the more demanding integrated circuits that were designed and tested for the SDC Silicon Tracker. In these implementations, the readout chain was split into two ICs: an analogue chip (preamplifier, shaper, comparator) utilising bipolar transistor technology, and a digital chip (data buffer with time stamp and readout) fabricated in CMOS. Both types of ICs have been fabricated with radiation-hard processes and tested. Suitable radiation-hard processes for both the bipolar transistor and CMOS chips are available from multiple foundries and performance has been demonstrated with various processes.

At this time chips have been designed and tested for ZEUS (Tektronix) [21], SDC (AT&T, Tektronix) [22][23][24] and NA50 (Tektronix) [25]. Both the Tektronix and AT&T processes allow circuitry on a 50 μm pitch and have been shown to be sufficiently radiation hard. Digital ICs using radiation-hard CMOS have been fabricated through Honeywell, Harris and United Technologies Microelectronics Center (UTMC). Clock-driven pipelines have been implemented using UTMC and Honeywell, and a data driven digital IC has been fabricated through Harris. The Honeywell process, by virtue of its small feature size of 0.8 μm and three metal layers, offers substantial advantages in power and die size.

Front–end ICs designed to SDC specifications have been designed at the University of California at Santa Cruz (UCSC) using the Tektronix process [24], and at Lawrence Berkeley Lab (LBL) using AT&T [22][23]. A chip for NA50 designed by Dabrowski [25] was fabricated through Tektronix. All chips use the basic block diagram shown in Figure 23. An integrating preamplifier is followed by gain stages and a comparator. The threshold is externally adjustable, so settings can be adjusted to the experimental environment, but also - and perhaps most importantly - to provide a diagnostic tool. Threshold scans perform a rudimentary analog-to-digital conversion that allows measurement of local noise levels, threshold settings, and signal levels.

All circuits use continuous, pseudo-gaussian pulse shaping without any switching activity in the analogue stages. The differentiation time constant is set by the RC discharge time constant in the preamplifier feedback path and the bandwidth of the gain stages is selected to provide the equivalent of multiple RC integrators. The circuit can also be tailored to tolerate substantial input currents with DC coupled detectors.

The signal gain up to the comparator input is set to yield a threshold level of >100 mv for 1 fC, to mitigate comparator threshold variations. If this gain were also effective at DC, the baseline drift would be excessive, so additional AC coupling with a long time constant is introduced. Figure 23 also shows the transition from single-ended to differential circuitry. All DC coupled stages are differential to maintain baseline stability and gain uniformity, both of which are critical to good threshold control. The externally controllable threshold is also applied differentially to eliminate the effect of voltage drops in the cabling. Detailed descriptions of the LBL design have been published [22],[23]. Table 3 summarises measured data for the ZEUS, SDC and NA50 chips. Preliminary simulation results for an improved IC currently under design as an ATLAS test prototype are also included.

The AT&T CBIC-U2 IC is substantially longer than the others. This has two reasons. The Tektronix process used in these designs does not include complementary *pnp* transistors and utilises oxide isolation. The AT&T process includes complementary *npn* and *pnp* transistors and uses junction isolation, which is more robust under irradiation. The AT&T CBIC-U2 chip was the first of the SDC designs and this process was chosen because it was best-characterised at high neutron and proton fluences, despite its lower density and marginal speed. Furthermore, since this circuit includes more redundancy to improve reliability and minimise radiation-induced circuit variations across a chip, the size of the die is increased.

An improved version of the SDC bipolar front–end is now being prepared for fabrication at AT&T [26]. This version has been designed specifically to meet the specification of operation in ATLAS at the LHC. This design utilises the CBIC-V2 process, which yields higher amplifier gain-bandwidth products and allows higher circuit density than the CBIC-U2 process used in the existing IC. The improved gain-bandwidth-product allows optimisation of the integration time constants in the individual gain stages to yield a very symmetrical pulse shape, so the double-pulse resolution is <50 ns (4 fC - 4 fC) for a 25 ns peaking time and realistic detector pulses. The newer process also allowed circuitry simplifications, which reduced the transistor count by about 30%. Together with the higher circuit density, this reduces the chip size to about half of the CBIC-U2 version, i.e. <4 mm. In addition this chip is designed to sink a detector current in excess of 2 μA per strip for use with the simpler and less expensive DC coupled single sided strip detector which are considered for use in the ATLAS SCT. Preliminary simulation results for this design are also included in Table 3.

Bipolar transistor circuitry inherently yields very good threshold matching. Since the gain in these designs is controlled by negative feedback, threshold and gain uniformity are quite good. Figures 24 and 25 show gain and threshold distributions measured on 8 chips of the ZEUS front–end [27]. Noise is also very well controlled. Similar results have also been reported for the SDC Tektronix chip [24]. Unlike CMOS, bipolar transistor noise level are very predictable, even after irradiation. AT&T preamplifiers irradiated to a fluence of about 10^{14}cm^{-2} (500 MeV protons) show an overall noise increase of about 7%. This noise increase is due to increased base current noise, because of degraded DC current gain. The base current at high fluence is typically several microamps, which is still larger than the detector bias current of $\sim$1 μA for a 12 cm long detector with 50 μm pitch, so that effects in the readout IC, rather than the detector, will dominate. A more detailed discussion with data follow in the section on radiation-hard electronics.

Two types of digital buffering have been implemented. The first is a simple clock-driven pipeline, where each bucket corresponds to a beam crossing. If a given channel is hit, it is registered in the appropriate time bucket. A trigger generates a pointer at the time bucket corresponding to the beam crossing, and the appropriate cells read out. The second is a data-driven circuit. Although the buffer area is smaller than for the clock-driven pipeline, the area required for the control logic is larger. Both schemes are competitive, but for trigger latency time in ATLAS, the high-density IC processes, such as the Honeywell 0.8 μm triple metal process, seem to tip the balance in favor of the simple clock-driven pipeline. Table 4 summarises the existing chips, two clock-driven pipelines (CDP64 and 128), the data-driven readout (DDR2), and projected data for an ATLAS-compatible clock-driven pipeline.

Simple data compression schemes have been implemented on the DDR2 to allow the use of inexpensive, low-speed fibre links. This allows the readout to be organised by module, to avoid the overhead of data collection chips to multiplex multiple module outputs to Gbit/s fibre links. The readout scheme also incorporates various failure mitigation techniques, for example circuitry to allow bypassing of defective ICs. More details are given in ref. [28]. This is also discussed in the readout section of this document.

A binary readout system has been in operation in ZEUS for about 6 months [20]. Although the crossing frequency is lower than at LHC, key characteristics, for example threshold control and diagnostic capability, are directly applicable to ATLAS. At a threshold level of 0.78 fC, the occupancy does not exceed 0.1%.

System operation is stable and predictable. SDC readouts using both the Tektronix and the AT&T ICs in conjunction with clock-driven pipelines operating at 40 MHz have been operated successfully in a beam test at KEK, yielding results that are consistent with bench measurements [29]. The new version of the AT&T IC is expected to be available in May, 1995 and at that time assembly of detector modules will commence for a beam test at CERN in late summer, 1995.

5.3.2 Fast–shaping CMOS preamplifier and pipeline architecture.

If achievable, the advantage of this option with respect to bipolar preamplifier designs is that the CMOS preamplifier can be integrated with the pipeline and other circuitry on a single chip. The advantage over slower CMOS preamplifiers is that the high speed allows minimal pile–up and correct BC identification, and the full pulse height information stored on a single pipeline capacitor.

This front–end design consists of a preamplifier having a fast rise slow fall time. The subsequent pipelining operation incorporates correlated double sampling (CDS) which results in a triangular 25 ns shaping time achieving the optimal noise performance for this technology, while at the same time maintaining robustness against correlated noise contributions. While a bipolar preamplifier may achieve better noise performance with the same risetime, two separate chips are required to achieve the same performance.

The preamplifier [30], called NICON, is a 3 GHz gain–bandwidth current mode preamplifier, with $<$15 nsec risetime, allowing full charge collection in a single bunch crossing. The circuit configuration, shown in Figure 26, has a grounded grounded gate input n–channel MOSFET transistor, operating in the weak inversion region. It is followed by a cascade with fast risetime and large gain (50mV for a $22000e^-$ input charge). The power consumption is only 300-500 μW, with a 5V supply voltage. The preamplifier input resistance is low (300Ω). The frequency stability is excellent because of the use of the grounded gate configuration without feedback. The operating DC conditions are controlled by a slow amplifier which forces the output DC voltage to be equal to a reference voltage, independent of the input current, supply voltage variations, and parameter drifts resulting from process variations or radiation damage. The preamplifier is therefore robust against environmental conditions. The availability of DC–coupling to the detector even at high leakage current is a potential system (economic) advantage. The NICON has been fabricated in 1.5 μm double–poly, double–metal CMOS, and separately fabricated using the Thomson SOI3HD process (see Section 5.4),

Figure 27 shows the measured pulse shape of the NICON preamplifier. A peak amplitude of 42 mV is obtained for 25000 e^-, with a risetime of 13 ns (10–90%). As shown in Figure 28, the integral non–linearity remains less than 2% in the range 0–15 mips input charge. After 25 nsec gaussian shaping, a noise level (ENC= 800+80 e^-/pF) is measured. This is shown in Figure 29. For an input capacitance of $<$ 10 pF, as originally intended for the ATLAS SCT, the noise is close to design requirements. For higher capacitance loads (the input capacitance is now expected to be 18 pF), the limiting parallel noise (800 e^-) becomes marginal.

Preamplifier and pipeline specifications for the SCT detector.	
Noise	<1300-1500e at 18pF input capacitance (single channel)
Maximum signal	8 mips (15 bunch spacings)
Pulse duration	<3 bc (binary) < 15 bc (analogue)
Risetime	< 25 nsec (fast–shaping option) < 25 nsec after deconvolution (slow shaping)
Bunch tag res.	1 bc
Double pulse resolution.	< 50 nsec (2 bunch crossings)
Max mean L-1 rate *L-1 trigger space*	100 kHz (stochastic) > 2 bc
Pipeline length	2μsec
Derandomising buffer	< 0.1% data loss.
Voltage compliance *Operational temperature*	0.3V nominal $\pm$ 2.5 $^\circ$ C in range -10 to $+30$ $^\circ$ C
Power	< 40 mW per cm^2 detector surface i.e. < 4.8 mW per channel total
Pad pitch	$\sim 44\mu$m

Table 2: Preamplifier and pipeline specifications.

Not yet available

Figure 22: Schematic of the On-Detector Readout Architectures (a)Analogue, (b)Binary, d(Digital)

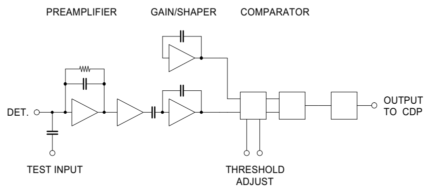


Figure 23: Block diagram of the analogue IC for binary readout

	ZEUS	SDC ATT	SDC Tek	NA50	ATLAS test
Noise	640e + 38e/pF	735e + 38e/pF	500e + 56e/pF	476e + 63e/pF	370e + 38e/pf
Power	2 mW/ch	1 mW/ch	0.8 mW/ch	1.3 mW/ch	1.8 mW/ch
Time walk	18 ns	16 ns	16 ns	10 ns	6 ns
Double-Pulse Resolution	170 ns	125 ns	80 ns	50 ns	50 ns
Readout Pitch	85 μm	50 μm	50 μm	50 μm	50 μm
Die size (64 ch)	5.4x4.9 mm^2	3.1x6.8 mm^2	3.1x4 mm^2	3.2x4 mm^2	6.2x4.5 mm^2 128 channels

Table 3: Analogue integrated circuits for binary readout. Time walk is stated for a signal range of 1.25 - 10 fC. Double-pulse resolution is for two successive signals of 4 fC (most probable energy loss). For the ATLAS test device, data is from simulation. Circuit is for use with n strips, p strip circuit has somewhat less noise. Power consumption is variable and noise can be changed. Numbers are for a single capacitor and no detector leakage.

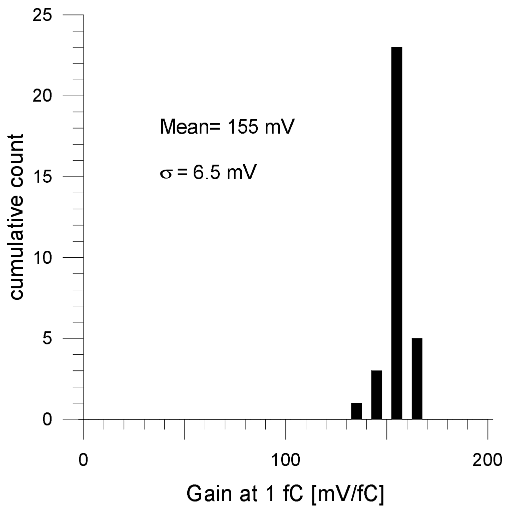


Figure 24: Gain uniformity of the ZEUS chip.

	CDP64	CDP128	DDR2	ATLAS
Clock Frequency	40 MHz	60 MHz	60 MHz	40 MHz
No. of cells/ch.	64	128	6	80
Process	1.2 μm double metal	0.8 μm triple metal	1.2 μm double metal	0.8 μm triple metal
Power [mW/ch]	0.66	0.25	1.24	0.15
Die Size [mm^2]	3.1 x 3.0	6.2 x 3.3	6.3 x 7.1	6.2 x 2.5 w. map+zoom

Table 4: Digital ICs for binary readout

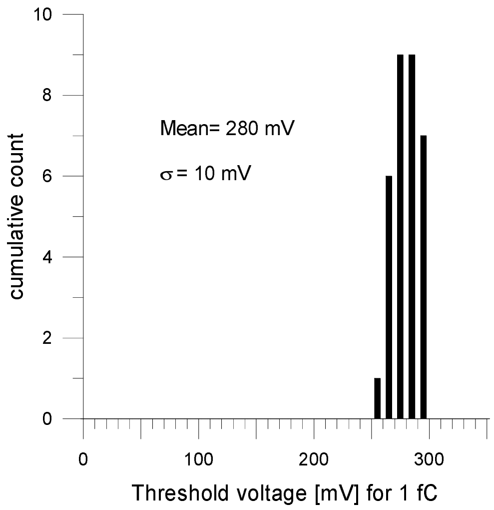


Figure 25: Threshold uniformity of the ZEUS chip.

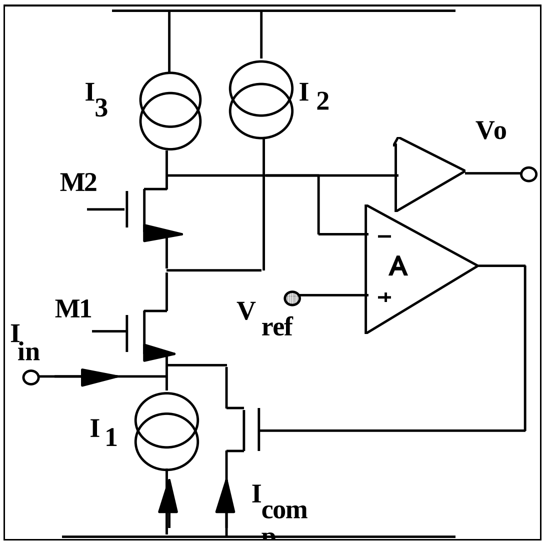


Figure 26: Schematic of the NICON preamplifier.

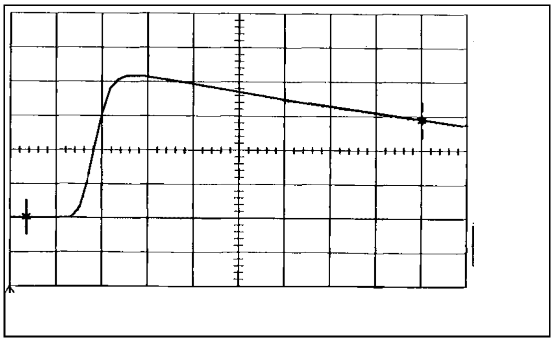


Figure 27: Voltage output of the NICON preamplifier (vertical scale = 10mV/div, horizontal 10nsec/div.

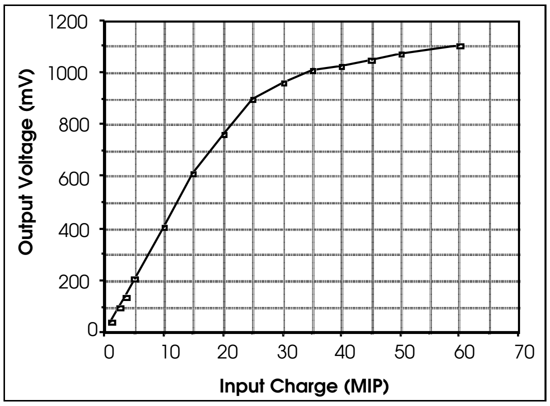


Figure 28: Linearity of the NICON (1 mip=4fC)

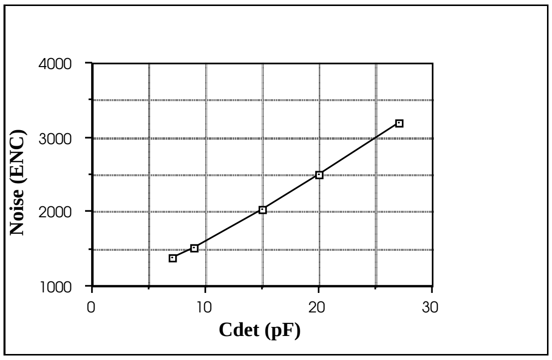


Figure 29: Noise of the NICON

Pipeline noise.			
Chip version	**Attributes**	**Fixed pattern noise**	**Baseline noise**
HARP	64 cell pipeline with limited digital on–chip control	1.4 mV rms	3.5 mV rms
APC3	128 cell pipeline with full digital control on chip. Static logic	5.6 mV rms [1]	3.5 mV rms
DYN–1	128 cell pipeline with full digital control on chip. Dynamic logic	1.9 mV rms	3.5 mV rms
DHARP2	64 cell pipeline with limited digital control on chip	< 2 mV rms [2]	< 1 mV rms [2]

Table 5: Summary of additional contributions to the output noise of channels in the pipeline function when operating under normal conditions (simultaneous read and write, at 40 MHz. ([1] non–gaussian [2] preliminary))

Several analogue pipeline studies adapted to the NICON have resulted in the DHARP2 pipeline, Figure 30, in which both the storage and shaping functions are simultaneously performed. In this scheme, all cells are periodically addressed; any cell may be protected however against sampling, and therefore skipped, following a trigger signal. At readout, the capacitors of successive time samples are subtracted, providing the required differentiation function (CDS) and common mode noise rejection.

Using a 32 channel DHARP2 chip with 64 time slots, and including the NICON front–end, the full signal can be collected in one time bucket, with a power consumption of < 2 mW. The sensitivity to signal jitter is measured to be $< 2\%$ of lost charge for a jitter of $\pm$ 2.5 ns. As shown in Figures 31 and 32, the requirement of a <50 ns double pulse reconstruction is easily achieved by the DHARP2 pipeline. The pulse shaping of the DHARP pipeline is also well adapted to bipolar preamplifier stages.

Analogue pipelines must be characterised for stability against perturbations affecting the samples during the three phases of operation: write, storage, and readout. Two of these perturbations are static effects which are inherent to the analogue structure: they are the fixed pattern noise which is the variation of pedestal from capacitor to capacitor within a given channel, and the channel to channel variation. Other perturbations are generated by dynamic effects, when the storage cells are addressed in either the read or write modes. The result is a variation of the the baseline, in a correlated way on all channels at the same time. Several sources of parasitic coupling can contribute to this common baseline effect [31]. Several realisations of the pipeline have been made since 1992; these are listed in Table 5, with their relative perturbative sources that contribute to the output noise in each channel:

Detailed tests in the laboratory, or with detectors in a source setup, laser assembly or test beam are made using a single DAQ system. In 1992–3, detailed testbeam measurements at a clock rate of 67 MHz, were made using 3 planes of DC–coupled, 200 μm–pitch detectors. In some runs, one of the detectors was irradiated to LHC levels. Figure 33 shows the residual distribution of tracks in each detector.

The DHARP pipeline is designed to be followed, on the same chip, by sparsification and digitisation circuitry, a readout FIFO, and readout control electronics. Figure 34 shows the schematic of the ADAM chip , a design which incorporates some of these integrated features.

The first 32 channel ADAM prototype has been fabricated and is being studied. Its architecture includes all the signal processing and control functions required for LHC except the signal sparsification. It uses the CRIAD piece–wise linear 6–bit ADC, described in Section 5.3.6.

When the ADAM receives a trigger signal an Address FIFO, which is part of the memory controller, stores the address of the corresponding locations in the analogue memory. Digitisation takes place if the address FIFO is not empty. During digitisation, the chip control copies the requested address to the output FIFO and points to the column location in the analogue memory. The analogue multiplexer then transfers data serially to the ADC, the analogue values corresponding to one or more samples of the signal. The output FIFO finally stores the digitised output, allowing readout without synchronisation to the internal chip operation. The set of 33 values is stored in the output FIFO during one digitisation cycle. An external

ADAM specifications.	
General	
Supply voltage	0-5V (±0.5 V)
Clock speed	40 MHz
Power consumption	¡ 2 mW per channel
Number channels	32
Rise time	13 ns (10-90%)
Coupling	AC or DC
Noise (NICON)	800+80/pF electrons rms
Detector dark current	10 μA
S:N	11 at Cdet=18 pf
Analogue Memory	
Dimensions	112 cells of 40 μmx 100 μm
Cap value	2 x 0.2 pF
Address FIFO	16 words by 7 bits
Analogue Mux	
Speed	5–10 MHz
Drive	10 pf max. Current drive ¡50 μA.
ADC	See CRIAD section 5.3.6
R/O buffer FIFO	66 words of 8 bit 10 MHz readout.

Table 6: Specifications and performance results for ADAM and components.

request signal generates the readout of one time frame on an 8–bit bus at 10 MHz. At the end of transmission the chip emits an output signal allowing the readout of a multi–chip daisy chain. Table 6 summarises the ADAM specifications.

Following a detailed study of this chip, a 128–channel version will be fabricated, compatible with both the NICON and bipolar preamplifier stages.

5.3.3 The Slow–shaping CMOS preamplifier and pipeline architecture.

In this approach, a very fast preamplifier (risetime $<$ 15 ns) circuit is combined with a slower (risetime 50–75 ns) shaping circuit in order to achieve both complete charge collection and optimal noise performance. The timing information required to associate an observed slow shaped signal to the 25 ns time slot of the corresponding Level–1 trigger is recovered using a deconvolution method [32, 33]. When leakage current from the detector is low, the parallel shot noise contribution is negligible and the full advantage of low noise from a slowly shaped sample, can be used to obtain optimal coordinate precision, while maintaining sufficient timing resolution to identify the crossing time bucket. In the case of high detector leakage currents, it has been demonstrated that deconvolution provides a very good filter (40% improvement over simple fast CR–RC reshaping), while the series noise is only worse by 6% compared with equivalent fast reshaping [33].

Implementation The deconvolution architecture is implemented in full CMOS technology. It essentially consists of 3 blocks, (a) a charge-sensitive preamplifier-shaper circuit, an analogue delay line buffer unit (ADB) and a simple deconvolution circuit (APSP). The preamplifier is laid out as a folded cascade with a pMOS input FET and a feedback, which allows operation of the circuit between ±2 VDC (Figure 35). The drain current of the input FET develops between 0 and 2 V resulting in a high transconductance at moderate power consumption. The voltage gain of the preamplifier is 1 mV/fC. The signal is then shaped in a CR-RC shaper amplifier with a further gain of a factor $\sim$ 20. At present two versions of this preamp–shaper circuit exist. One uses a 75 ns shaping time (FELix) and is implemented in a 1.2 μm rad-soft AMS (Austria Mikro Systeme) process, the other version (APV3) has been produced in a 1.2 μm rad-hard process (AVLSIRA by Harris) with a peaking time of 50 ns.

The second block of this system is an analogue pipeline (ADB) which samples the voltage output of the shaper every 25 ns. It consists of an array of capacitors and a digital control section. The analogue part is laid out as an array of capacitor storage cells, which receive the voltage samples directly from the low impedance output buffer of the shaper circuit (Figure 36). The samples are read, on request, by the APSP amplifier.

The necessary digital control signals are generated in the ADB control unit [34]. This unit controls the flow of data in the analogue section. It provides the 40 MHz clock to sample the input signal and controls the buffering of triggered samples and the readout of the samples into the APSP. The control logic is subdivided into a control unit and registers which contain pointers used to control writing, reading and skipping; the main registers are illustrated in Figure 37 for the case of the AMS version. In the simplest scenario, deconvolution makes use of 3 voltage samples. Thus for every validation from a Level–1 trigger, the 3 cells containing the samples needed for deconvolution, together with the cell containing the peak sample are marked as buffer cells. Up to 4 events can be buffered waiting for readout.

The readout of the reserved buffers is sequential, i.e. the first event in becomes the first event out. This is achieved through the first sample register. After readout the buffer zone will be released by removing the tags in the trigger-tag register.

Writing and reading are independent. The buffer zone cells are part of the pipeline. A skip mechanism is implemented which forces the write pointer to skip all cells which are in use as buffer zones until they are freed again after readout.

The deconvolution of the sampled signal is performed in a simple circuit (APSP) (Figure 38), consisting of an amplifier, which receives under the control of the readclock (1 MHz) the stored voltage signals from a given buffer zone in the pipeline. Five clock pulses are required for the full sampling operation. A set of switches place the three voltage samples sequentially on three capacitors. The relative sizes of theses capacitors provide the weights needed in the deconvolution sum. While the first 3 read clock pulses transfer the voltage samples for deconvolution, the 4th read clock pulse transfers the peak amplitude and on the 5th pulse the deconvolution sum appears on the output of the APSP amplifier.

The FELix circuit has 34 channels and and 84 cell deep pipeline. Of these cells, up to 16 can be used as buffers for trigger labelled samples. The capacitance of a single cell is 220 fF. The APV3 chip has 34 channels and 160 cells of depth of which up to 24 can be used as intermediate buffers. The capacitance of a single cell is 105 fF. Pointer and readout control are somewhat different in the two implementations. In both chips the upper two channels have output and input pads between shaper and ADB and between ADB and APSP to allow probing of the circuits at intermediate nodes.

Expected Noise Performance and Power Consumption Taking the ATLAS specifications on detector capacitance for 12 cm long modules of C_d = 18 pF and the noise requirement of Noise $\leq$1300 e^-, one can evaluate the required transconductance g_m of the CMOS preamplifier input FET. It is assumed here the optimal magnitude of the preamplifier input capacitance C_i and the detector capacitance C_d are related as $C_i = C_d/3$.

In this calculation the relationship before (slow) and after (fast) deconvolution in the series noise is ENC(Fast)= $1.84 * ENC(Slow)$ for deconvolution from 75 ns to 25 ns. The noise slope

$$ENC/C_{tot} = \frac{e}{q}\sqrt{\Gamma kT/3g_mT_p}.$$

Taking the total capacitance of the system to be $C_{tot} = C_d + C_i = 18 + 6 = 24$ pF, a noise slope of 50 e^-/pF at 25 nsec is required to reach the specification of ENC = 1300 e^-. In reality, the noise slope should be slightly less to compensate for the effect of parallel noise from a leakage current of about 3 μA per strip corresponding to the innermost layer after 10 years. Here the relation between parallel noise before (slow) and after (fast) deconvolution is ENC (Fast) = 0.35 ENC(Slow) [33] for the case of 75 nsec shaping time.

The corresponding transconductance is approximately 6.5 millisiemens for an ideal CMOS process. This can be achieved with a drain current of 700 μA. The corresponding noise slope of the slow shaping circuit is 27 e^-/pF. The corresponding power consumption in the preamplifier shaper circuit is 2 mW leading to a total power consumption per channel for the FELix chip of about 2.5 mW. To allow for a non ideal CMOS process, one has to plan for I_d = 800 μA giving 2.7 mW for the full channel.

To reach the full requirements of 1300 e^- after 10 years of running for a complete channel including the readout circuitry with sparsification, one will have to allow around 3 mW power consumption for the final chip.

First Test Results with Prototype Chips The different circuits used in the FELix chip, the AMP, ADB and APS, have been first produced as test chips and evaluated. All circuits were shown to work

Cell variations	
V_{in}(mV)	σ (mV)
0.0	0.61
10.0	0.63
20.0	0.49

Table 7: Measured storage cell voltage variations for 3 input dc voltage levels.

separately [34]. In particular, the ADB performance was studied to check for uniformity and linearity in response. The clock speed used was 67 MHz. A sine wave and a triangular wave form were injected into the pipeline spaced by only 2 BCO clock pulses. Figure 39 shows the ADB response and demonstrates both the linearity and the functioning of the skip logic. Table 7 shows the measured storage cell voltage variation for 3 different input DC voltage levels over all 84 cells. The rms variation from cell to cell are independent of DC input level and are around 0.6 mV. This indicates that the variations are not due to a spread in the capacitance of individual cells but probably are caused by variations of charge injection from switches at storage cells [32]. The corresponding voltage amplitude for a minimum ionising particle in the system is 40 mV.

The feasibility and functionality of the full chain was demonstrated in a test beam experiment by combining the three separate chips into one chain connected to a Si strip detector [35]. For this test, an amplifier with 45 ns peaking and deconvolution to a 15 nsec time interval was used.

More recently, first lab and beam tests with the complete 34-channel chips, the FELix and the APV3, have been performed. Full functionality has been demonstrated. In both chips, a number of features have been found which need improvement. Some of the results are presented here.

Results from first test with FELix: The noise as a function of load capacitance is shown in Figure 40 without deconvolution. The bias current of the preamplifier was set to 600 μA. For load capacitances above 5 pF, a slope of 34 e^-/pF is fitted. The zero capacitance point in this plot corresponds to the input capacitance of the preamp and some stray capacitance of the setup, totaling about 8 pF. The measured point at C=0 lies above the noise slope intercept indicating an additional constant noise contribution, most likely from the shaper, of about 300–350 e^-. In this measurement, which is made upon one channel, common external pickup noise cannot be eliminated.

In Figure 41 a measurement of the response of the preamplifier–shaper to a ^{241}Am 60 keV γ is shown. This is done using one of the probe-able channels. The upper trace shows a single pulse and the lower trace an average. The chip was connected to a strip detector (from CSEM) with 50 μm pitch and readout pitch, 6 cm long strips, and biased at 70 VDC. The observed pulseheight is compared with the baseline rms in this channel. This leads to a noise of ENC=640 e^-, compatible with the noise slope measurement in Figure 40.

Different assemblies of front–end chips with detectors were then tested in the H8 ATLAS test beam facility. Pulse height spectra before and after deconvolution were recorded. Signal to noise values (most probable) around 20:1 before and 10:1 after deconvolution were observed (Figure 42). Signal to noise values are always given for the signal of a cluster of strips above some threshold and for the noise of one strip in the cluster.

As expected, the S/N value after deconvolution depends upon the correct correlation between trigger time and start of the BCO clock pulse. In the asynchronous test beam situation this is achieved by measuring the relative time between the BCO pulse and a trigger hodoscope signal with a TDC. Selecting events within an interval of $\pm$1 ns around the BCO clock pulse achieves optimal S/N values. The timing precision from the scintillators is in this case the limiting factor.

Results from tests with APV3 First laboratory measurements on 5 APV3 chips gave a noise performance of ENC = 550e^- + 50e^-/pF. The shaper noise contributes $\sim$300e^- of the 550e^- base value. All 5 samples show virtually identical results. The measurements show that the APV3 has only a 10% higher noise value than observed with the FELix chip, after scaling the observed values with the relative peaking time and drain current conditions in the two measurements.

This result is better than expected from transistor characterisations on samples from a previous production run at Harris. Further measurements on transistors from the latest run confirm that the analogue performance in this run was considerably improved. For transistors with channel lengths of 2.4 μm and larger, almost ideal noise factors of $\gamma = 0.7$–1.0 have been measured ($\gamma = 2/3$ ideally in the strong inversion regime). The full functionality of the chip was proven in a beam test. There are remaining problems: the read amplifier design of the pipeline, and cross-talk between the read and write lines in the pipeline. Relevant design improvements have been made in the APV5 design presently under production at Harris.

Results on the radiation tolerance of this chip are expected by early 1995.

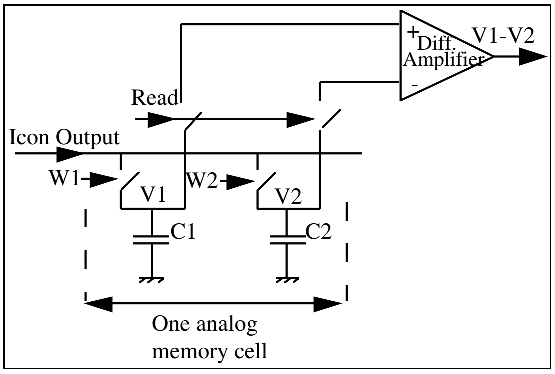


Figure 30: Principle of DHARP2 pipeline.

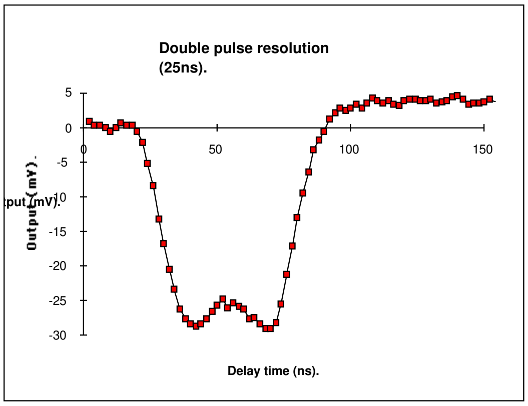


Figure 31: DHARP2; separation of 2 signals 25 ns apart

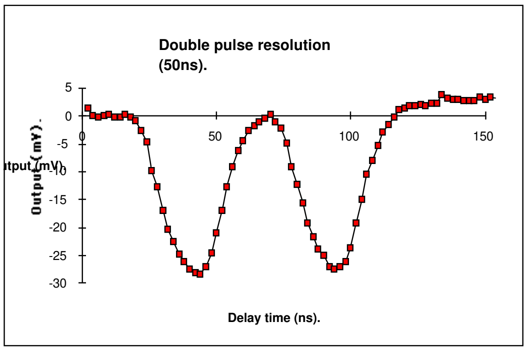


Figure 32: DHARP2; separation of 2 signals 50 ns apart

Not yet available

Figure 33: Residuals from test–beam using DHARP2

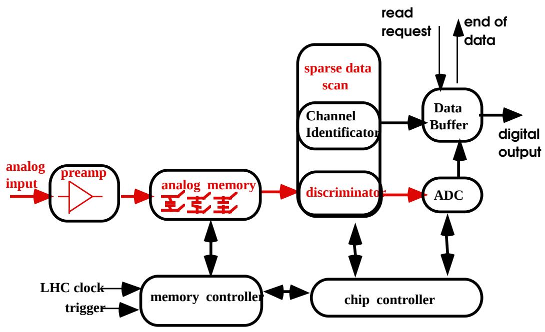


Figure 34: Schematic of the ADAM chip.

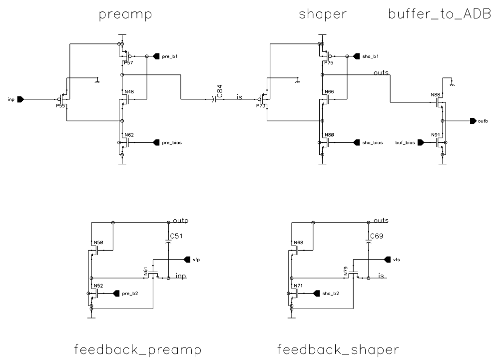


Figure 35: Layout of the preamplifier–shaper, analogue delay line and deconvolution circuit for the slow–shaping option

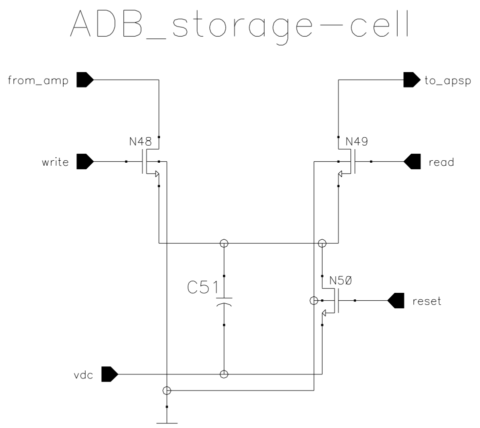


Figure 36: Cell architecture of the ADB pipeline

Not yet available

Figure 37: Architecture of the ADB pipeline element

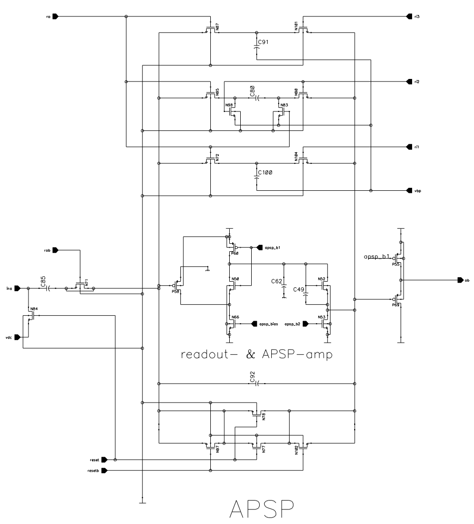


Figure 38: The APSP circuit architecture

Not yet available

Figure 39: ADB Response

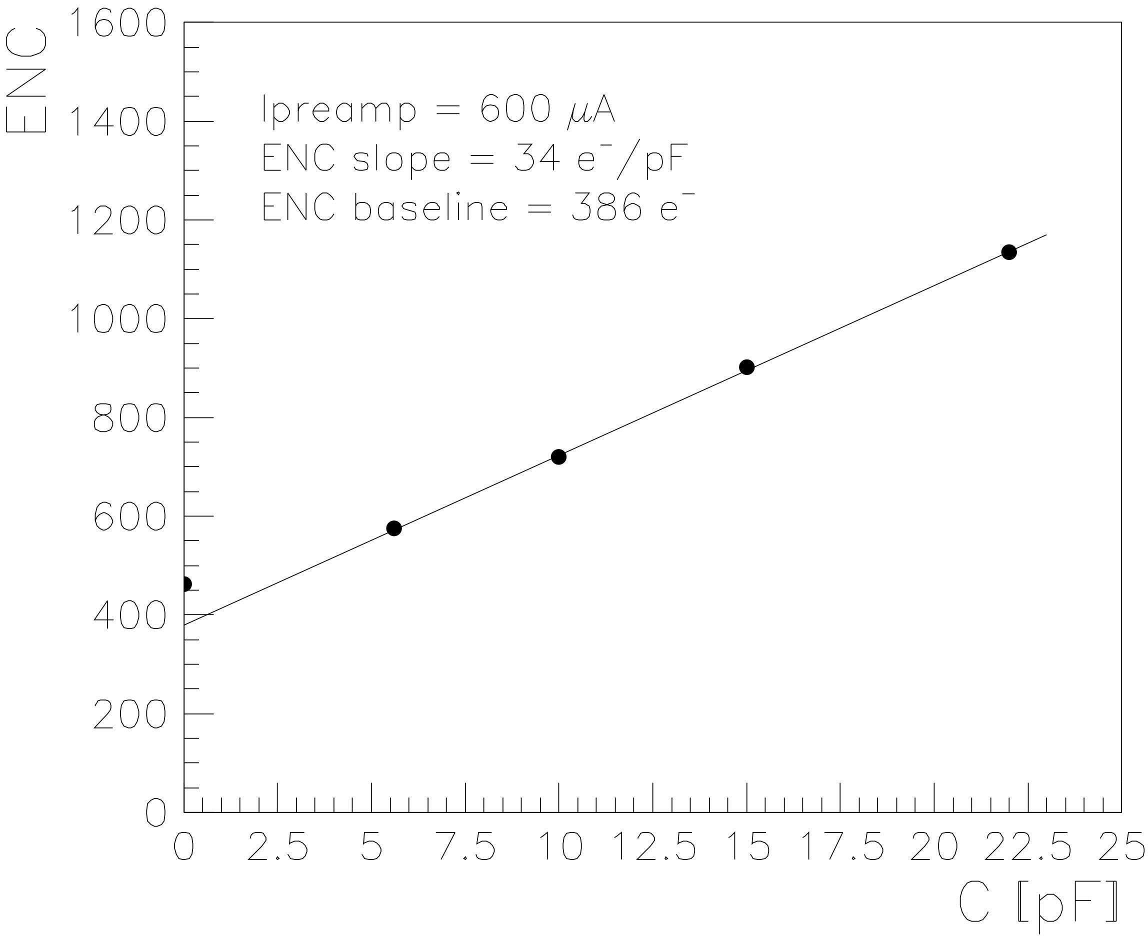


Figure 40: Noise performance of FELix

Not yet available

Figure 41: Response of the FELix preamplifier–shaper to a ^{241}Am 60 keV γ.

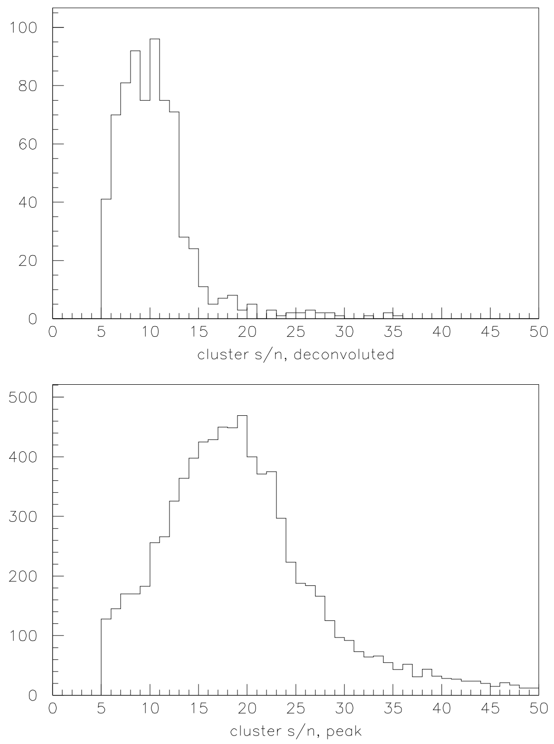


Figure 42: Pulseheight spectra obtained using FELix readout in a high momentum testbeam

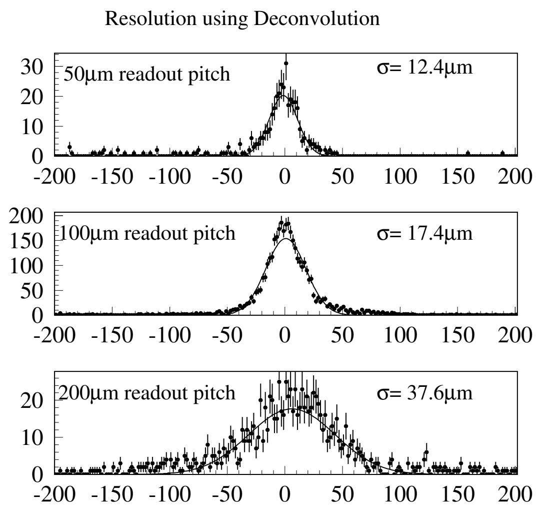


Figure 43: Spatial resolution measured in the H8 testbeam using detectors equipped with FELix and APV3 with deconvolution

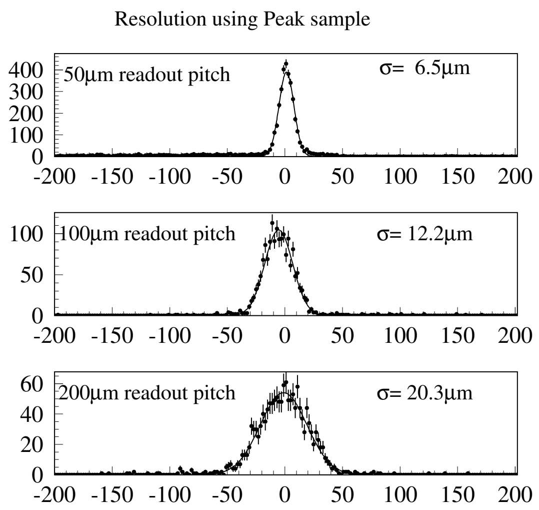


Figure 44: Spatial resolution measured in the H8 testbeam using detectors equipped with FELix and APV3 using peak values

Spatial resolution The spatial resolution was measured with both FELix and APV3 chips for detectors with 50 μm, 100 μm and 200 μm readout pitch. The 50 and 100 μm devices had one intermediate strip while the 200 μm device had two intermediate strips. Measured residual distributions for the different detectors are shown in Figures 43 and 44. The 50 μm pitch and 100 μm pitch detectors were connected to FELix chips and the 200 μm pitch detector to an APV3. The resolutions are σ = 12.4 μm, 17.4 μm and 37.6 μm when using the deconvoluted amplitude and σ = 6.5 μm, 12.2 μm and 20.3 μm when using the peak value of the amplitude. The resolution of each detector can be directly taken from the observed residual distributions, since the position of the extrapolated track at the detector under test is measured to a precision of better than 2 μm in both x and y coordinates using the H8 Si beam telescope.

5.3.4 A Bipolar preamplifier–shaper and CMOS analogue pipeline architecture.

The performance of different fast bipolar preamplifier-shaper-comparator circuits with peaking times of around 20-25 nsec, which have recently been designed for SDC and NA50 and are at present under design for ATLAS have been described above. In particular, the measured noise performance of the TEKTRONIX amplifier for NA50 and the simulated noise performance of the ATT design for ATLAS look promising. A programme to build such a bipolar front–end without discriminator and combine it with an analogue pipeline has therefore been initiated.

In a first step, a version of the new ATT design for ATLAS has been produced in which the comparator is omitted and a high impedance current output at the output of the shaper has been implemented. The useful linear range of the preamplifier has been extended for analogue applications, and the shaping has been made more symmetrical to meet double pulse resolution requirements. The preamplifier is designed to remain stable with detector leakage current of several microamps.

The chip is laid out in 128-channel form with 128 parallel output pads. The expected noise slope obtained from a simulation is shown in Fig. 10. With a power of 1.6 mW/channel for the preamplifier-shaper alone, one expects with a 12 cm long detector a noise performance of 1400 e^- including shot noise from 2μA leakage current.

Signals from this chip will then be fed to the inputs of specially adapted pipelines of existing design.

- The analogue pipeline of the FELix chip has been adapted by adding a current to voltage converter at the input (Figure 10); the output will be simply the peak sample of the fast shaped pulse from the bipolar input circuit.
- The shaped bipolar output is well suited with minimal modification to the DHARP pipeline developed for the fast-CMOS application (NICON).

5.3.5 Milestones towards an ATLAS analogue front–end.

The ADAM development: The ADAM chip architecture for digitised readout consists of 4 architectural entities; the fast front–end preamplifier and shaper, the analogue pipeline and pipeline control, the digitisation and sparsification step, and the readout buffer/protocol. It is foreseen to develop each section as part of a fully integrated 128–channel chip, during 1995, using as basis the developments already described above. As a result, the strategy of component improvement and detailed chip submissions are described separately below. As already noted the 32 channel ADAM chip is under test, and will be used for the readout of RD2 LHC detector modules early in 1995. It is expected to equip the second generation ATLAS Z-modules (see Section 6 for a description of module configurations) with 128 channel rad–soft chips before the end of 1995, and to submit designs for rad–hard fabrication in this period.

The implementation of each entity is envisaged as follows:

1. The front–end preamplifier and shaper: following tests combining the DHARP pipeline with a bipolar front–end preamplifier in March 1995, a choice will be made on the basis of performance between this option and a lower–noise development of the NICON preamplifier, in the final 128 channel chip.
2. The key to analogue and digital readout architectures is to minimise the total (correlated and uncorrelated) noise in the fully integrated chip and under realistic test conditions. The key weaknesses so far have been the sensitivity to external pickup, and to cross-talk from the digital control. The DHARP double correlated sampling architecture with digital control will be further optimised to minimise these effects.

3. The ADAM chip already includes a CRIAD ADC operating within design specifications. Following the decision, to pursue a Wilkinson–type ADC and sparsification architecture, it is hoped to equip the 128–channel ADAM chip with that ADC; pending the development of this, however, the CRIAD non–sparsified readout will maintained as backup.

4. A buffered output, and a prototype readout protocol has been developed by the RD2 Collaboration for multi–chip readout of a detector module, on the basis of the 32 channel ADAM with digitised but non-sparsified readout. It is foreseen to extend and adapt this to fulfill the protocol specifications discussed in Section 5.6.

On the basis of these developments, the following chip fabrications are under way, or planned:

1. A DHARP 32 channel pipeline has been adapted for bipolar front–ends, and is under fabrication; the resulting tests of this chip with the bipolar front–end will help to determine the baseline ATLAS front–end choice (see below).

2. A 128 channel MIETEC 1.5 μm layout of the ADAM chip is starting and can be expected to be submitted in May 1995; the final front–end choice will be chosen following the tests above, and the choice of Wilkinson or CRIAD in the 1st chip submission will depend on the status of the former. An analogue O/P will also be available. Considerable emphasis will be placed on development of the pipeline control section.

3. It is expected that submissions of design optimisations for the Thomson HSOI3HD rad–hard process, and possibly the Harris CMOS process, will be made later in 1995.

The FELix development: A new version of the FELix–32 chip was submitted in November 1994. Major modifications were implemented in the control section mainly to improve robustness with respect to timing variations. Recognised weaknesses of the analogue circuitry of the ADB-APSP section have been eliminated in the new design. The biggest improvement is expected from a redesign of the APSP amplifier which is laid out similarly to the circuit used in the preamplifier section. Tests with this chip are expected in the period February–April 1995. In parallel the circuit will be laid out in a 128–channel version to be ready for submission in March 1995. This will allow incorporation of possible shortcomings of the FELix–32 design. The 128 channel version will be laid out with parallel output pads to have the greatest flexibility in joining it to either a 128 channel multiplexer chip or to a sparsification chip. Detailed laboratory tests with the 128 channel chip are expected to start in Summer 1995, and beam tests with a set of different ATLAS prototype detectors are foreseen in the last H8 test period. The schedule relies on having no major failure in the foreseen processing runs.

The APV development: The existing APV3 32 channel version is currently under extensive test. These tests include first irradiations with neutrons and electrons. Some design weaknesses similar to those found in the FELix–32 have been corrected in the APV5 128 channel chip now under fabrication. This chip, which has a 128 channel analogue multiplexer integrated in the full layout will be available for testing early in 1995. The chip will be both laboratory and beam tested during 1995. Since further improvements may be required, an updated version can be submitted, following temporary closure of the Harris plant, late in 1995

On the basis of rad–soft technologies, an assessment of the status will be possible in 1995. At the same time, the first full chip implementations in rad–hard SOI, and DMILL technologies are expected; that will enable a comparison with the Harris process used for APV5 fabrication.

Bipolar developments: As noted above, an ATT version of the 128 channel bipolar preamplifier–shaper front–end will be available for tests about March 1995. At the same time optimisations of the ADB and DHARP pipelines will be available to test the full bipolar–CMOS analogue and digital architectures. Given satisfactory results a development of the ADAM 128 channel chip which is matched to the bipolar preamplifier–shaper is expected to be submitted in May–June 1995.

5.3.6 On–detector sparsification and digitisation.

The specification for a sparse readout circuit is given here:

- For digitised output architectures, the circuit is part of the pipeline chip or adjacent to it (i.e. one sparse circuit per pipeline chip with no multiplexing into the sparsification circuit).
- The selection of channels to digitise and transmit off- detector is made on the basis of a signal threshold. The possibility of discrimination with respect to 3–channel analogue sums exists, but requires a feasibility and performance study.
- Neighbor logic is required as an option set at initialisation. This implies that the channels immediately adjacent to a channel exceeding threshold will also be digitised.
- An option to read out all channels is required.
- It must be possible to mask channels before digitisation and data transmission.
- The sparsification threshold is required to be programmable chip–by–chip, either via an analogue differential level or an on–chip DAC.
- The resolution of the threshold setting must be <25% of the rms. noise for the total front–end chain.
- The rms. variation of sparse threshold across a chip (discriminator to discriminator) is required to be <50% of the total rms. noise.
- The power consumption of the sparsification circuit should be minimal, and <50 mW/chip or 500 μW/channel (see above).
- For digital or analogue architectures, an effective 7–bit resolution is acceptable for charge division studies, but this must be extended (nonlinearity or bits) if noise measurements are requested.
- The total time for sparsification and digitisation time must be $< 3.3\mu$sec , averaged over the expected occupancy.

ADC implementations have already been made [36]which satisfy (or close to satisfy) the requirements of ATLAS. Other implementations are also being considered, the aim being to optimise the performance of the full ADC and sparsification module for speed, power and stability. As a baseline concept ATLAS has chosen to emphasise the development of a Wilkenson type ADC with sparsification. Work is still planned on other implementations, as a back–up to this. Below we discuss the various ADC–sparsifications circuits which have been considered.

The Wilkenson ADC–sparsification circuit is based upon a concept developed for the CDF SVX–II program. In this circuit, all channels are digitised in parallel by comparing a global chip wide ramp to the each channel's output. Upon a crossing, a counter value is latched for that channel. Readout is performed only on those channels which exceed a programmable digital threshold.

The CRIAD ADC, in its present implementation, is a piecewise linear ADC with four segmented dynamic ranges, each with a 6-bit precision. The lowest range (0-64mV) provides maximal resolution (1mV), while the highest range (512 mV-2048mV) has a 32 mV resolution. In any final implementation, these ranges can be optimised. The architecture, shown in Figure 45 consists of one track and hold (T&H), and a 2-step A-to-D converter. The dynamic range is then determined by a resistor string. The second step is performed by a 6-bit linear charge redistribution ADC in the range indicated by the flash. To maintain the required throughput, seven time-interleaved redistribution units operate in pipeline mode. The comparison clock of the ADC is linked to that of the main 40 MHz clock; 6 clock cycles are required for 6-bit digitisation, but the result is a full 8-bit digitised output for each clock cycle (2 bits indicate the range and possibly in future the sparsification level, with 6-bit output. The CRIAD has been fabricated and tested. Results are shown in Table 8. Figure 46 shows a 6KHz signal of 800mV amplitude reconstructed from the ADC digital output, running at 5M sample/sec with automatic segment switching. The CRIAD has also been implemented on the ADAM chip, and detailed tests are in progress.

CRIAD results	
Max resolution	11 bits
Sampling rate	5 MHz
Differential non-linearity	$\pm$ 0.5 LSB
Integral non-linearity	$\pm$ 0.5 LSB
Input range	0 to -2048 mV
Power supply	-2.5 V to 2.5 V
Power consumption	25 mW
Active area	5 mm^2
Technology	1.5 μm CMOS

Table 8: CRIAD experimental results.

Not yet available

Figure 45: Architecture of the piece-wise linear CRIAD ADC.

Not yet available

Figure 46: Output code for 800mV sine wave. Horizontal scale is time, vertical scale is binary o/p converted to decimal code.

5.4 Implementation of radiation hard electronics in the ATLAS SCT.

Radiation hard ICs are crucial for reliable operation of local electronics in the ATLAS tracker. The SCT requires mixed mode integrated circuits operating after a total ionising dose of 10 Mrads and fluence in the order of 10^{14} hadrons cm^{-2}. These front–end IC's must employ radiation hard processes which are carefully selected and qualified by ATLAS in order to guarantee reliability for an operating life of 10 years. No experience in space or defence applications with operation at these radiation levels [37]–[38] exists. Consequently extensive testing is required by the HEP community. Not all architectures can be implemented in the same way on different technologies. This leads to a difficulty of simple second sourcing of components. Put differently, some architectures are optimised for specific technologies so it is difficult to select technologies and architectures independently. The only useful assessment is to compare real implemented prototypes of the proposed system. Therefore, a preliminary prototyping phase with a few silicon foundries, from which a first and a second source have to be selected, would be certainly necessary for the full qualification of rad hard front–end chips. This will enable us to guarantee their reliability with the combination of irradiation and aging over the expected operating life of LHC. There already exist several bipolar and several field effect technologies which will withstand the required radiation levels with acceptable performance degradation. The technologies useful for the tracker are probably all known to us already. The job to be done is to gain more experience with these processes and using prototypes prove that the proposed system architectures can be implemented reliably.

5.4.1 Bipolar transistor processes

Bipolar transistors are inherently radiation hard. The basic transistor is unaffected by ionisation effects. At the fluences of interest in ATLAS, displacement damage has no practical effect on the transconductance nor output resistance. The parameter that is affected is DC current gain, which degrades and increases the shot noise contribution of the base current.

The mechanism causing current gain degradation is the same as in detectors. Defects created by displacement damage increase the generation current in the base, which adds a constant term to the base current and reduces the DC current gain. The magnitude of the generation scales with base volume, so transistors with smaller base widths, i.e. higher speed, are more radiation resistant. Operation at higher current densities also reduces the relative degradation. Transistors with high pre-radiation current gains will exhibit a larger relative degradation than those with lower initial current gains. However, the absolute magnitude of the post-radiation current gain will tend to be quite similar in both types of devices as the generation current becomes significant, i.e. of order 1% of the collector current.

Although the basic transistor is radiation resistant, the manner in which the device is incorporated in an integrated circuit may cause severe radiation effects. Lateral transistors can be implemented rather simply (and are found in many CMOS processes), but they have large base volumes and are very sensitive to surface effects. Isolation techniques also affect radiation resistance. Oxide isolation increases circuit density over junction isolation, but can make the circuit very sensitive to ionisation. Lateral transistors and structures with oxide isolation often exhibit severe degradation after ^{60}Co irradiation, so it is always prudent to compare the results of gamma and neutron or proton irradiations.

In general, fully complementary processes with vertical *npn* and *pnp* transistors and junction isolation offer superior radiation resistance. Furthermore, faster processes operated at high current densities are less susceptible to current gain degradation, so industry trends toward higher speeds and smaller devices will also increase the radiation resistance.

Full custom bipolar transistor ICs are available from several companies as foundry services. Bipolar ICs for HEP applications have been fabricated primarily through AT&T and Maxim (formerly Tektronix), for which processes and results are described below. Harris, IBM, and Westinghouse also offer suitable processes.

Bipolar Transistor Processes–AT&T: AT&T Microelectronics offers fully complementary bipolar processes with junction isolation and two metal interconnect layers. The CBIC-U2 with 4.5 GHz vertical *npn* and 3.75 GHz vertical *pnp* transistors has been in production for several years. This process has been utilised for several HEP designs, notably the front–end ICs for the SDC Silicon Tracker and the SDC straw tracker [22][23][39]. The faster and higher-density CBIC-V2 process was released in 1993 and provides 10 GHz npn and 4.5 GHz pnp transistors, both rated at 10 V. The minimum transistor size is 35.0 x 37.5 μm^2 for both *npn* and *pnp* devices. The minimum emitter area is 1.5 x 15 μm^2. A typical input transistor for a silicon strip readout would use a 3× device with an emitter area of about 70 μm^2. Two types of diffused resistors are available, one with a sheet resistance of 80 Ω/sq. and the other with 1800 Ω/sq. A fusible link trim capability is also provided. Capacitors can be implemented either as metal-nitride-oxide-silicon (MNOS) or metal-on-metal (MOM) structures. MOM capacitors exhibit the best characteristics with a capacitance density of $3.75{\cdot}10^{-5}$ pF/μm^2.and an absolute tolerance of $\pm$20%. Higher capacitance can be realised with the MNOS structures, which provide a capacitance density of $2.3 \cdot 10^{-4}$ pF/μm^2 ($\pm$35 tolerance) and allow both rectangular and comb configurations.

Experimental Results: CBIC-V2 devices have been irradiated to a fluence of $1.2{\cdot}10^{14}$ cm^{-2} (800 MeV protons). The DC current gain of *npn* and *pnp* transistors before and after irradiation is shown in figs. 47 and 48.

These data are for an emitter area of 45 μm^2. At a current of 100 μA, corresponding to the current density typical of a strip detector front–end, the DC current gain after irradiation decreases to about 60% of its original value for the *npn* and to 50% for the *pnp* transistors. Transconductance and output resistance are hardly affected. The noise measured on individual devices matches the calculated value for the transconductance and current gain. An increase in base resistance could not be detected. Of more interest is the measured noise of a complete charge-sensitive amplifier as implemented in a full analogue IC for binary readout. Fig. 49 shows the measured spectral noise density at the output of the preamplifier before and after irradiation to a fluence of about 10^{14} cm^2 (500 MeV protons). The noise increases by about 7%, without taking a slight increase in gain into account.

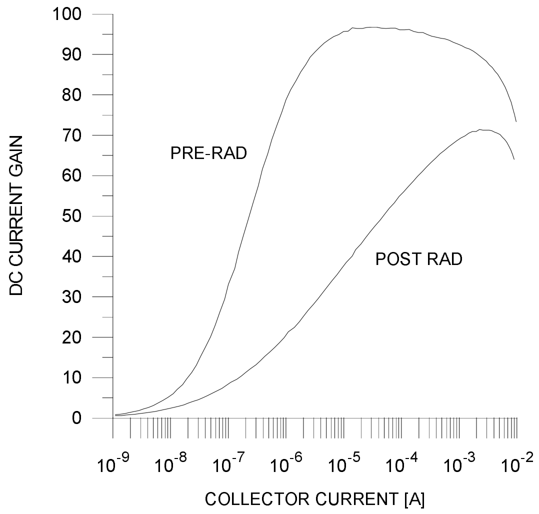


Figure 47: DC current gain vs. collector current of npn transistors before and after irradiation to a fluence of $1.2{\cdot}10^{14}$ cm^{-2} (800 MeV protons).

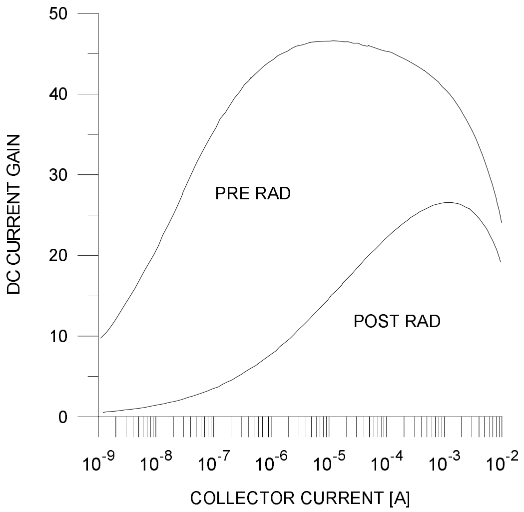


Figure 48: DC current gain vs. collector current of pnp transistors before and after irradiation to a fluence of $1.2{\cdot}10^{14}$ cm^{-2} (800 MeV protons).

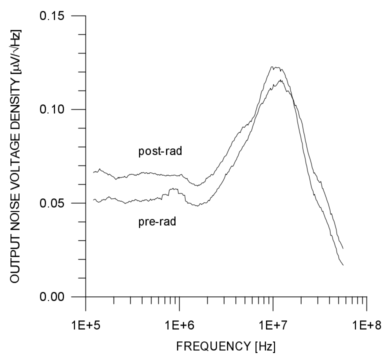


Figure 49: spectral noise density at the output of the preamplifier before and after irradiation to a fluence of about 10^{14} cm^2 (500 MeV protons).

Bipolar Transistor Processes–Maxim (formerly Tektronix): Maxim has recently taken over the Tektronix bipolar IC foundry operation. Several HEP designs have utilised the Tektronix SHPi process [27][21][25][24]. This process is well-characterised for analogue applications and provides excellent device matching. Active devices include vertical *npn* transistors with 9 GHz unity gain bandwidth, peaking at a collector current of 500 μA for a 1$\times$ device (emitter area 0.75 x 4 μm^2). A minimum size *npn* transistor occupies about 12 x 24 μm^2. Oxide isolation provides high circuit density, with some attendant radiation effects. Only lateral *pnp* transistors are available, with rather poor radiation characteristics [40][24], but *p*-JFETs available on this process offer a usable alternative for some purposes [25][24]. Resistors can be implemented as active base, *p*+ diffusion or nichrome structures. MOM and MOS capacitors are also available. Interconnections can be made in two metal layers with 2 μm layout rules. HEP circuits designed with this process include straw-tube [39] and silicon strip detector front–end ICs [21][25][24]. With proper design, this process provides excellent results and yields, with low-noise analogue circuitry that can operate to fluences of 10^{14} cm^{-2} [41][24]. Recently Maxim has announced the CHPi process, which adds vertical *pnp* transistors to the SHPi process.

Bipolar Transistor Processes–DMILL: DMILL provides CMOS, JFET and bipolar transistor circuitry in a process specifically designed for mixed analog-digital applications. This process is being investigated by several HEP groups, one example is for the pixel readout in ATLAS. A description of the DMILL process with radiation data are given in ref. [42].

NPN bipolar transistors from DMILL technology are designed with a vertical structure with a very thin base in order to obtain a high frequency transition (f_t > 5 GHz) together with a high radiation hardness against displacement damages in silicon caused by hadrons. The design of these devices is also optimised for high radiation hardness against ionising effect. Several values have been considered for the current gain: from 100–200. The final version of DMILL bipolar transistors will be designed for a current gain range of 150–180.

Neutron hardness: Bipolar devices have been irradiated and characterised after neutron irradiation using a 1 MeV source (nuclear reactor) with no bias during irradiation. For NPN transistors designed with a 1.2 x 1.2 μm^2 emitter surface and exhibiting an initial current gain of β=100, the relative variation of the gain $\Delta\beta/\beta$ (Ic = 100 μA) after 10^{14}n cm^{-2} is about -35%, corresponding to a final current gain of 70, see Figure 50.

Gamma hardness: Bipolar devices have irradiated and characterised under gamma irradiation, using a ^{60}Co source at a dose rate of 3 Mrad/hour, with collector bias = 2 V and base bias = 0.75 V (referred to the emitter) during irradiation. For NPN bipolar transistors designed with a 1.2 x 1.2 μm^2 emitter surface and exhibiting an initial β of about 210, the relative variation of the gain $\Delta\beta/\beta$ (Ic = 100 uA) after 17 Mrad is about -20%, corresponding to a final current gain of about 170, see Figure 51.

Bipolar Transistor Processes–Other Processes: As mentioned above, suitable processes are also available from Harris, IBM, and Westinghouse. Some radiation data for Harris are given in ref. [43]. Westinghouse devices show characteristics similar to AT&T CBIC-V2 [44].

Summary of bipolar demonstration circuits:

5.4.2 Radiation hard CMOS bulk processes.

There are two types of radiation hard silicon field effect processes bulk and silicon on insulator. Several bulk processes have been investigated in the HEP community specifically: Harris AVLSI-RA, UTMC UTE-R and Honeywell RICMOS. These have been evaluated to various levels and prototype ICs fabricated. All these processes are available as foundry products and allow designers in the community to access them with the appropriate non disclosure agreements in a similar fashion to non radiation hard technologies.

Radiation hard CMOS bulk processes–Harris AVLSI–RA: The Harris AVLSI-RA process is a 1.2 μm bulk (epitaxial) CMOS process with two metal layers and a high voltage independent precision capacitor. The process is very simple to design with needing no special considerations to achieve radiation

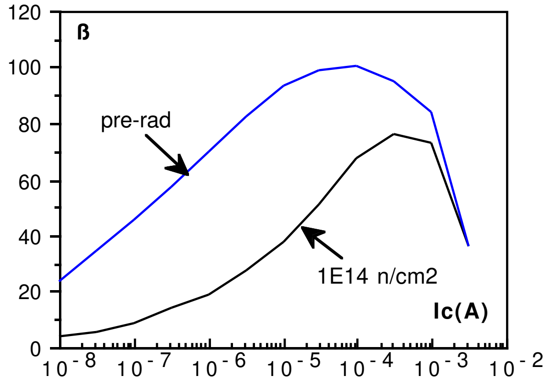


Figure 50: Variation of current gain with neutron irradiation for NPN transistors

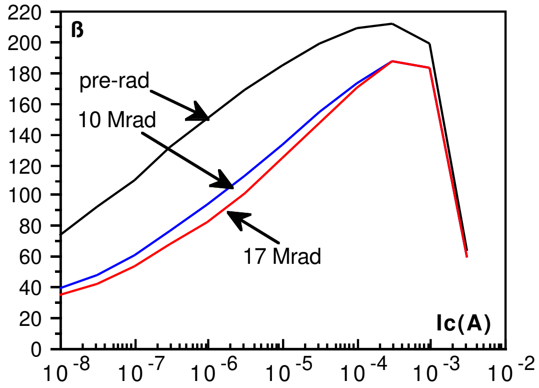


Figure 51: Variation of current gain with gamma irradiation for NPN transistors

hardness. Harris specify a radiation hardness of 1 Mrad and 10^{14} n cm^{-2}. Although guaranteed only to these radiation levels, the process seems to be much harder. Test transistors and a complete mixed signal integrated circuit have been irradiated and shown to function well at total doses greater than 100 Mrad. Full prototype chips appropriate to Atlas applications have been built and function [37],[45].

Gamma irradiation of Harris AVLSI–RA transistors: Transistors fabricated on the Harris AVLSI-RA technology were irradiated using a ^{60}Co source. The transistors were irradiated biased with 5 volts and unbiased to 10Mrad (100kGy) at 55 rad/s (0.55Gy/s). The devices are grounded after irradiation and kept at room temperature. After 7 months the devices were heated to 120 °C for 14 hours and remeasured and then again stored at room temperature. Only a few devices were measured from one processing batch.

The biased 80:9 micron NMOS transistors after 10 Mrad (100kGy) show a very small reduction in gain and also a −400 mV shift in threshold which anneals to around −300 mV after 10 months. The sub threshold leakage current of these devices at 0 volts gate bias shows a three order of magnitude increase but they are still switched off sufficiently for operation in digital and most analogue applications. The unbiased 80:9 micron NMOS transistors after 10 Mrad (100kGy) show a −200 mV shift in threshold but a small reduction in gain. This then produced a shift in the characteristic of just more than −100 mV at 200μA drain current. This then anneals slightly after 10 months to less than −100 mV. The biased 5600:1.5 micron PMOS transistors after 10 Mrad (100kGy) show noreduction in gain but a −200 mV shift in threshold. There is no change in the characteristics after 10 months. The unbiased 5600:1.5 micron PMOS transistors after 10 Mrad (100kGy) show a negligible reduction in gain. The measurement for the 1 hour after irradiation was done with the drain at 5 volts. All other measurements were done with −2 volts. Correcting for this then gives approximately −100 mV shift in threshold and no change in the characteristics after 10 months. This is a very positive result and gives every indication that circuits built on this technology will operate successfully even after a 10 Mrad irradiation.

Further Gamma Irradiation of Harris AVLSI–RA transistors: A different batch of devices were irradiated using a ^{60}Co cell at AEA Technology. The transistors were irradiated biased with 5 volts and unbiased. The doses used were 69 Mrad (0.69 MGy) at 217 rad/s, 100 Mrad (1.0 MGy) and 125 Mrad (1.25 MGy) at 390 rad/s. The ambient temperature during irradiation was 20 °C. Dosimetry was carried out with Red 4034 Perspex dosimeters and is accurate to better than 10temperature. The first measurements were carried out during a three hour period after irradiation. The biased 80:9 micron NMOS transistors after 69 Mrad and 125 Mrad (0.69 MGy and 1.25 MGy) show a very small reduction in gain and also a −400 mV shift in threshold. The sub threshold leakage current of these devices at 0 volts gate bias shows a two order of magnitude increase (Figure 52) but they are still switched off sufficiently for operation in digital and most analogue applications. Similar devices measured previously showed a −400 mV threshold-shift after 10 Mrad.The unbiased 80:9 micron NMOS transistors after 69 Mrad and 125 Mrad (0.69 MGy and 1.25 MGy) show a negligible shift in threshold but a small reduction in gain.

This then produced a shift in the characteristic of about −300 mV at 1mA drain current. The sub–threshold leakage current of these devices at 0 volts gate bias show little change. Similar devices measured previously showed similar small threshold shifts after a 10 Mrad irradiation. The biased 5600:1.5 micron PMOS transistors after 69 Mrad, 100 Mrad and 125 Mrad (0.69 MGy, 1.0 MGy and 1.25 MGy) show a negligible reduction in gain but a −200 mV shift in threshold (Figure 53). Similar devices measured previously showed the same threshold shift after a 10 Mrad irradiation.

The unbiased 5600:1.5 micron PMOS transistors after 69 Mrad and 125 Mrad (0.69 MGy and 1.25 MGy) show a negligible reduction in gain but a −100 mV shift in threshold. The sub threshold leakage current of the 125 Mrad device at low gate bias shows three orders of magnitude increase which is not evident at 69 Mrad. It is not clear whether the difference is due to this particular device or the increased dose. This should not affect the device in most applications except possibly when used in long term switched capacitor storage switches. Similar devices measured previously showed the same threshold shift after 10 Mrad. This is a very positive result and gives every indication that circuits built on this technology will operate successfully even after 100 Mrad (1MGy).

Neutron irradiation of Harris AVLSI–RA transistors: Transistors fabricated on the Harris AVLSI-RA technology and which had previously been irradiated to 10Mrad (100 kGy) of gamma radiation were irradiated using the ISIS neutron source at RAL. The transistors were irradiated unbiased to 10^{14}n cm^{-2}

Preamplifier noise and gain.				
Chip Number	Noise mV rms Prerad	Noise mV rms 100 Mrad	Gain Prerad	Gain 100 Mrad
1	0.17	0.19	0.059	0.055
2	0.17	0.22	0.0525	0.0513
3	0.16	0.26	0.054	0.050
4	0.18	0.25	0.055	0.0525
5	0.14	0.17	0.050	0.0463

Table 9: Radiation effects on noise and gain of Harris MX7-RH

with 1 Mev neutrons. The devices were grounded during and after irradiation and kept at room temperature. None of the samples showed any appreciable change in their I/V characteristics. The noise of the unbiased 5600/1.5 micron PMOS device showed a 2 dB reduction in flicker noise. This may be due to the annealing of the damage caused by the gamma irradiation. This is however a very good result showing a small effect on the flicker noise mechanism. The noise above the flicker noise corner frequency due to thermal noise is unaffected by either irradiation.

Gamma irradiation of Harris MX7-RH: As well as the measurements on individual transistors it is important to confirm that a large system fabricated on the process will work after irradiation. RAL has built an MX7-RH chip for High Energy Physics applications which contains 128 low noise amplifiers, filters and an analogue output multiplexer. This 36 mm^2 mixed analogue and digital chip is used to accurately measure charge deposited by ionising radiation in silicon detectors. It has been extensively studied and instruments many thousand electronics channels. Five MX7-RH chips fabricated on the Harris AVLSI-RA technology were irradiated using a ^{60}Co cell at AEA Technology. The transistors were irradiated biased with 5 volts to 100Mrad (1.0MGy) at 390 rad/s. The devices were grounded after irradiation and kept at room temperature. The first measurements were carried out during a twelve hour period after irradiation. Within the first hour after irradiation two chips [38],[46] were working but had a factor two increase in noise. The other chips had dead channels but the shift registers worked. Chip 1 had one third of the channels dead. After 12 hours all channels worked on all chips. The noise and gain were slightly affected as shown in Table 9.

This is a very positive result and gives every indication that complete mixed analogue and digital systems can be built on this technology will operate successfully even after 100 Mrad (1MGy).

Summary of Harris AVLSI-RA process: In summary, test transistors and MX7-RH integrated circuits, fabricated in the Harris AVLSI-RA process, have been radiation tested in a ^{60}Co environment to total doses of up to 100 Mrad (1.0 MGy) and, after short annealing, shown to exhibit relatively little degradation. This is believed to be the first time that results on this process at this level of total dose have been reported. The test data give very promising indications that devices can be built in this technology for in-cell remote operations and future hadron colliders at total doses of the order of 100 Mrad (1 MGy) and 10^{14} n cm-2. Furthermore, this is a well supported commercial process.

Radiation hard CMOS bulk processes–UTMC: United Technologies Microelectronics Center (UTMC) offers a radiation hardened bulk CMOS process that has been utilised by high-energy physics groups for several years [20][47][48][49]. This is a p-well process with a 1.2 μm minimum feature size. The process uses polysilicon gates with an oxide thickness of 20 nm. The sheet resistance of the polysilicon gates is 20 Ω/sq. and by careful design the parasitic noise resistance can be kept quite small [47]. The minimum effective channel length is 0.8 μm for the NMOS and 1.0 μm for the PMOS devices. Until now this process has been available with two layers of metallisation, but recently UTMC has announced the introduction of a third metal layer. The maximum die size is 1.49 x 1.49 cm^2. Using triple metal, UTMC's gate array technology supports up to 100000 gates. This is a "digital" process, but it has been demonstrated to be quite usable for analogue applications that do not require precision capacitors or resistors.

Type	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
Width	75	75	1332	1332	888	888	1332	1332
Length	1.2	1.2	1.2	1.2	2.2	2.2	3.2	3.2
I_d/W= 0.03								
0 Mrad			0.81	0.61	0.64	0.59	0.66	0.50
5 Mrad			2.17	0.84	1.00	0.58	1.50	0.69
I_d/W= 0.10								
0 Mrad	1.10	0.70	1.20	1.10	0.80	0.80	0.80	0.60
5 Mrad	3.80	1.10	3.40	1.60	1.30	0.90	1.70	0.70
I_d/W= 0.30								
0 Mrad	1.60	1.30	2.00	1.70	1.10	1.00	1.10	0.77
5 Mrad	5.00	2.90	4.80	2.70	1.60	1.40	1.20	0.81

Table 10: Noise coefficients $\gamma_n = v_n^2 \cdot g_m/4kT$ for NMOS and PMOS transistors of various widths and lengths, operated at current densities I_d/W= 0.03, 0.10 and 0.3 A/m, before and after irradiation to 5 Mrad(Si).

Experimental Results: Threshold shifts of 200 mV for NMOS and 150 mV for PMOS have been measured after ^{60}Co irradiations to 5 Mrad [50][51][52]. About 2/3 of these shifts are observed after the first 250 krad. Figs. 54 and 55 shows the scaled transconductance g_m/I_d vs. I_d/W before and after irradiation [52]. For the selected channel length this representation allows direct scaling to any device width at constant current density. The PMOS devices are quite insensitive to radiation, with only a few percent decrease in transconductance at 5 Mrad, while the NMOS devices suffered a 30% degradation. About half of the observed change at 5 Mrad occurred after a dose of 1 Mrad.

Extensive noise measurements have been performed at the University of Pennsylvania [47] and by a UCSC/LBL group [52]). In the latter, spectral noise density was measured over a frequency range of 10 kHz to 10 MHz before and after ^{60}Co irradiation to a dose of 5 Mrad(Si). Shaping times for typical ATLAS tracking front–ends correspond to peaking frequencies in the range 2 to 10 MHz. The noise was measured at three representative drain current densities I_d/W, i.e. the drain current per unit width. Again, these data can be scaled to any device width, where the noise scales with $1/\sqrt{W}$. The difference between the NMOS and PMOS results is striking. The NMOS devices show a much greater degradation and the PMOS devices also exhibit substantially less low-frequency noise. The low- frequency noise spectral density of the NMOS devices can be described by

$$v_n^2 = \frac{A_f}{f^q} + B$$

where q ranges from 0.8 to 1.0 and is constant for all currents for the same geometry. The changes in q after a dose of 5 Mrad are of order 0.1. The noise coefficient A_f is about 1.0 to $1.5 \cdot 10^{-30}$ V^2 pre-rad and 5 to $10 \cdot 10^{-30}$ V^2 post-rad. Before irradiation, A_f scales well with inverse gate area, but no clear pattern is observed after irradiation. The low frequency noise behavior of the PMOS devices is more complex and cannot be parametrised in this simple manner, but the devices exhibit substantially better noise than the NMOS transistors.

White noise was evaluated at high frequencies and is characterised by the noise coefficient

$$\gamma_n = \frac{v_n^2 \cdot g_m}{4kT}$$

to assess the inherent noise properties independent of transconductance. Results for various device geometries and current densities are shown in Table 10. For these measurements the substrate was biased at the source potential.

Again, we see substantially less post-rad degradation for the PMOS devices. One can also observe the higher intrinsic noise in NMOS short channel devices. Although the observed degradation is quite small in some cases, typically it is quite substantial and would need to be compensated for by a considerably higher operating current.

Several HEP ICs have been fabricated using the UTMC process. Clock-driven pipelines for ZEUS and SDC have been fabricated and tested, and are in use in a running experiment [20]. The SVX IC designed

Not yet available

Figure 52: Change of sub–threshold leakage current at 0V gate bias for NMOS transistors (Harris AVLSI-RA) irradiated at 69, 100 and 125 Mrad

Not yet available

Figure 53: Threshold change of PMOS transistors following ^{60}Co irradiation of 60, 100 and 125–Mrad, for the Harris AVLSI–RA process

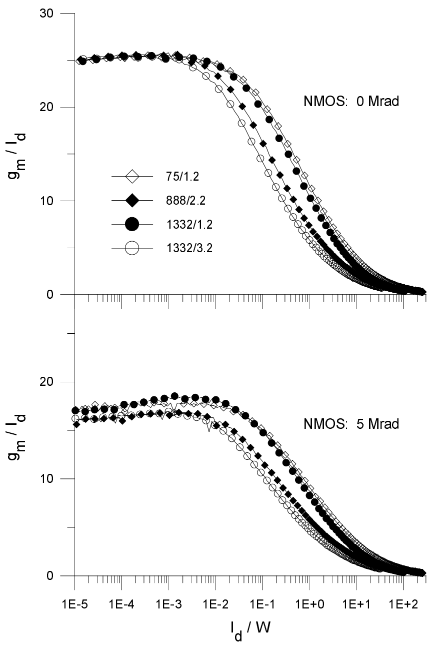


Figure 54: Change in NMOS transconductance vs. normalised drain current before and after irradiation to 5 Mrad(Si).

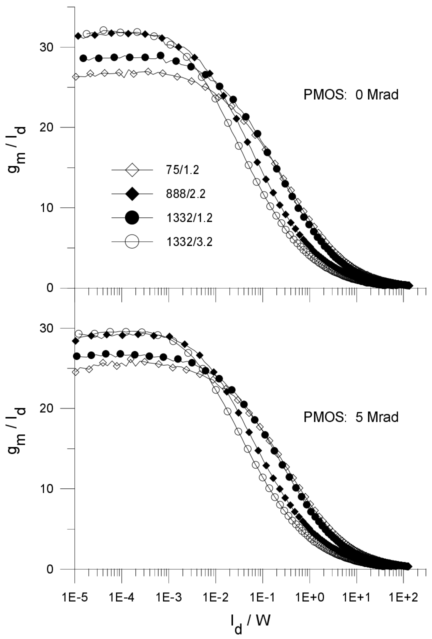


Figure 55: Change in PMOS transconductance vs. normalised drain current before and after irradiation to 5 Mrad(Si).

for the CDF silicon vertex detector has been transferred to the rad- hard UTMC process. SVX-H ICs are installed in both CDF [48] and L3 [49] and are providing excellent results.

Radiation hard CMOS bulk processes–Honeywell: The Honeywell Solid State Electronic Center offers a radiation-hard bulk CMOS process with a 0.8 μm feature size and triple metal. The maximum die size is 15.5 x 15.5 mm^2, but maximum resolution is only guaranteed within a circle of 8 mm radius, which allows dice of about 12 x 12 mm^2. This is an n-well process, which provides a good match to the HP 0.8 μm process offered by MOSIS, so that prototype circuits can be designed and debugged in rad-soft multi-project runs. 64- and 128-channel clock-driven pipelines have been designed at UCSC and fabricated for the binary readout [53]. Test structures and preamplifiers have also been designed (LBL) and fabricated to assess analogue design models and noise performance. Measurement data are not yet available. This is also primarily a 'digital' process, but it is anticipated that results similar to UTMC will be obtained in analogue applications.

Summary of CMOS bulk demonstration circuits:

5.4.3 CMOS SOI processes

CMOS-SOI radiation hard processes are studied at CERN in the framework of two DRDC programmes; RD9 for the HSOI3HD 1.2 μm and HSOI4HD 0.8 μm processes, and RD29 for DMILL BiCMOS. A common key feature of SOI processes is their intrinsic robustness against Single Event Latch up(SEL) and Single Event Upset (SEU). The typical energy threshold for SEU is in the order of 120 MeV.mg^{-1}.cm^{-2} (HSOI3HD) and the total dielectric isolation of each transistor makes ICs on the SOI process latch–up free. In addition, the dielectric isolation of source and drain junctions ensures a very low leakage current after bulk damage.

CMOS SOI processes–Thomson HSOI3HD thin film process: The HSOI3-HD technology is a polycide gate CMOS-SOI process. The SIMOX substrate is realised by 200-202-KeV O^+ ion implantation at 500-700 ^{o}C followed by annealing at 1300-1350 ^{o}C. Isolation between devices is achieved by conventional LOCOS, and a LDD (Lightly Doped Drain) structure is used to lessen impact ionisation due to high electric fields. The thickness of the gate oxide, silicon body and buried oxide are respectively 23 nm, 150 nm and 380 nm, and the transistor body is partially depleted. The minimum line width in this technology is 1.2 μm (size of contacts, bias and interconnections), but the minimum drawn gate length allowed is 1.4 μm, and 0.8 μm for HSOI4HD for all line width.

^{60}Co irradiation of transistors: The analogue performance of HSOI3-HD has been measured on various test transistors up to a total dose of 12 Mrad(Si) of ionising radiation (^{60}Co)[54],[55]. The threshold voltage shift is -170 mV for p-channel and -20 mV for n-channel transistors. Transconductance degradation is respectively 4% and 17%. Noise has been measured in the 500 Hz–25 MHz bandwidth. The white noise increase after irradiation is not significant; n-channel transistors, and particularly p-channel transistors exhibit a low 1/f noise . The corner noise frequency is 800 KHz for n-channel and 100 KHz for p-channel for an aspect ratio W/L=1000/2 and a drain current of 250μA. It has been verified that the additional G-R noise caused by the back-gate is not significant for a proper biasing condition of the body.

DC parameters and noise measurements: The noise is studies using large transistors with high transconductance to have a low equivalent input voltage noise density. The MOS-SOI n- and p-channel transistors studied were actually five-terminal devices. The width of the measured transistors was W=1000μm, and two different gate lengths were studied: L=1.4μm and L=2μm. Gamma irradiation took place with a calibrated ^{60}Co source at room temperature. It was performed in four steps with pauses at 1.5, 3.5, 7.5 and 12 Mrad(Si). The dose rates were respectively 18, 31, 54 and 54 rad/sec, and 5 hours interruption between each exposure was kept to perform static measurements. The uncertainty on the dose was 20%. In Table 11 we summarise the threshold voltage, transconductance and mobility of n- and p-channel transistors before and after a 12 Mrad(Si) exposure

In Table 12, the results of the white noise measurements before and after irradiation are reported. The measured thermal noise is given together with the corresponding g_m and the extracted G. The body potential

Transistor static parameters.			
	V_{th}(V)	gm (mS)	μ (cm^2/V_s)
p-channel prerad	-1.31	3.76	200
p-channel 12Mrad	-11.48	3.63	194
p-channel variation	-0.17	-3.4%	-3%
n-channel prerad	0.85	4.12	357
n-channel 12Mrad	0.83	3.41	263
n-channel variation	-0.02	-17%	-26%

Table 11: Summary of the static parameters of transistors with W=1000 μm and L=1.4 μm before and after irradiation to 12 Mrad and subsequent annealing, using the HSOI3HD process.

Transistor noise parameters.				
	n-chan L=1.4μm	n-chan L=2μm	p-chan L=1.4μm	p-chan L=2μm
S_v prerad [nV/Hz]	1.98	2.02	2.16	2.21
S_v 12Mrad [nV/Hz]	2.73	2.66	2.50	2.57
ΔS_v	+38%	+32%	+16%	+16%
gm prerad [mS]	4.58	4.22	4.52	3.90
gm 12Mrad [mS]	4.00	3.55	4.20	3.60
Δgm	-13%	-16%	-7%	-8%
G prerad	1.07	1.03	1.26	1.13
G 12 Mrad	1.78	1.49	1.56	1.41
ΔG	+66%	+45%	+23%	+25%

Table 12: White noise, transconductance and G factor for n- and p-channel transistors before and after irradiation and annealing for the HSOI3HD process. V_{body}=-0.7 V for the n-channel and 0.9 V for the p-channel transistors. V_{bg}=0 V in all cases.

during measurement was -0.7 V for the n-channel and 0.9 V for the p-channel transistors. The back-gate potential was in every case 0 V.

Table 13 shows the variation of the 1/f coefficient before and after a 12 Mrad irradiation.

Neutron irradiations: The results of the neutron irradiation at 3.10^{14} $n.cm^{-2}$ should be compared with data following a 1 Mrad ^{60}Co irradiation, since this dose accompanied the neutrons. For the n-channel device we measured a smaller threshold voltage shift for the neutron exposure. After the 3.10^{14} $n.cm^2$ with its accompanying 1 Mrad ionising radiation, the ΔV_{th} of p-channel devices,was larger by 160 mV when compared with a pure 1 Mrad gamma irradiation. This indicates that the net effect of the neutron damage is small. The variation of the threshold voltage and drain leakage current caused by neutrons of the back-gate is similar to the ^{60}Co results for n and p-channel. This indicates little neutron contribution for those parameters.

Proton irradiations: Bot n- and p- transistors of dimension 1000:1.4 and 1000:2 have been irradiated to 2.2 10^{14} $protons.cm^{-2}$. The drift of the DC parameters exhibits a figure almost equivalent to the 10 Mrad ^{60}Co irradiation. This indicates no specific damage due to proton. The noise has been studied and results indicate that the p-channel device exhibits an additional increase of corner noise frequency; 100 KHz prerad, 250 KHz post rad. The n-channel device exhibits no specific corner noise frequency increase due to protons certainly because of the higher pre–rad value, 800 KHz which hides irradiation effect.

1/f parameter before and after irradiation.				
	a prerad	a 12Mrad	$K_f(fC^2/\mu m^2)$ prerad	12Mrad
n-chan $L=1.4\mu m$	0.88	0.94	2.53^{10-9}	8.33^{10-9}
n-chan $L=2\mu m$	0.88	0.92	5.38^{10-9}	9.13^{10-9}
p-chan $L=1.4\mu m$	0.88	0.92	4.7^{10-9}	1.35^{10-9}
p-chan $L=2\mu m$	0.90	0.96	7.2^{10-9}	2.22^{10-9}

Table 13: Extracted 1/f coefficients before and after irradiation and annealing, for the HSOI3HD process.

Summary of SOI demonstration circuits: A 8–bit flash ADC designed by TCS has been irradiated up to 100 Mrad [56],[57] and is still functional with a loss of precision of less than 1 bit. The CERN RD9 [58] program has developed a folded cascade preamplifier, an analogue memory similar to what is required for tracker electronics, a NICON preamplifier, a fast transimpedance low noise preamplifier and with RD2 an 11 bit multi–range ADC (CRIAD) [59]. Preliminary results on the analogue memory demonstrator indicate an excellent robustness after 10 Mrad irradiation. In particular, the absence of an increase of leakage current of source and drain junctions in SOI technology (total isolation) plays an important role in the stability of the droop rate of the memory capacitor after irradiation. After 10 Mrad, the analogue memory demonstrator was fully functional at 40 MHz write frequency, without noticeable degradation of analogue performances.

BiCMOS SOI processes–the DMILL BiCMOS SOI thick film: DMILL technology integrates monolithically mixed analog-digital CMOS, PJFET and NPN bipolar transistors, on a SOI substrate [37]-[55]. CMOS and bipolar transistors, exhibit electrical characteristics close to that of advanced BiCMOS technologies. This technology is optimised to obtain on the same SIMOX substrate an excellent and hardened NPN device and a $0.8\mu m$ CMOS process which enable the design of mixed mode rad hard circuits in BICMOS. In addition a PJET is available. DMILL technology use in a thick active silicon film (1.2 μm thickness) which allow for a low resistivity body contact for back-gate shielding. This low resistivity body contact is also required to bias the body of CMOS in order to obtain a very low thermal noise.

Neutron irradiation (1 MeV): As it is expected for rad hard CMOS processes, the transistor characteristics were not significantly affected by neutron irradiation, even after a fluence of 5.10^{14} $n.cm^{-2}$.

Gamma irradiation (^{60}Co): Irradiation conditions: dose rate = 3 Mrad/hour, gate bias=+5 volts for NMOS and −5 volts for PMOS (worst cases), drain bias = 0 volts, body bias = 0 volts. After a total dose of 17 Mrad (SiO2), the threshold voltage shift measured on these devices is less than 90 mv for NMOS and less than 250 mv for PMOS (Figure 56).

Noise measurements: Measured before irradiation on PMOS W/L= $5000\mu m/3\mu m$ at drain current of $100\mu A$ and a substrate voltage of −4 volts (NMOS) and +4 volts (PMOS), show a thermal noise Sn = 2.5 nV/Hz for NMOS and Sp = 2.4 nV/Hz for PMOS, which correspond to noise factors Gn = 0.93 for NMOS and Gp = 0.66 for PMOS(The ideal theoretical noise factor is 2/3. Corner frequencies measured before irradiation are Fc = 16 KHz for NMOS and Fc = 2.6 KHz for PMOS.

After 10 Mrad (SiO_2), the noise factor of the thermal noise, and the corner frequencies slightly vary: Sn = 3.3 nV/Hz, g = 1.19 and Fc = 40 KHz for NMOS (Figure 57), Sp = 2.9 nV/Hz, g = 0.88 and Fc = 6.5 KHz for PMOS (see Figure 58). This corresponds to an excellent low noise characteristic, even after very high irradiation doses.

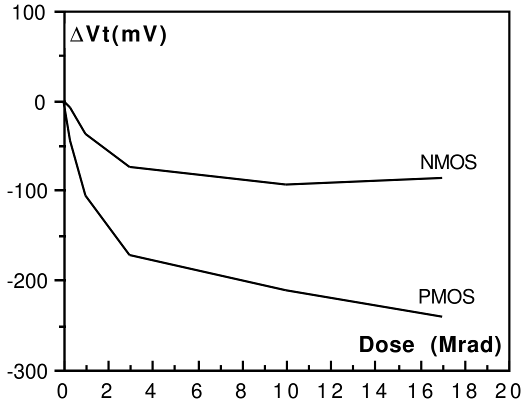


Figure 56: Variation of ΔV_T with dose of up to 20 Mrad, n- and p- channels, for the DMILL BiCMOS process.

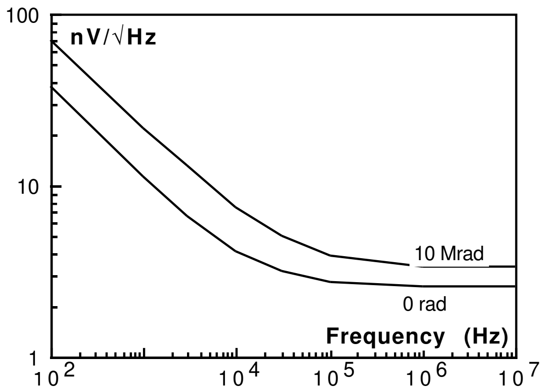


Figure 57: The noise spectrum for an N-channel device of length:width 5000:3 μm., using the DMILL BiCMOS process.

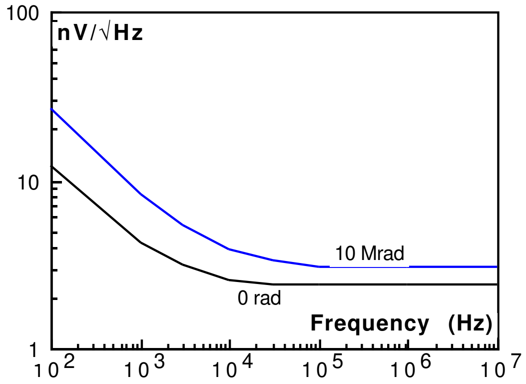


Figure 58: The noise spectrum for a P-channel device of length:width 5000:3 μm., using the DMILL BiCMOS process.

Summary of DMILL BiCMOS SOI demonstration circuits: The DMILL technology exhibits an excellent radiation hardness with in addition a compatible high current gain 6 Ghz NPN bipolar and an extremely low noise figure for the PMOS and NMOS. This technology is presently in development in LETI-CEA-F, with an expected industrialisation phase beginning in 1996. Many demonstrator chips, such as low noise bipolar transimpedance preamplifier and flash ADC have submitted and will be tested end of 1994.

5.4.4 Radiation hard silicon on sapphire processes

A Silicon On Sapphire (SOS) technology from ABB Hafo, Sweden, is being investigated for radiation hardness. Results from the present version of the process, SOS4, has shown significant threshold voltage shifts after gamma irradiation, 1 to 2 V at 10 Mrad, probably caused by insufficient control over the Silicon-Sapphire interface in the processing. Its packing density at feature sizes of 2 μm results in a chip area 2 to 3 times larger than other proposed technologies when applied to a FELix-type preamplifier shaper with a channel pitch on chip of 49 μm (this includes the increase in input transistor size to decrease low frequency noise). Neutron fluxes of up to $2 \cdot 10^{13} n/(cm)^2$ has not shown significant damage effects. The low frequency noise of SOS-processes is somewhat higher than for bulk processes, but would not be a major problem in high frequency applications. Other groups, working with digital applications, have shown in initial tests with presamples of the new ABB Hafo process SOS5 with feature size 1.25 μm (expected to be commercially available in the beginning of 1995), that full functionality of digital circuits is maintained at gamma doses of approx. 10 Mrad. From this can be inferred that threshold shifts in SOS5 are significantly lower than in SOS4. To calculate the increase in packing density with SOS5 compared to SOS4 for an analogue front–end requires detailed study of a specific topology, but for random logic the factor is approximately 2 to 3. Measurements of detailed characteristics of test structures for analog electronics, before and after various irradiation conditions, will commence as soon as the developments of the new process are finished (expected in December 1994).

5.4.5 Procedure of qualification.

Qualification is done primarily by assessing relevant prototype devices fabricated on the process. The criteria and methods for this assessment follow:

1. Power pre/post radiation.
2. Circuit density.
3. Cost, yield.
4. Radiation hardness overall change in parameters.
5. Ultimate failure dose.
6. Repeatability of IC performance.
7. Foundry access and traceability and process monitor information.

5.5 OptoElectronic data transmission options and development for the ATLAS SCT

Data transfer from the front–end is only feasible via optical fibre links, which reduce the volume of material by a factor of ten compared to copper twisted pairs and the number of radiation lengths of material by a factor one hundred. Ground loops, electronic pickup between channels and high power drivers are eliminated. Possible means of converting the electronic signals at the front–end to optical signals rely on either active elements (lasers or LEDs) or passive modulators. In the case of modulators the light source is a laser at the electronics barrack and the light is piped by fibre to the modulator. Electronic signals applied to the modulator affect the light level it transmits or reflects; and this modulated signal is returned to the electronics barrack via optical fibre. Of the two possible modulator technologies investigated the RD23 collaboration (refs. [60]– [61]) has found that, while both are functionally satisfactory, that based on Multi-quantum well modulators [62] is preferred because the device size is compact and the potential for using monolithic opto-electronic integrated circuits at the electronic barrack offers cost and manufacturing advantages. The design, construction and proving of this technology has been carried out by the RD23 Collaboration [60]–[63] which includes GEC-Marconi Defence Systems at Stanmore UK.

The development of LED based read-out has been carried out by the universities of Bern and Oxford in collaboration with GCA Fibre-optics Ltd. and GEC-Marconi.

The modulator approach offers an intrinsically radiation hard technology, whereas diodes, which rely on minority carrier processes, are in principle more susceptible to radiation damage. Nevertheless, damage measurements on GCA LEDs show good light efficiency following an irradiation of 10^{14}n/cm^2 and measurements of the LED lifetime after irradiation are encouraging. To date no diode failures have been reported from ten diodes after the equivalent of 10 LHC years running.

A further advantage of modulators is that the level of power consumption and heating is negligible. An LED system for use with digital or binary architectures would require at least 5 μW/channel and an analogue transmission system up to 200 μW/channel. The latter figure could be as much as 10% of the total power requirement for the system.

The LED approach, on the other hand, is a mature technology which allows more confident predictions of performance and cost. Further, the number of transmitters in a physical package can be small, which makes matching the geometry of digital and binary systems easier.

5.5.1 Proposed modulator system.

Data transfer from the front–end will be via monomode optical fibre, with the conversion from electronic to optical signals being performed at the front–end by Multi-quantum Well (MQW) modulators at an optical wavelength of 1550 nm. At the electronic barrack photodiodes on the readout board will recover the electronic signals. Also located at the backend (readout board) will be the lasers that pump the light to the modulators. A single fibre carries the light from the laser to the the modulator and also returns the reflected modulated beam to the photodiode. At the backend an optical splitter couples the laser and photodiode to the fibre. Figure 59 shows the layout. The lasers, photodiodes, couplers and light guides will be constructed in compact Opto-Electronic Integrated Circuits (OEICs).

The optical link transfers the electronic signals from the front–end to the electronics barrack with frequency independent unit gain, zero cross-talk and negligible increase in noise level. It is essentially transparent. There are several architectures being considered for the front–end requiring either analogue or digital signal transfer from the front–end, all of which can use MQW modulators for signal transfer. The link components are intrinsically radiation hard, and a full link will be tested to levels of radiation expected from 10 years of LHC operation. Several modulators are expected to be packaged together, the package having a footprint of 10 mm $\times$ 8 mm or less and a height of 6 mm or less. The whole system - modulator, packaging and fibre does not appreciably add to the radiation lengths of material or mass of the detector. Figure 60 shows the components of a 4-channel modulator package that will be available for beam tests in 1995.

5.5.2 Modulator device characteristics.

The MQW modulators have input capacitance of 1 pF and input resistance of 1 MΩ. Operation is at a bias of -7 V currently with response being linear to 2% for signals of 1.5 V amplitude. Dynamic range is currently $\approx$100:1 with improvement expected for devices now being manufactured. The electric current drawn through a modulator under bias is $\approx 10\ \mu A$. Overall heating from all processes is only 0.3 mW per modulator. This includes Joule heating, dissipation of absorbed light, and light lost in couplings. Bench tests currently in progress at 40 MHz analogue operation are encouraging. The operational temperature range is likely to be asymmetric, 5° to +15°C around the selected optimal temperature; this variation allows both low luminosity running without detector cooling and high luminosity running with cooling. In magnetic fields of up to 4 T there is negligible change in the MQW characteristics. The device structure in III/IV material is inherently radiation hard. Thus far negligible change in the characteristics of MQW material reflectivity has been observed at neutron fluences of a few $10^{14}\ n/cm^2$; exposure to gamma radiation at LHC ten year levels is now in progress. Radiation hard fibres and connectors are available [64]

5.5.3 Alternative modulator architectures

With analogue signal transfer from the front–end each modulator would have 128 electronic analogue levels multiplexed onto it at 40 MHz (i.e. each level presented to the modulator for 25 ns in sequence). This signal

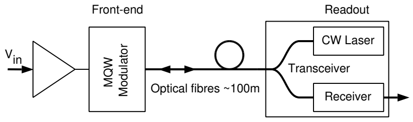


Figure 59: Schematic of the MQW Optical link.

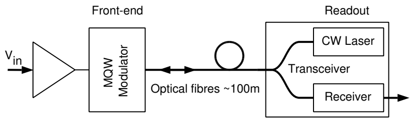


Figure 60: 4-channel MQW package.

train would be preceded by a header of four clock cycles with levels provided by the front–end chip-set: full on, full off, and then two voltage levels (the first high, the second low). The transitions from the first pulse pair would provide an initialisation signal at the readout board, while the second pair would provide calibration pulses for each analogue channel. A package of 8 modulators would service one silicon module of the SCT, that is 1024 strips (electronic channels). Corresponding packaging is foreseen for the GaAs and MSGC readout. The total number of modulator packages is therefore 3500 for the SCT, 1700 for MSGC and smaller numbers for GaAs. This includes spares and reserves at the level of 20%.

With digital signal transfer the output of ADCs on the front–end chip will be multiplexed onto modulators after sparsification. This will typically involve multiplexing from 512 or 1024 electronic channels onto one modulator. In this case there would be either two modulators or one modulator for digital signal transfer in each package mounted on a single SCT silicon module. The total number of modulator packages would be unchanged from the analogue readout case; however each package would have fewer fibres emerging.
Additional fibre connections are necessary between the front–end and the electronics barrack to bring in clock and control signals. It is possible that these lines may use MQW photodiodes at the front–end mounted in the same packages as the modulators or adjacent to them. The design of this facet of opto-electronic signal transfer is not as far advanced as that for signal transfer from the front–end.

5.5.4 BSP testing of a modulator system in 1995 and 1996

The programme already approved by the PPESP in the UK and by the CERN DRDC for RD23 is to have ten complete 4-optical channel links available for testing in Spring 1995. These will be used partly for tests of the full scale analogue readout chain from a single silicon module using the APV5 chip; Birmingham, Cambridge, Liverpool and Oxford Universities, Queen Mary Westfield College, London University, and Rutherford Appleton Laboratory are building this full scale analogue chain and will install it for beam tests at CERN in Spring 1995. These opto-electronic links will have discrete components (lasers, photodiodes and optical couplers) at the readout board. In parallel the RD23 collaboration will be designing, fabricating and testing OEICs for the back–end. By the end of 1995 RD23 aim to complete all the necessary tests on prototype MQW multichannel modulator packages, including irradiation, temperature and magnetic field response, and also to have prototype OEICS. The final prototype MQW modulators and packaging incorporating experience from bench and beam test results will then be available. Thereafter, 50 of the final protoptype 8-channel links would need to be purchased by ATLAS for the BSP programme in 1996 [65].

For the full analogue readout architecture these links (one per silicon module of 1024 strips) would have 8 fibres carrying analogue signals out. In the event that the full digital readout architecture is preferred then a similar number (50) of links would be needed for the BSP programme, each with 1 or 2 fibres carrying out the digital signals.

Additional control, Level–1 trigger and clock signals would be routed on optical fibre from the back end to the front–end, and one possibility being actively investigated is that MQW photodiodes (possibly on the same substrate as the modulators) could be used to convert the inboard optical signals to electronic signals. Additional modulators and fibres could be included in the modulator packages to provide spares should this appear prudent. In the next section a detailed description of a specific, but not final, analogue link is given.

5.5.5 Full analogue link.

These are the components of a possible full analogue link for a single silicon module with 1024 strips and area 6 cm x 12 cm:

- 8-channel modulator and 2 channel photodiode assembly with 2 spare modulators; with 10 m long radhard monomode 12–way fibre ribbon cable terminated by MT 12–way snap fastening connector, also radhard.

- 90 m long 12–way monomode optical fibre ribbon cable with MT 12–way snap fastening connectors at each end; ribbons will have protective sheathing and be merged into ∼1000 fibre cables over the run from the detector to the electronics barrack,

- opto-electronic integrated circuit (OEIC) incorporating 8 channel photodiode receiver and transimpedance amplifier, bandwidth 100 MHz, 1550 nm laser source(s); 2 channel transmitter for clock and control signals transfer to front–end at 40 MHz; 12–way monomode optical fibre pigtail terminating in 12–way MT snap connector on an optical backplane.

Parameter	Range/Performance and Commentary
Dynamic range	$\geq$128:1
Bias	-7V now, target -3V
Max voltage swing with 2% linearity	$\pm 1.5V$
Operational modulator:	
temperature range	$\pm 10^0$C around selected temperature
humidity range	$40 - 50\%$ expected
Operational transceiver:	
temperature range	-5 to $+15^0$C
humidity	$\leq$ 50% humidity
Magnetic field range	0 to above 4T
Radiation resistance	resistant to 2 $10^{14} n/cm^2$ and to γ/ electron and charged hadrons to 10 MRads; the full 10 year LHC dosages
Operational wavelength	$\approx$ 1550 nm
System bandwidth	$\geq$ 100 MHz with cutoff at $\geq$ 10 KHz for test purposes with ripple $\leq \pm 2dB$
Rise/fall time	$\leq 5ns$ from 10 to 90%
Cross-talk	$\leq$ 1%
Analogue output	defined with RAL ($\sim 0 - 2V$)
Modulator heating per optical channel	$\leq 300\mu W$
8-modulator package	8 mm x10 mm footprint $\leq$ 6mm high fibre entry through 6 mm x10 mm face nonmagnetic low atomic number material silicon and ceramics preferred
Transceiver unit (at readout end)	No external cooling shall be required
Optical Connections	All ribbon fibre cable connectors with angle polished ends MT-12 Europtics connectors
Interchangeability	All modulator and transceiver to be interchangeable without adjustment
Link tolerance	all gains within $\pm$10%
Laser Safety	Link will be a designated class 1 laser product; with interlocks if a class 3b laser is needed
Link length	for purposes of bid assume 100m overall
Failure rate	at level of 1% over 10 years of LHC operation

Table 14: Table of modulator system characteristics.

Table 14 summarises the design system characteristics for a modulator system.

5.5.6 The LED–PIN system.

Industry has been using LED-PIN diode links for data transmission for many years. The standard is no longer such links, but rather laser-PIN systems because they have higher bandwidth. In the ATLAS application to the inner tracker the requirement is not for high bandwidth, but rather a distributed system of medium bandwidth. This makes the use of old fashioned LED-PIN links with multi-mode fibre the better approach. Another advantage over lasers stems from the desire to turn the diodes on/off to minimise power consumption. Relaxation oscillations in lasers at turn-on mean that they are run C.W. and modulated, which is not compatible with minimum power usage. In addition lasers often have limited lifetimes and their properties under irradiation are less well known than for LEDs.

Radiation hard LEDs and PINs, characterised by the manufacturer have been available for several years. The issues in developing such links for use in ATLAS are not ones of fundamental process and device development, but rather the details of packaging and of amplifiers and driver circuit development. In particular it is necessary to develop a suitable non-magnetic, low mass and economic package. Two such packages have been developed; one by GCA and the other by GEC. The GCA package, with a ABB HAFO diode has been radiation- and lifetime–tested. The link is capable of 80Mbit/s digital data transfer and has also been run as an analogue link with sampling rates compatible with the inner detector. with its amplifier mounted on the same substrate as the diode. Figure 61 shows the performance of the GCA diode for analogue data transmission. Figure 62 shows the same diode performance for digital data transmission. Figure 63 shows the light output from GCA diodes after irradiation, as a function of time.

The diodes can be fabricated with multiple LEDs per physical package, much like modulators; The only concern is caused by not knowing the ATLAS requirement in sufficient detail to go to the final design.

5.6 Read–out control for the ATLAS SCT

5.6.1 Read–out and control protocol for the ATLAS SCT

This section describes a proposed protocol to be used by the ATLAS SCT for transmitting data from the detector to the DAQ/Trigger and monitoring systems located off detector. Another protocol is also proposed for sending trigger and control function commands to the on-detector electronics. This work was started in order to provide guidelines for on-detector ICs which must be designed and prototyped in the near future to meet the expected construction schedule of the silicon microstrip subsystem. Thought was given to the particular requirements of that subsystem where the on-detector part of the data transmission system must be inside the active volume of the detector requiring low mass, low power and radiation hard components. However, the protocols at this level of definition are quite general so that they can accommodate all three of the readout architectures still being developed for the microstrip detector. In a sense they provide a shell to surround data packets of potentially different content and form. Because of this generality, the protocols could be adaptable to other subsystems of the ATLAS Inner Detector.

Several assumptions were made in designing these protocols and these are enumerated below.

1. The data will be transmitted from and to the detector via serial links, most probably optical. And these links will be highly segmented around the detector.
2. The bit error rate of these links is expected to be very low. Certainly $< 10^{-11}$ but more likely $< 10^{-14}$. The links going to the detector are expected to have considerably lower error rates since the transmitters can be operated at higher light outputs from outside the detector where power is not a great concern.
3. Because of this low error rate, no attempt is made in the data transmission protocol to detect or correct errors. Some attempt is made, however, to minimize the confusion of data packet boundaries due to single bit errors. A single bit error should not cause the loss of more than one data packet or the erroneous merging of two.
4. Likewise, with the control protocol, no attempt is made to include error detection or correction but some attempt is made to minimize problems due to single bit errors by choice of bit patterns.
5. One possible implementation of optical links will use LED transmitters. To remain compatible with that technology and maximize LED lifetime, a data format is used to allow idle states which do not drive light or make level transitions.
6. It is assumed that the clocks at both ends of the data transmission line have a common source and are, therefore, synchronized. Synchronization of received data with the local clock only requires establishing a constant phase shift which can be done via an initialization process. The data format does not have to provide a clock recovery scheme and is chosen not to since most such schemes require more transitions per bit transmitted and larger bandwidth.
7. The frequency of issuing Control Function Commands to the on-detector electronics will be low enough and the length of the command sequence short enough that disabling Level–1 Triggers during the transmission of such commands (not the execution of the commands necessarily but only the transmission of the command to the on-detector electronics) will not introduce an unacceptable dead time.
8. The trigger logic will force a gap of at least 2 beam crossing between any two Level–1 Triggers. The combination of (5) and (6) allows the Control Functions to be transmitted on the same link as the Level–1 Triggers.

The specifics of a proposed protocol meeting these requirements is given in Appendix B.

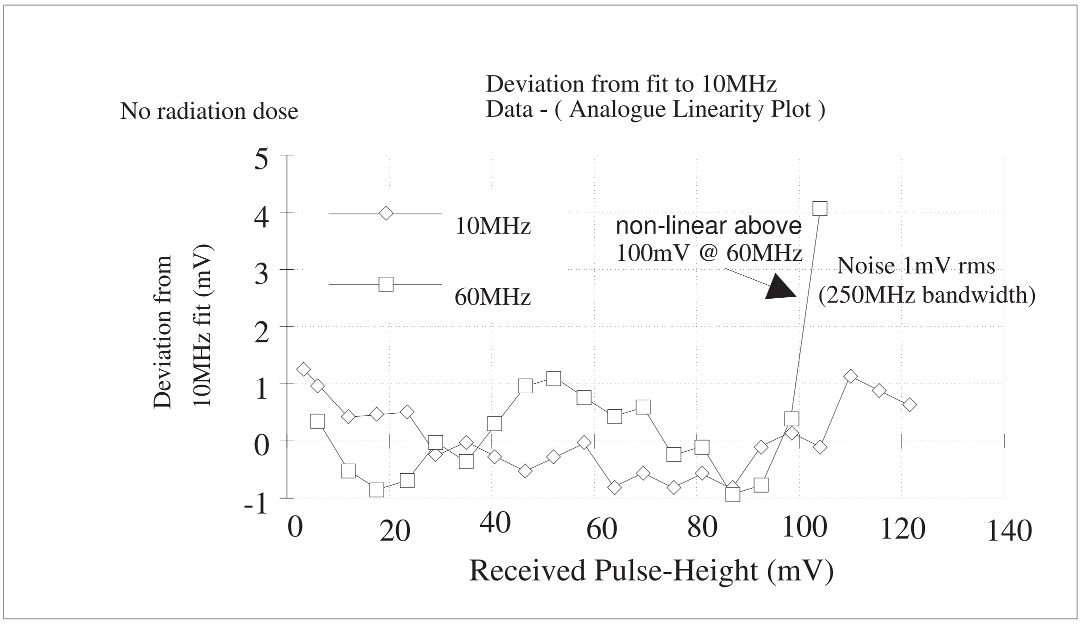


Figure 61: The GCA diode analogue performance.

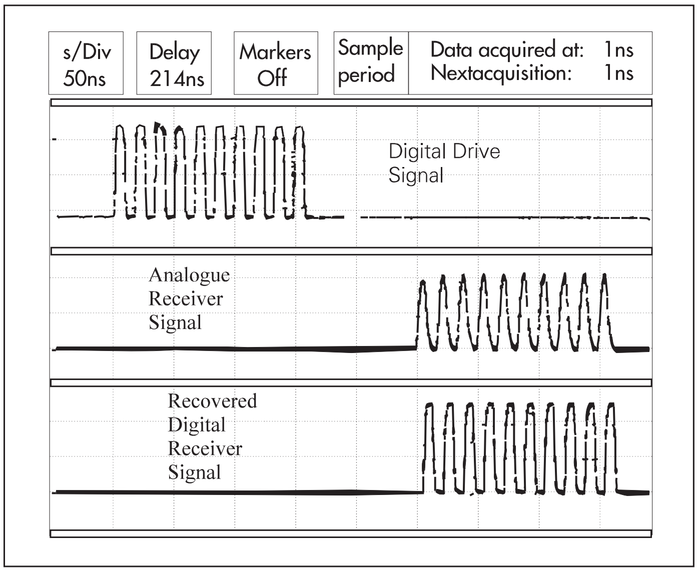
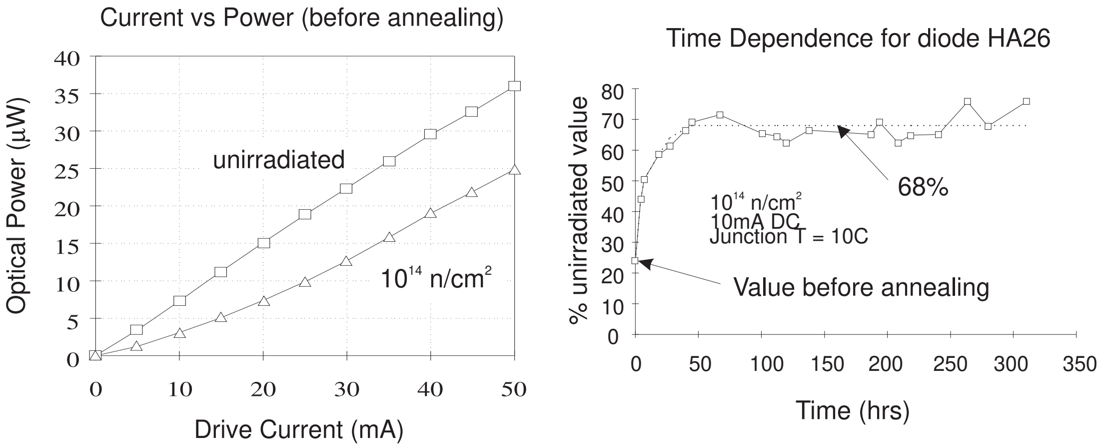


Figure 62: The GCA diode digital performance, operated at ATLAS speed.



Data for a single diode showing response to a dose of 10^{14} n/cm^2

Figure 63: The effect of neutron irradiation on the GCA diode.

5.7 Monitoring, calibration, and control

The control and monitor (CM) of a detector module can be seen as a set of functions operating the module. The functions can be of two kinds: control when data is sent to the module and monitor when data is received from the module. A control function can be a bias setting of a front–end chip or a masking pattern of bad channels. A monitor function can be a reading of a bias current or a temperature reading. The changes in the operating conditions of a detector module can be split in two types: global and local. The global changes are generally slow, caused by radiation or heat. If a global change is excessive the physics data from the tracker may be degraded. A global change may damage the tracker completely. A local change is fast, typically propagated by a single channel, a front–end chip or a detector module failure. Such a failure is local and and will only partly degrade the data. A CM circuit which can check the module regularly but not necessary constantly is sufficient to avoid global problems. On the other hand the number of CM functions needed will be big, requiring each function to be fast. A minimum set of functions necessary for a reliable operation of a SCT module are stable biasing, thresholds for discriminators, masking of bad channels and calibration test of the front–end. At present LEP experiments the CM circuits are located in the control room with separate wires connecting the experiment for each function. In the LHC environment the material budget and the complexity of the experiment will limit the amount of cables that connect the control room to each detector module. The CM can be integrated on the detector module either in a separate CM ASIC chip or integrated as functions in the front–end chip. Some control functions such as chip biasing and thresholds are probably better to implement in the front–end chip. The functions such as bias and temperature monitoring are preferably made only once on every hybrid.

5.7.1 Integration of CM on detector modules

Figure 64 shows a schematic layout of the building blocks of the CM circuit. The question of whether different functions should be built on separate chips is still to be decided.

A circuit controlling the optical link will split the CM data from the BCO (Beam Cross Over) and Level–1 trigger (T1). The same circuit will also control whether data from the CM chip can be sent to the control room.

The CM circuit, shown in Figure 65, is built with a serial bus which serves both the slave circuits on the CM chip and the distributed slave circuits on the front–end chips. A Bus Handler (BH) controls the traffic on the CM bus and the communication with the control room. The slave circuits performing specific control or monitoring functions are placed on the CM serial bus. Each slave circuit has an address that can be recognised in the data protocol. The modular structure of the CM concept allows extensions to satisfy future needs. The slave circuits are specialised either for control or monitor tasks. The aim is to minimise the amount of data that must be sent on the optical link during data taking. This can be achieved by using two protocols. The minimal communication that is needed during data taking will follow a "short" protocol allowing transmission of warnings and important control commands. The circuit will however continuously monitor biasing and temperature conditions on the detector module and send warnings to the control room when a value is outside a predefined range. On request from the control room all biases and temperatures monitored by the CM chip can be transmitted following a "long" protocol.

5.7.2 Protocol for CM

The CM signals will be transmitted to the control room through a separate line for each detector module. The CM data will be folded with the 40 MHz BCO clock. The communication between the control room and the CM chip will always follow a pattern with a request sent from the control room and a replay returned from the CM circuit.

There are two types of protocols needed for the CM. The "short" protocol can easily be sent between triggers. The "long" protocol is needed when the bias settings are transmitted to the front–end DAC and when the monitoring data is transmitted to the control room. The amount of data to be transmitted is big and the transmission is preferably done outside data taking. Low priority data is always sent by the "long" protocol.

The "short" protocol consists of:

- 5 bits preamble, 11101
- 1 bit data type identifier, 1
- 4 bits header/command, 0000-1110
- 9 bits trailer, 1000000000

During data taking the CM circuit is regularly receiving data from the control room a few times every minute. The data sent from the control room may be any of the commands listed in Table 15 except for the long protocol header. After data is received from the control room the CM circuit will return high priority data. The possible high priority data is listed in Table 16. If the HPB is empty the idle command is returned to the control room. A typical state during data taking is that no actions are requested from the control room in which case a idle command is sent. If the detector module is running well there is no high priority data and a idle command is returned to the control room.

The "long" protocol consist of:

- 5 bits preamble, 11101
- 1 bit data type identifier, 1
- 4 bits header, 1111
- (12 bits header)
- 4 bits book keeping number

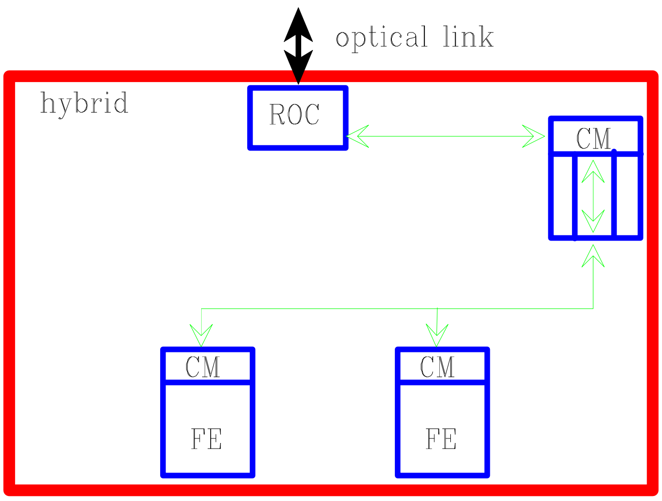


Figure 64: A schematic layout of a hybrid showing the serial bus and the distributed CM.

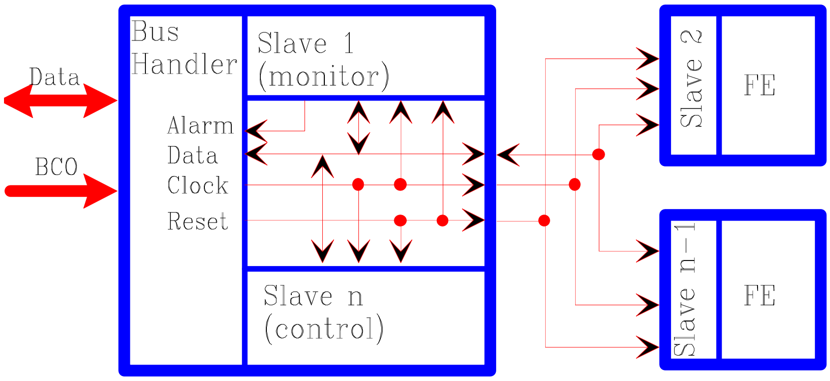


Figure 65: A schematic layout of the CM bus.

Idle	0000
Reset	0001
	0010
	0011
	0100
	0101
	0110
	0111
	1000
	1001
	1010
	1011
	1100
Long protocol from CM disabled	1101
Long protocol for CM enable	1110
Long protocol enable	1111

Table 15: Broadcast command sent from the control room.

Idle	0000
High temperature warning	0001
High Priority Buffer overflow	0010
	0011
	0100
	0101
	0110
	0111
	1000
	1001
	1010
	1011
	1100
	1101
Low priority data waiting	1110
Long protocol follows	1111

Table 16: Short protocol headers received by the control room.

- 4 bits address
- unlimited bits data
- 9 bits trailer, 1000000000

With the "long" protocol all the tuning of the detector module can be done. The "long" protocol is flexible and allows very long strings of data to be sent to the detector module. This is needed for setting up bias voltages and currents and masking of bad channels. When a request or a setting is sent to the CM circuit high priority data is returned to the control room as default. If no high priority data is present then low priority data is returned. If as well low priority data then a idle command is returned. It takes typically some time for the CM circuit to process low priority data after a request from the control room. Some requests may as well need less time for execution than others hence every request from the control room has a book keeping number. The request will be cleared when a reply arrives from the CM circuit with the same book keeping number.

5.7.3 Technical description of CM sub–elements

The Bus Handler (BH): A functional layout of the BH is shown in Figure 66. The BH controls the communication between the control room and slave circuits on the CM bus. The data coming from the control room is put into a buffer and is forwarded to a subelement in the CM circuit when the serial CM bus is free. Warnings and error messages coming from monitoring functions will be sent to the control room on the first transmission available. When the detector module is working properly no error messages will be issued, hence no data will be sent to the control room unless data is explicitly requested.

Clock Input Module (CIM): This subcircuit inside the CM chip receives the BCO clock from the optical link. The signal is transformed to a frequency suitable for the CM circuit and driven out to the other subcircuits in the CM chip. The signal is labeled CLCK.

Input/Output Module (IOM): The CM data traffic inside the bus-handler is controlled by IOM. The data sent from the control room is sent to the Input Buffer. If the High Priority Buffer, HPB, has data High Priority Data, HPD, will be sent to the control room. If "long" protocol is allowed and the High Priority Buffer is empty Low Priority Data, LPD, from the Low Priority Buffer, LPB, will be sent to the control room. The highest priority is given to the High Priority Buffer overflow that will be transmitted before any HPD, LPD. The output buffers, HPB and LPB, sets a flag in the IOM if data present. The IOM controls the buffers by a local a control bus.

Local Data Controller (LDC): The bi-directional data traffic on the CM serial bus is handled by this subcircuit. Data from the input buffer is sent to the slaves and the LDC will wait for a return signal from the requested slave before any other actions taken. Data arriving from the slaves are put into appropriate buffers.

Input Buffer (IB): The data is kept in the IB until the LDC is ready for transmission. If the IB is overflow a OVFLW signal is sent to the HPB.

Low Priority Buffer (LPB): The output data from a monitoring slave will be stored in the LPB until "long" protocol allow the data to be transmitted to the control room. If the LPB is overflow a OVFLW signal is sent to the HPB.

High Priority Buffer (HPB): The important monitoring data are stored in the HPB awaiting transmission to the control room. This buffer will be constantly read by the controller. However if this buffer overflows a OVFLW signal is generated and sent to the control room before any other messages.

Internal Bus (IB): The internal bus consists of essentially two lines one transmitting the serial data to the slave circuits and one transmitting a reset and other control signals. The clock can be taken from the common clock on the detector module.

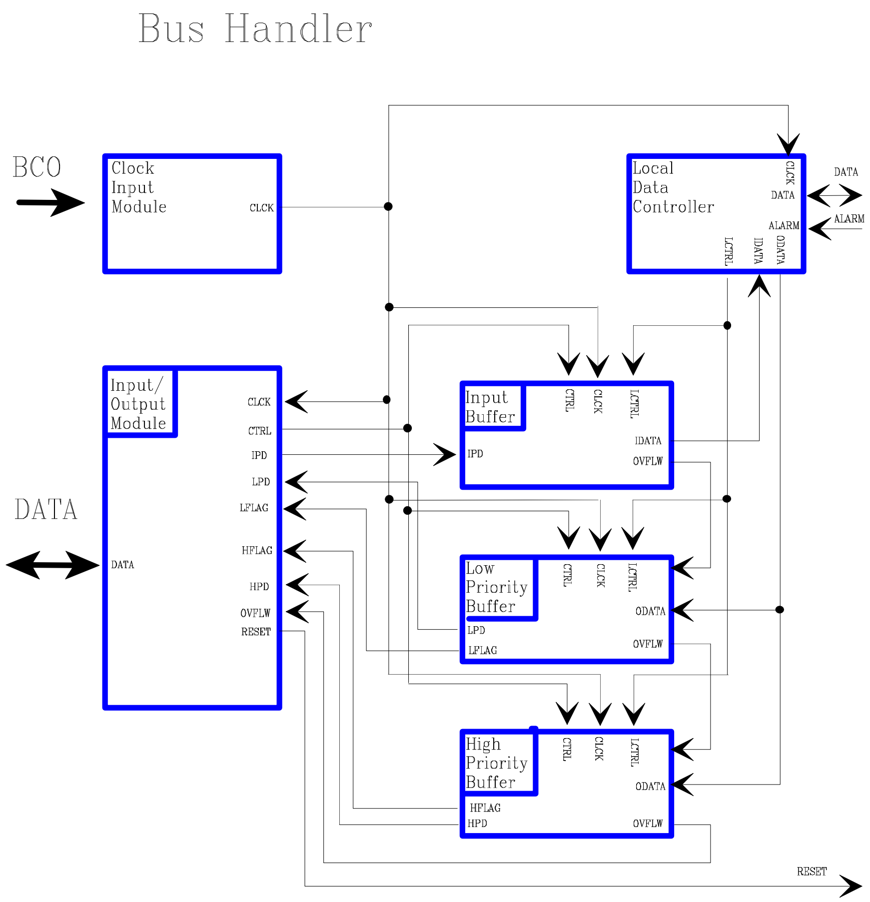


Figure 66: A functional layout of the Bus Handler.

Slave Modules: The slave modules are specialised on a single function. Every slave has a genuine hard-wired address. The slave elements can provide biasing, functionality test, monitoring of voltages, thresholds etc. Table 17 lists the different CM functions with an indication where on the detector module the function is implemented.

Figure 67 shows a slave module for control functions. The control function generator will be a DAC or a register for channel masking etc.

Figure 68 shows a slave module for monitor functions.

- front–end bias, figure
- Thresholds, figure
- Masking of bad channels, figure
- Monitoring and alarm-generator, figure
- Functionality test and testpulse, figure

Control Signals: The internal control signals in the CM circuit are explained in Table 18.

5.7.4 R&D for Slow control and monitoring

In the BSP project prototypes of the CM scheme will be tested. The CM circuit will in the first phase be separated and have no part integrated with the front–end chips. The main advantage with such an arrangement is that it is more flexible, allowing for development of the CM and front–end circuits to be carried out separately. The CM chip development has naturally to follow closely the development of the front–end chips. As there are presently several different designs of front–ends the CM circuit has to be modular and general in the R&D phase so that the CM chip in can accomodate to the needs of different schemes under development.

Milestones:

- 1995

April: final design for CM R&D chip

July: first prototype with basic functions ready mounted on a PCB connected to the hybrid, providing CM for a ATLAS module

- 1996

June: design for specific front–end CM architecture ready

August: second iteration of CM chip ready

- 1997

January: integration of some CM functions in the front–end (if necessary) third iteration of CM chip ready, rad hard technology

5.8 Off detector electronics for the ATLAS SCT

The silicon off-detector electronics interfaces the readout electronics on the silicon modules to the Trigger/DAQ System, providing the silicon module with clock and control signals and receiving event data from the module. It also provides the silicon module with power and other services.

The material in this section is under revision and will expand in a future version of this document.

The off-detector electronics consists of three basic card designs, SCT Service Cards, SCT Clock and Control Cards, and SCT Data Receiver Cards, corresponding to the three basic functions of the off-detector electronics, services, clock and control, and data receipt. Sets of these cards are housed in crates, each controlled by a single board computer or crate CPU. The principal interconnection of these crates is via the SCT Controls Network linking the crate CPU's. The SCT Controls Network may be a subnet of the ATLAS Controls Network. The basic layout of the system is shown in Figure 69

Function	Place
front–end bias	FE
Discriminator threshold	FE
Masking of bad channels	FE
Monitoring front–end bias	CM
Monitoring temperatures	CM
Alarm generator	CM
Functionality test	CM and FE

Table 17: The implementation of CM functions on the detector module. FE= front- end, CM= control and monitoring chip.

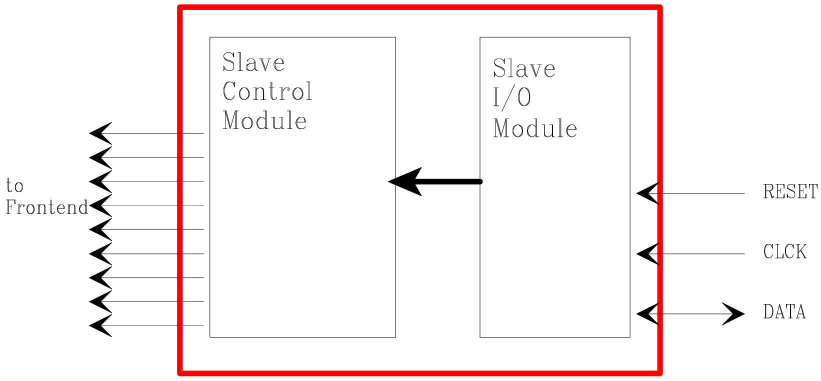


Figure 67: A schematic layout of a controlling slave module.

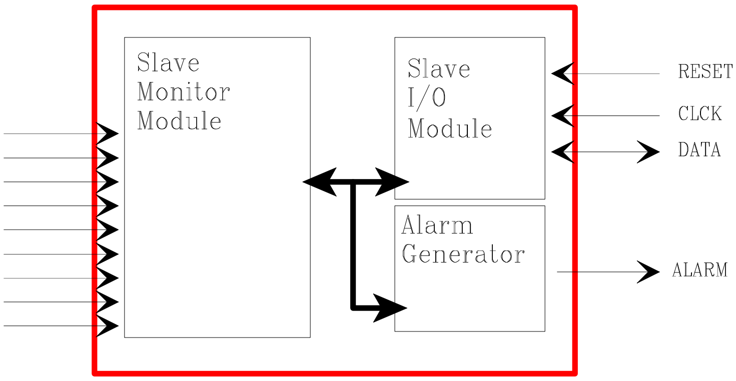


Figure 68: A schematic layout of a monitoring slave module.

5.8.1 Services.

The off–detector electronics provides the following services to the on-detector readout electronics:

- detector bias voltage,
- power for readout electronics,
- readout threshold voltage,
- electronics calibration level,
- readout for module temperature monitoring.

It also provides monitoring and control for these services.

These services are provided to each detector module via a low–mass flat ribbon cable, one cable per module. This cable will carry only DC voltages. It will be a custom design fabricated of kapton with aluminum conductor. The width and radiation length specifications are to be defined at a future time.

5.8.2 Detector bias voltage

Requirements on detector bias supplies, numbers, control, and other requirements.

5.8.3 Power for readout electronics

5.8.4 Readout threshold voltage

The readout threshold voltage will set the minimum signal size required for readout of detector channels. Its use will vary according to the readout scheme adopted.

1. **Binary readout:** For the binary readout scheme, the readout threshold voltage will set the threshold of the comparator at the output of the amplifier-shaper circuit, preceding the Level–1 buffer.

2. **Digital readout:** For the digital readout scheme, the readout threshold voltage controls the sparsification process. Its use will vary according to the digitisation scheme adopted. If all channels are digitised for each Level–1 trigger, then the threshold voltage may be required to set the threshold of a comparator which sets a bit in a register which determines which channels will be read out. If channels are digitised according to the size of the analogue pulse height stored in the Level–1 buffer, then the threshold voltage will be used to set the threshold of the comparator which determines whether the channel is digitised.

3. **Analogue readout:** For the analogue readout scheme, in which it is expected to read out all channels without sparsification, the readout threshold voltage will not be used.

Delivery of the readout threshold voltage via the services cable requires less on-detector complexity than would generation of the threshold voltage by a DAC on the detector module. External generation requires very little increase in size or mass of the cable.

5.8.5 Electronics calibration levels

The electronics calibration level is a DC voltage which determines the size of the calibration pulse to be injected into the input of each preamplifier channel. One voltage sets the charge for all channels in a module, although calibration mask registers control which channels receive pulses. This voltage will be independently settable by the DAC for each module.

Delivery of the electronics calibration level via the services cable requires less on-detector complexity than would generation of the level by a DAC on the detector module. External generation requires very little increase in size or mass of the cable.

5.8.6 Readout for module temperature monitoring.

Readout of DC voltages from thermistors or thermocouples which measure module temperature will be via the service cable. It is expected that there will be one, or at most two, temperatures measured per module. Transmitting these voltages from the detector module for digitisation by the off-detector electronics avoids the complexity of an on-detector ADC for this purpose. This service requires very little increase in size or mass of the cable.

5.8.7 SCT service cards.

Services to the detector modules are provided by a set of SCT Service Cards located in crates near the detector. Each card services a number of detector modules. A single board computer connected to the SCT Controls Network is located in each crate in order to provide control and monitoring for the Service Cards. The crate will be of a low-noise design, possibly VXI, in order to minimise digital noise on the cables to the detector modules. Services to each detector module are provided by a point-to-point connection.

5.8.8 Clock and Control

In this sub-section we review the Clock & Control system proposed for the SCT. In section 5.8.9 we define the requirements of the system. The overall architecture is described in section 5.8.10 and finally possible physical implementations are discussed in section 5.8.11.

5.8.9 System requirements

The primary requirements for the Clock & Control System are that it should distribute the following information to all front end modules:

- The 40 MHz Bunch Crossing Clock. The precision of the clock (i.e. pulse-to-pulse jitter) must be better than 2 ns and the drift rate must be less than 1 ns/hour. The relative phases of clocks should be adjustable so that slow drifts can be corrected for.
- L1 Triggers.
- Control Information. This consists of two types:
  - Fast Control. e.g. global resets and bunch counter resets.
  - Slow Control. This should be able to provide all the information to set-up and operate the front end chips. The exact instructions will off course depend on the type of front end architecture adopted but possible commands would be mask off bad lists of bad channels, force readout of all channels (pedestal mode), change to X-ray alignment mode etc.

5.8.10 Overall architecture

The overall architecture must of course satisfy the requirements given in section 5.8.9 above but it should also have the following features:

- Uniformity. The DAQ system should see a maximally uniform system from the point of view of hardware and software.
- Minimum complexity on the detector. From the point of view of the design of the front end electronics, the amount of electronics that is on the detector must be minimised in order to reduce the cost of the radiation hard electronics, the radiation length and the power consumption. In addition, the use of very high speed signals in close proximity to sensitive amplifiers is undesirable.
- Redundancy. The system should be immune to single point failures.
- Timing adjustment. It is desirable that the clocks for the front–end readout chips can be synchronised without beams so that useful Cosmic Ray data can be taken.

The last two items are discussed in [66]. A possible scheme to satisfy these constraints is shown in Fig. 70.

Signal	From	To	Explanation
CLCK	CIM	IOM,LDC,IB,LPB,HPB	internal clock in the CM circuit
OVFLW	IB,LPM	HPB	buffer overflow signal
OVFLW			
IPD			
LPD			
HPB			
RESET			
LFLAG			
HFLAG			
CTRL			
LCTRL			

Table 18: Explanation of internal signals.

Not yet available

Figure 69: Layout of off detector electronics.

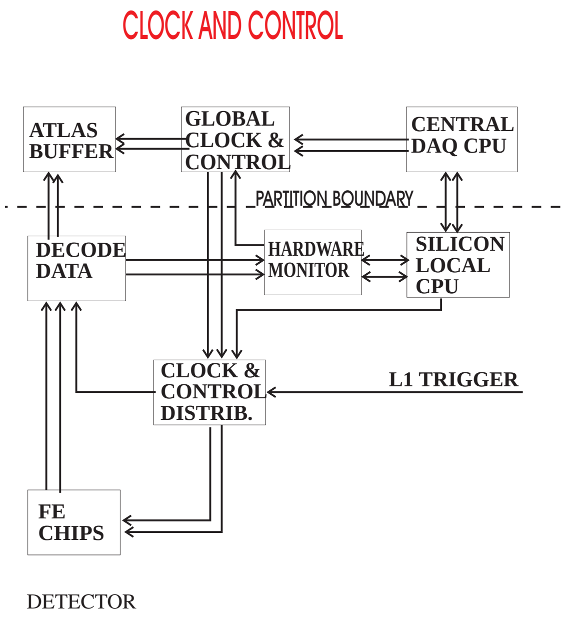


Figure 70: Overall Clock & Control Architecture

The DAQ sends the clock and control information to the interface module using the ATLAS wide DAQ standard. Thus the DAQ system sees a completely uniform system. The inputs to the interface module contain clock clocks, L1 triggers and fast control data which will be fanned out and sent to all modules. The interface module also receives slow control information from the DAQ via a software link. This can be used to send different control data to the different modules. The clean partition boundary between the DAQ and the detector will allow for easy operation of the SCT in standalone mode which will be useful for testing the detector during installation.

5.8.11 Possible implementations

Two possible implementations for the fibre links carrying the control data from the interface module to the front end modules are under development in ATLAS. One scheme is based on passive optical fan-outs and is very well described in [67] so only its main features are outlined here. The control data is encoded and then sent to a modulator which controls the output of a laser. The output of the laser is then fed into an all-optical passive fan-out and the outputs of the fan-out are sent by optical fibre to the front end modules. Currently a system is being evaluated whereby two data channels are time division multiplexed (TDM) and an encoded biphase mark (BPM) at 160 MHz is used. Since a phase lock loop (PLL) (or equivalent) is required in this system the timing of the clock is very precise. An alternative scheme is based on LEDs and is described in [68]. Pulse height modulation (PHM) is used to encode the control data on top of the 40 MHz clock. The data is sent down multi-mode fibre to the front end modules. After optical to electrical conversion by a radiation hard PIN diode the clock and control information are recovered by the DORIC chip [69] which consists of an amplifier and two discriminators. The output of the low (high) discriminator is the clock (control data).

The primary advantages of the passive fan-out scheme are:

- Lower fan-out time which would this reduce the latency of the L1 trigger by a few clock cycles.
- the use of longer wavelengths may make the system more radiation hard.
- it uses mono-mode fibres which are cheaper than multi-mode fibres used in the LED scheme.
- The passive scheme allows one to send two bits of control data with the clock every 25 ns. This would allow one to trigger on successive bunch crossings. This could be achieved in the LED scheme but at a cost of a few bunch crossings of additional latency. However the analogue pipelines will require at least one (two) empty bunch crossings after a L1 trigger for the case of correlated double sampling (pulse deconvolution). In any case the effect of not allowing L1 triggers in the two bunch crossings following an L1 trigger is to introduce a dead time, in the worst case of the maximum L1 trigger rate, of only 0.5%.
- The passive fan-out scheme has an excellent timing accuracy due to the use of th a PLL circuit on the detector. A similar precision could be achieved with the LED scheme if a PLL was used however given the relatively modest timing requirements of the SCT this is probably an unnecessary complication.

The main advantages of the LED scheme are:

- It requires the minimum amount of on-detector electronics. It also requires lower speed and power on-detector electronics.
- In the LED scheme synchronisation of neighbouring modules can be achieved as described in [68] and although this could be achieved in the passive fan-out scheme it would require additional logic on the detector. The same argument also applies to the redundancy issue.
- The LED scheme allows different control data to be sent simultaneously to different modules and it therefore has a higher effective bandwidth although it is not clear if this a very important advantage.

5.8.12 Data receipt

Event data is transmitted from the detector modules via fibre optics in response to a Level–1 trigger. The data is received by SCT Receiver Cards. These cards consist of electro-optics, logic to check data integrity, buffers to make data available parasitically to the crate CPU, and logic to concentrate the serial data streams from several modules into high-bandwidth serial streams which transport the data to generic DAQ Readout Cards, which provide Level–2 buffering and which transmit data to Level–2 and Level–3 processors.

5.9 Quality assurance and testing.

This section is yet to be written.

6 Modules

6.1 Module overview

In this section we shall discuss the specifications, design and development of the basic module sub–units of the SCT barrel silicon strip layers. The development of appropriate modules is a major component of the SCT prototyping program which should be completed about the end of 1995.

At the present time, ATLAS is considering two different module concepts. The first concept is called the "Z–module" because it features electronics placed along the z axis of the barrel. The second concept is called the "Rϕ module" because it features electronics placed along the azimuth. To some degree there is a coupling of these module designs to options in the design of the mechanical support for the barrel layers (see Section 7). For example, once the support structure is fully specified there will be an impact on these module designs. It is however widely believed, within the SCT community, that either module concept could be made to meet ATLAS specifications. The final choice will probably be based upon cooling issues, materials budgets, cost and second order electrical performance issues.

Below we will present first a general discussion of requirements and specifications on module design and then describe the two concepts. A key aspect of the module design is cooling efficiency. An important issue to be discussed is the possibility of thermal runaway effects in irradiated silicon detectors which are not cooled sufficiently.

6.2 Requirements and specifications

6.2.1 Detector

One module is made of two (four) 6 cm $\times$ 6 cm double (single) sided silicon microstrip detectors. In the double sided case the two detectors are to be butt glued to form a 12 cm long mechanical unit and strips of two the detectors are electrically connected to form 12 cm long strips. In the single sided case, the butt gluing is followed by a back to back attachment. One side of the module measures ϕ (the "axial strips") while the other side is rotated by the small stereo angle of 40 mrad (the "stereo strips").

The area of 6 cm $\times$ 6 cm is nearly the maximum to be fabricated in a standard 300 μm thick, 4 inch silicon wafer. The length of one detector can be chosen at 6 cm based on an argument of pitch and a stereo angle of 40 mrad [70]. The active width of the detector depends upon the choice of strip pitch, number of strips and the required spaces for bias ring, guard ring and edge clearance.

6.2.2 Stereo strip detector

There are two ways to make the stereo strips using back to back detectors: using identical sets, mechanically rotate one side by the stereo angle (rotated solution); skew the stereo strip pattern by the stereo angle in a unique rectangular detector (stereo detector solution). In the case of double sided detectors, the latter is the only available solution.

To have continuous coverage in both the ϕ and z directions, adjacent detectors are required to overlap. The stereo solution requires an extra overlapping in ϕ; the rotated solution requires it in z. When strips are read out at the centre of the 12 cm module, two solutions give the same area for the overlap because a detector is square [70]. The cost for having two masks is negligibly small for the production of nearly 18 m^2 of detectors. The stereo solution may be easier for back to back alignment since the fiducial marks can be exactly at the same location. Handling may be easier for the stereo solution since the edges of the detectors are aligned. We may make u and v stereo measurements in different radial layers by flipping the axial and the stereo sides over.

The overlapping of modules in the ϕ direction is shown in Figure 71. For the stereo solution, the required area for the ϕ overlap is a triangular region of width of 2.4 mm. The overlap fraction is 4% over the nominal 6 cm width, in addition to the overlap for alignment and for inactive edge regions. In the z direction, since the edge of the axial and the stereo detectors are aligned, no extra overlapping is required other than the inactive edge region.

As long as the strips are not read out from the ends of the module, the width of the stereo detector can be the same as the width of the axial detector with the same number of strips and pitch. The pitch of the stereo strips (p_s), measured normal to the strips, is slightly narrower than the axial strip pitch (p_a) as

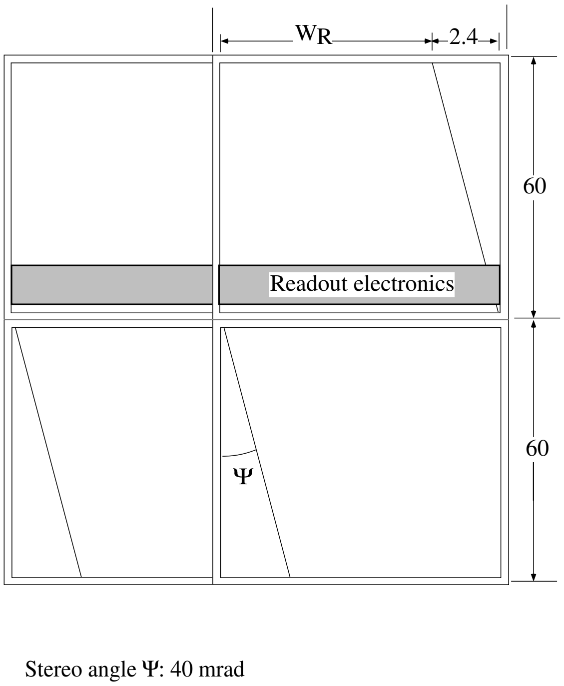


Figure 71: Overlap of the modules in the ϕ direction. Required overlapping is 2.4 mm, in addition to the edge space.

$$p_s = p_a cos\psi. \tag{1}$$

where ψ is the stereo angle.

6.2.3 Tilt angle of the detectors

There are two ways to overlap adjacent detectors in ϕ direction: (1) **tilting**, wherein detectors are rotated by a small angle about a line parallel to the beam at fixed radius; and (2) having adjacent modules tangential to a beam centred cylinder but **offset** in radius. There is an advantage in tilting since it will improve the spatial resolution by creating charge sharing regions.

The maximum charge sharing is achieved by making the charge carriers always spread between two strips. Without a magnetic field, the swath of the carriers $\Delta_\circ$ at a tilt angle χ is

$$\Delta_\circ = d \tan \chi \tag{2}$$

where d is the thickness of the detector (d=300 μm). Equating the swath to the strip pitch p gives the tilt angle of 20.5° (p=112 μm), or 14.0° (p=75 μm). A magnetic field deflects the carriers and changes the relation.

In the case of a single sided detector, only one type of bulk, and thus one type of charge carrier, can be chosen. We take for example, p–bulk and electrons. The deflection distance of the furthest carrier is

$$\Delta_- = \mu_B^e B d. \tag{3}$$

Since the Hall mobility of electrons is large, μ_B^e=1391 cm^{2}V^{-1}s^{-1} [71], the deflection is also large, 83.5 μm, at B=2 Tesla. A tilt angle is required to compensate the deflection so that the corrected swath equals the pitch,

$$p = \Delta = \Delta_- - \Delta_\circ \tag{4}$$

and,

$$\chi = \arctan((\mu_B^e Bd - p)/d). \tag{5}$$

The tilt angle is -9.3° (p=112 μm) or -2.4° (p=75 μm).

In the case of a double sided detector, there are two types of charge carriers. Instead of making the swath equal the pitch, we need to choose the swath of electrons and the swath of holes to be equal to give equal resolutions,

$$\Delta_+ + \Delta_\circ = \Delta_- - \Delta_\circ. \tag{6}$$

The tilt angle becomes

$$\chi = \arctan((\mu_B^e - \mu_B^h)B/2) \tag{7}$$

which is 6.1° using the Hall mobility of holes, μ_B^h=325 cm^{2}V^{-1}s^{-1} [71]. The swath is about 50 μm.

Clearly the tilt angle is a key parameter in the design of the tracker and will, once chosen, effect the overall mechanical design as well.

6.2.4 Required mechanical precision

The required mechanical precision for the module assembly is derived from the required point resolution and positional tolerances on mechanical interference. As discussed in Section 1 and Section 3, ATLAS specifies a statistical point resolution of 20 μm with a systematic error of 10 μm to be added in quadrature. This systematic includes misalignments within a module as well as in the placement of modules on the support. For the construction of a module we specify that the individual crystals should be shifted in ϕ by less than $\pm$12 μm over the length of the module. We further specify that the z position of the crystals should be correct to $\pm$25 μm. Furthermore we expect that an optical survey should be possible to an accuracy of $\pm$5 μm.

6.2.5 Thermal issues

The thermal performance requirements of an SCT module can be specified by considering temperature distributions across the detectors and within the whole module.

Detector Temperature Distribution The temperature distribution on the silicon detector should be such that the annealing and reverse annealing processes of radiation induced defects are well balanced to give a maximum lifetime for the detector. One measure of this balance is the depletion voltage of a detector after it has been exposed to a given fluence of neutrons, as shown in Figure 72 [75]. If we aim for operation always below 200 V, the average temperature should be in a band between -15°C and +5°C. The tolerable deviations from the mean temperature are determined by the maximum acceptable leakage current variation across the detector. To obtain less than a factor of two variation in leakage current due to temperature variations, these must be less than $\pm$3.5°C. All detectors with a common bias supply should be kept in the same temperature envelope.

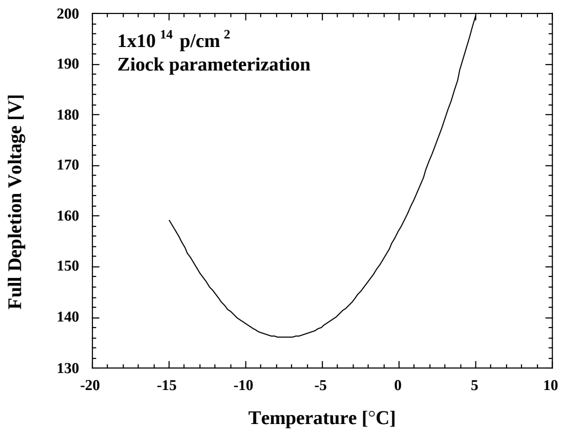


Figure 72: Operating temperature dependance of the full depletion voltage. No warm-up.

A second issue effecting the required detector temperature distribution is the possibility of thermal runaway, introduced in Section 6.9. The process can be understood by a simplified model [74]. We assume some region of the detector is at 0°C, in close contact with a cooling line. Other regions of the detector will, in general, be warmer. Let I_0 be the leakage current in the detector at $T = 0$°C. Then at some other temperature ΔT away we have,

$$I = I_0 e^{\alpha \Delta T}.$$

Here $\alpha = 1/11$ for silicon. The power dissipated for a bias voltage of V_b is IV_b. Because of this, the heat generated in the silicon will be

$$Q = Q_0 e^{\alpha \Delta T}.$$

We assume the silicon is serviced by a cooling system relying on some combination of conduction and convection. The heat removed will be

$$Q_{out} = k_{conduction}\Delta T + k_{convection}\Delta T = K\Delta T.$$

In equilibrium we have (expanding the exponential),

$$\Delta T = Q_0/(K - \alpha Q_0)$$

which clearly diverges for insufficient cooling or excessive heat. The key points are to have a low enough absolute temperature and a small enough gradient across the module. It is not possible to set strict limits based upon a simple estimate. In Section 6.9 we will present results of finite element calculations which indicate various bounds on these temperatures.

Module temperature distribution Restrictions on the temperature profile of the module arise from two design criteria:

- The lifetime of the front–end electronics should not be limited by its operation temperature.
- The mechanical stability of the module should not be affected by the temperature profile.

The first point clearly depends upon the choice of front–end electronics. We expect however that the restrictions due to the detector effects discussed above will be sufficient to protect the electronics. This point awaits further study and confirmation. The second point can only be tested via combined thermal and mechanical analysis of the module through both finite element calculations and simultaneous measurements of temperature and dimensional variation.

6.2.6 Operating temperature

Based upon the considerations given above, we are led to a specification of the operating temperature of the module. As seen in Figure 72, the depletion voltage after a large integrated dose is broadly minimised for an operating temperature of -7°C. For this, and for practical reasons, we specify the operating temperature of the silicon to be 0° C. We specify the gradient across the silicon to be less than 5° C. It is essential that the depletion voltage of the detectors to be used in ATLAS be confirmed at the expected fluence. Some uncertainties in these magnitudes have led to considerations of lower operating temperatures. These discussions are on–going.

On the hybrids and particularly in the readout chips, there will be larger temperature variations. We defer setting a limit on these until a later point in the module and front–end electronics prototyping programs.

Irradiated detectors will be warmed up for repair or maintenance. Both the duration of the warm period and the temperature should be minimised. For example, after the 10 year operating period of ATLAS, a 13 day maintenance at 15°C will result in a 5 V increase in depletion voltage [77].

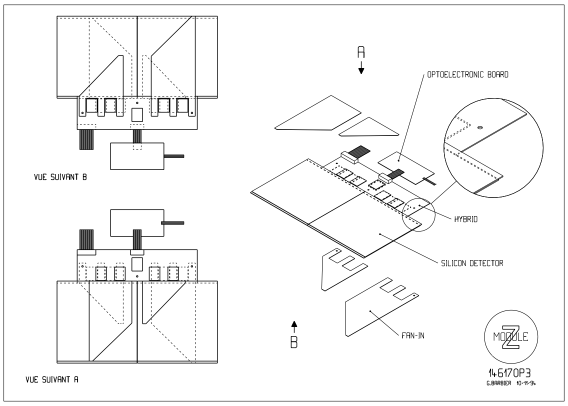


Figure 73: Z module concept. Electronics is off–board of the silicon. A beryllia fan-in makes the interconnection to the readout hybrid assembly.

6.3 Cooling concepts

There are two major problems that each cooling concept has to address.

- The front–end electronics produce a well localised, large power density which has to be removed before it can heat up the silicon detectors. The available surface area for cooling is small (less than 1 cm^2 per chip) and mechanically delicate. The typical power load is around 5 W per module.
- The silicon detectors also dissipate power. The power is distributed across the module but is maximum in the warmest regions of the silicon. For a fully irradiated module, at 30 cm radius, this power can be 400–600 mW/module depending upon the operating voltage.

The mechanical support and cooling concepts under consideration for the barrel silicon modules are described in Section 7. In all cases though, a cooling fluid is piped across the module, passing close to the readout chips. The challenge for the module design and the cooling systems is to ensure that there is also proper cooling of the most distant corners of the module and that thermal runaway is prevented.

One mechanical support concept is based upon longitudinal staves which carry coolant. This works most naturally with a Z module design. The other concept is based upon the use of 1 mm^2 aluminium tubes which carry coolant. In principle these tubes can contact the module in a number of spots. This scheme can work with both Z and Rϕ module designs.

6.4 The Z module

6.4.1 Module concept

Figure 73 shows a design for a "Z" module. The detectors and the hybrid assembly are adjacent to each other. The cooling runs along the hybrid in z and makes contact between readout chips. The front–end electronics and silicon strips are interconnected by beryllia fan–ins which also serve to cool the detectors. Detectors are either back to back single sided or double sided. The hybrid supports the front–end electronics and provides the control, readout and bias lines. The front–end chips are placed on both sides of the hybrid.

One advantage of this design is that the front–end electronics is decoupled electrically and thermally from the silicon. Another advantage is that the cooling paths are short into the detectors.

6.4.2 Hybrid and interconnections

The hybrid is constructed of alumina and patterned probably with a thick film process. On each side of the hybrid, an on–board Read Out Controller (ROC) receives the clock, trigger and control signals and transmits them to the front–end chips. The ROC also collects, formats and transfers to an external board (Opto-electronic Board) the data from the chips. In case one ROC fails, data can be routed through the second, located on the other side of the module. The Opto-electronic Board holds the drivers for the optical links and, profitting from the modularity of the optical cables, can serve up to two modules.

The use of fan–ins to interconnect the sensors and the electronics permits the chips and strips to be on a different pitch. The space available between chips can be used for bonding and biasing of the electronics and for the insertion of thermal fingers to improve heat extraction from the detectors. Two types of fan–ins are currently being studied, either in ceramic (e.g. beryllia) or in metallised kapton. Beryllia fan–ins are very effective for cooling the detector whilst providing additional mechanical support. They do, however, add to the overall material budget.

One major concern in the design of the module is the simplicity of construction and stability of the system. Both can be enhanced by limiting the number of components used. Table 19 lists the components of the Z-module.

Table 19: Components of the Z module.

component	number
hybrids	1
front–end chips	8 to 12
ROC chips	2
fan–ins	4
connectors	2
SMD components for biases	to be determined

6.4.3 Prototyping

An extensive prototyping program is underway to test the various options considered for the cooling and the fan–ins. Fully instrumented prototype modules are being developed with the aim of operating at LHC speed at the H8 test beam in 1995. These prototypes will also provide a testing ground for the evaluation of the different front–end electronics concepts being considered.

Within the RD2 collaboration prototype modules are being produced that hold 4 silicon detectors with 256 strips and 8 signal processing (ADAM) chips of 32 channels each. The electronics will therefore equip only one half of the strips of each detector. A telescope of three prototype modules will be installed on the H8 beam line.

6.4.4 Material budget

A critical aspect of the module design, and a basis for comparison, is the material budget. An estimate of this is given in Table 20 for the Z module.

6.4.5 Cost

A further aspect of the module design is the cost, an estimate of which, is given in Table 21.

6.5 The R/ϕ module

6.5.1 Module concept

The Rϕ module concept is shown in Figure 74. It is a natural configuration for use with axial or small angle stereo strips, particularly on a cylindrical support structure. This is because the electronics is oriented parallel to the strip direction. The module is designed to: (1) be self supporting, (2) minimise the number of components (material), (3) maximum the signal-to-noise ratio, and (4) have open edges for ease of overlap.

The point of (1) is to decouple the barrel cylinder structure from the modules allowing the module to carry only its own stresses. This is in contrast with the SDC ladder design where the module is a load carrying element[78]. Anchoring the module near the centre eases the thermal stress in the module (compared with support at two ends). Such a design, with a support cylinder, does add extra material but may have advantages in assembly and mounting.

Table 20: Material budget of the Z type module for both stave and cylinder mechanical supports. Units are percent of a radiation length.

item	material	stave	cylinder
detectors	silicon	.60	.60
glue	adhesive	.08	.08
hybrid	alumina	.11	?
electronics	silicon	.05	.05
fan-ins	BeO/kapton,Au,	.08	.043
adhesive		.04	.04
printing	thick film	.06	.06
power cables	aluminium, kapton	.08	.08
stave	Be,water	.33	
cooling pipe	Al,water		.03
general support			.25
fasteners			.02
carbon fibre			.054
heat sink		.02	.02
total		1.43	1.2+...

Table 21: Cost estimate for the Z type module in CHF.

item	CHF
hybrid	370
fan-ins	200
opto-electronics	210
total	780

Points (2), (3), and (4) are realised by placing the front–end electronics parallel to the strip direction and near the middle of the module unit. By eliminating the long fan–in of the Z module, input capacitance is reduced. Reading signals out at the middle of the strips results in the minimum noise and the maximum signal for a given strip length. This is because the resistive input load on the preamplifier is 1/4 of the load with readout at the ends and the signal dispersion through the strips is also reduced. Placing the hybrid at the middle puts the vulnerable wire-bonds to the hybrid away from the edges. Open edges may also simplify the module overlap in both ϕ and z directions.

In the Rϕ design the front–end electronics is closer to silicon and so this design may be inferior to the Z module for reasons of cooling and electronic noise pickup. It should have less material. These are however detailed questions which will be examined in the prototyping program.

6.5.2 Hybrid assembly

The baseline hybrid configuration is composed of a multilayer metallised kapton substrate and a beryllium shield plate (Figure 74). The kapton substrate provides electrical bus lines between the front–end chips, interconnections to the off detector electronics and also distributes bias potential to the detector. The industry standard for this metallisation is copper. Some development work might lead to the use of aluminium instead. The kapton substrate is backed by the beryllium plate, which shields the electrical activity of the front–end electronics from the detector strips and provides a heat path from the chips to the cooling pipes. The cooling pipes contact the end of the plate, with an offset to clear the electrical connections and components. The possibility of imposing a gap in the plate, below the region of the chips is under consideration as a means to reduce heat flow into the silicon. In this case, the plate would be cantilevered, making contact to the silicon only near the cooling pipe with perhaps a small spacer to stabilise during wire-bonding. The cooling pipe contacts only one side of the module but the temperature drop across the

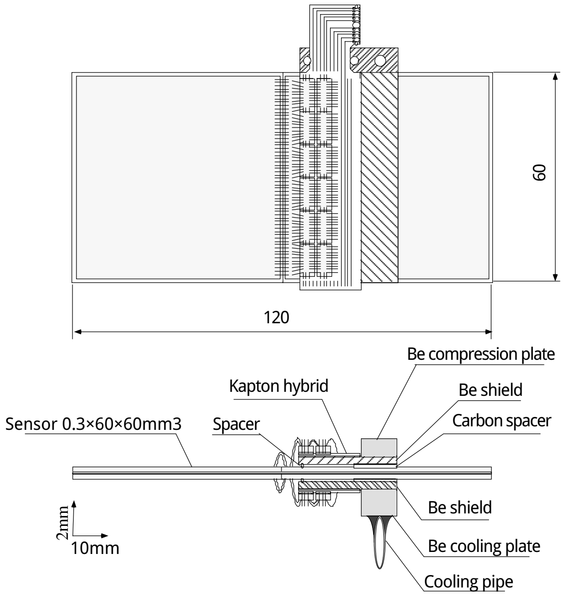


Figure 74: Rϕ module concept. A kapton multilayer substrate and a beryllium plate makes the readout hybrid assembly.

silicon to the the other side is expected to be small, a fraction of a degree C. The overall arrangement with cooling pipe and contact points can also be seen in Figure 75 and Figure 76.

The readout chips are glued either to the kapton substrate, or directly to the metal, face up for wire-bonding. Depending upon the final pitch of the detectors and the choice of readout chips, bonding will either be direct to the detectors or through a small fanout piece, a few millimetres in length. Bonding will be done with aluminium wire, a well established process. It may be required to provide some space between readout chips for bypassing components or power connections.

The assemblies on the top and on the bottom of the detector are made of one piece of kapton substrate/cable, which wraps around the detector and is mounted directly on the detector as shown in Figure 74. If this wrapping is difficult, a fallback would be to make the two assemblies separately and connect them at one side. The kapton substrate will anyway have an extension (pig–tail) on one side to service off detector voltages. These include power supplies for front–end chips and detectors, slow control signals and monitors. In addition a set of fibre drivers and receivers will be placed on the assembly for digital signal transmission.

The beryllium shield on one side of the module is extended out to form a precision alignment "ear", with which the module is positioned to the mechanical support structure. For example, in a scheme which uses a cylindrical support for the modules, the force to hold the module is provided by compression between the cooling line and the compression plate (see Figure 75).

As a fallback solution we are considering also a hybrid assembly made of ceramic, probably beryllia. This would replace the kapton and beryllium components. The possible advantages are in reliability, manufacturing, electrical performance, and cooling. The material burden would probably increase however.

6.5.3 Material budget

Material budget is assessed for the Rϕ module model (Figure 76) and for a model of a support cylinder (Figure 77). The heat exchanger is included in the calculation of the module. The averaged material in the module for normal incidence is 1.01 % of a radiation length (X_o). The averaged material of components is summarised in Table 22. The largest local spot is where the front–end chips are in the material distribution along z (Figure 78). The module is ϕ symmetric in general.

The average material of the support cylinder is 0.51 % X_o. Although the largest local element is the cooling pipe (1.2 % X_o), its average is small (0.1 %X_o) due to its small volume. The CFRP skin and the core are the dominant element on average: both are 0.17 %X_o. The average material of components is also

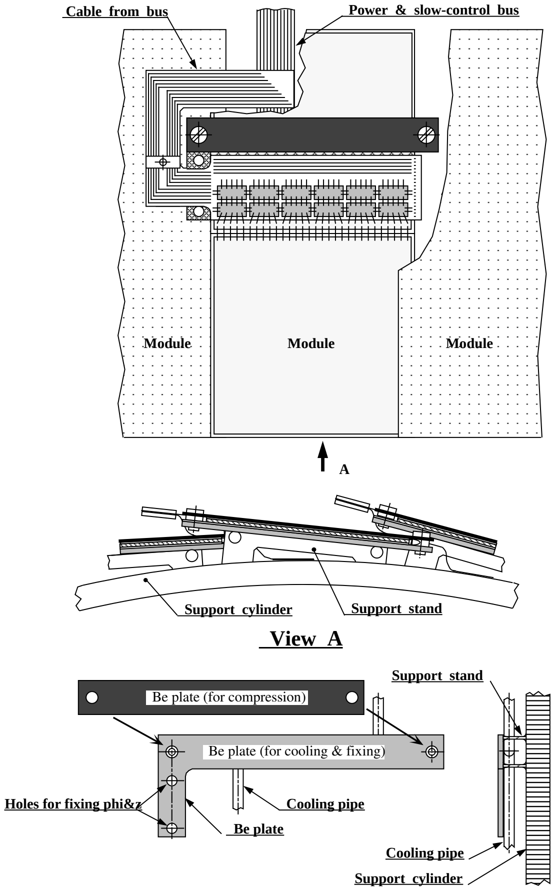


Figure 75: Support mechanics and cable routing assuming a cylindrical support

Table 22: Materials of the Rϕ module and a support cylinder. The scaled thickness (t') is defined by t'= t$\times l/L\times w/W$ where t, l, and w are real thickness, length, and width, and L and W are the module width (60 mm) and length (120 mm).

Component	X_o [mm]	t' [mm]	$< X >$ [% X_o]	Comment
Module			1.01	
Si detector	93.6	0.6	0.641	
Al wire	89	0.0025	0.003	
Epoxy glue	250	0.11	0.045	
Be shield	353	0.29	0.083	
kapton hybrid	284	0.097	0.034	
Si LSI	93.6	0.098	0.104	500μm$\times$13mm$\times$54mm
Be cooling	353	0.325	0.092	
Be pipe	353	0.033	0.009	
Cylinder			0.51	
Al pipe	89	0.044	0.050	
Coolant	361	0.250	0.069	
kapton cable	284	0.042	0.015	
CFRP stand	249	0.067	0.027	
CFRP skin	249	0.440	0.177	
CFRP core	249	0.430	0.173	

summarised in Table 22.

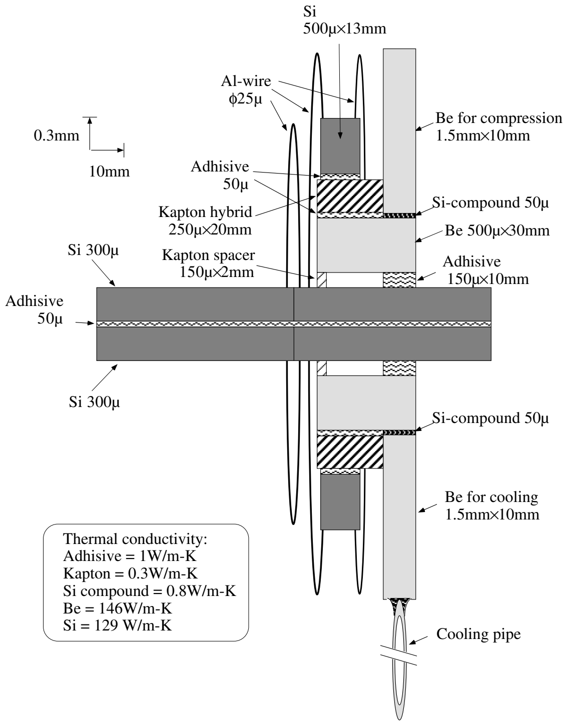


Figure 76: Material distribution model of the Rϕ module.

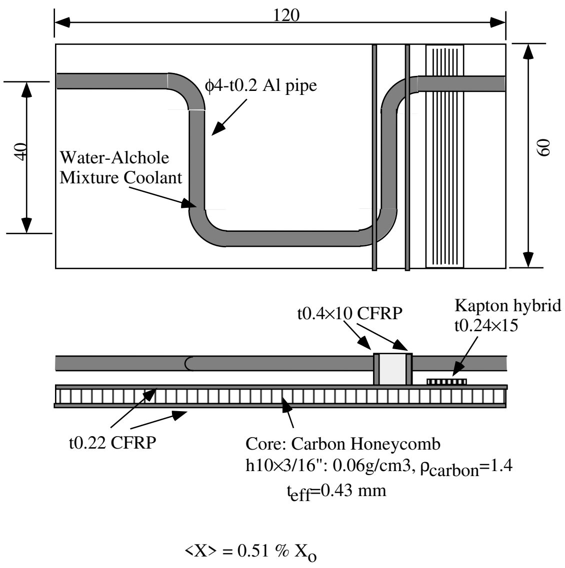


Figure 77: Material distribution model of support mechanics.

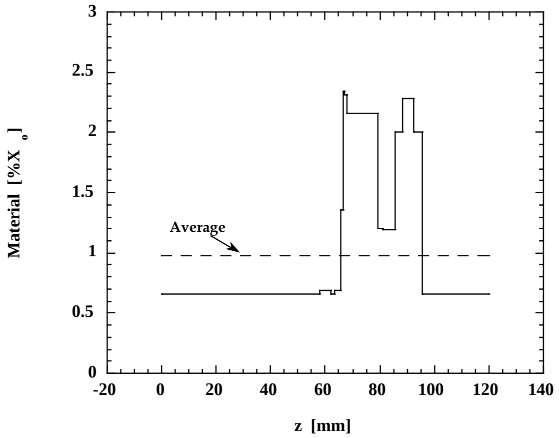


Figure 78: Material distribution along z in the module.

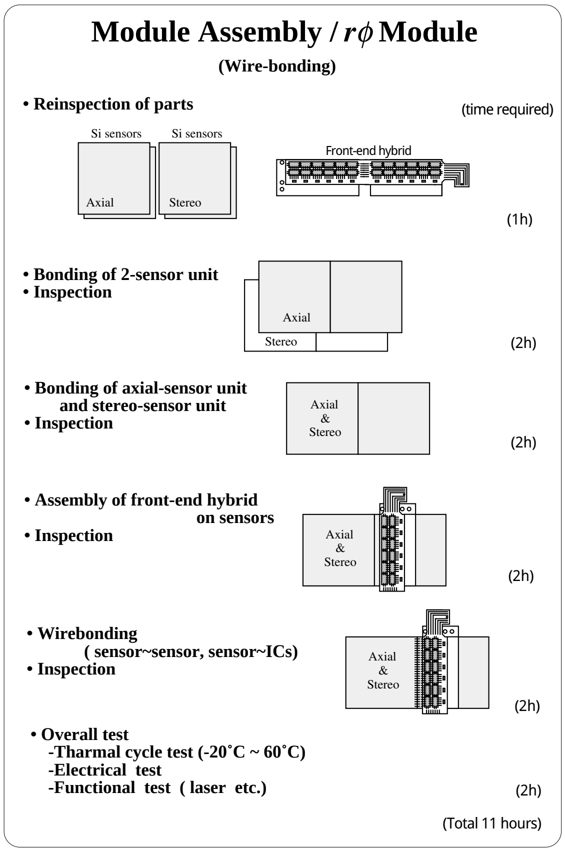


Figure 79: Module assembly sequence.

6.6 Assembly

In this section we shall discuss assembly procedures for silicon barrel modules. Both Z and Rϕ modules will be discussed. Many of the assembly issues are common. The module assembly represents a significant overhead in the construction of the SCT barrel. We envisage multiple sites each constructing a subset of the modules such that the total construction can be completed in about three years. Hybrid assembly and component testing will probably be centralised.

The module assembly can be divided into seven main steps. The first is assembly and test of the hybrids and associated chips. The second is the re–inspection and test of silicon detectors. The third is the gluing together of 6 cm detectors to form the 12 cm unit. The forth is the gluing of hybrid assemblies and fanouts to the silicon. The fifth is the wire-bonding of strips to front–end chips. The sixth is a mechanical survey. Finally, the seventh is an electrical test of the completed module and an extended burn–in cycle. Figure 79 is an example of these steps, with rough time estimates, for a version of the Rϕ module. The module assembly process is discussed in more detail below.

Step 1: Hybrid assembly Hybrid substrates will be either photo–etched or deposited by a thick film process. Hybrid assembly consists of mounting components, cables, and connectors, die attach and wire-bond, test and rework, and die encapsulation. Regardless of the etch or print of the substrates, we estimate 3 hours for the assembly in production. (Required working time: 3 hours)

Step 2: Detector inspection Detectors will be inspected by the manufacturer for a range of characteristics such as depletion voltage, leakage currents, noisy strips, and shorted coupling capacitors. The full set will be specified in the contract. We expect to check a subset of the detectors in detail as a confirmation and perform a reduced set of tests on each device. This will involve a multi–pin probe card and an automated probe station. (Required working time: 2 hours)

Step 3: Glue detectors together This is a key step because it establishes the basic mechanical precision of the module. We consider here the case of back to back single sided detectors since that is more complex than the double sided case. A major issue to be resolved in the prototyping program is whether optical inspection will be required during this step or if crystals can be cut with sufficient precision to use instead mechanical jigs and pins. More specifically we consider the following options.

A) Alignment with dowel pins: Detector crystals will be diced by the manufacturer. We would require precision on the distance and parallelism of the cut edge to the first sensor strip. Evaluations on samples from one vendor indicate this may be possible. It remains to be seen whether this precision can be maintained in production and obtained from other vendors as well. Given a precise crystal, all the alignment could be done off appropriate mechanical jigs. This would greatly simplify, accelerate and cheapen this step of the assembly process.

B) Manual optical alignment: Detector crystals are first positioned (by dowel pins and held in place with vacuum) on jigs as in A). Crystals could be shifted and rotated with micrometer driven stages by an operator. This work would be done on a coordinate measuring system (CMM) using magnified optics to measure the position of fiducials. Data could be logged to the CMM computer. This method is independent of detector cutting precision. This method has been used successfully in the construction of most of the existing silicon vertex detectors [79]. Measurement accuracy is about ± 5 μm.

C) Automatic optical alignment: This process is similar to B) but would occur within a robotic assembly work-cell [80]. This work-cell contains optics with digital imaging and pattern recognition software. The system could be programmed to find and log the crystal fiducials. Positioning micrometers would be actuated by drive signals from the work-cell processor. Survey data would be automatically logged. These systems may also be able to dispense and cure adhesives.

The first step in the detector assembly is the gluing of a butt joint between the two sets of crystals. This will probably be done rapidly with a UV cure adhesive. Following this, a thermally and electrically conductive adhesive has to be spread on the back of one set of crystals. This layer has to be uniform to ensure good heat transfer between the detectors. This layer will be applied either by a stencil/screen technique or with an adhesive film. Once the two sides are pressed together and aligned we will lock them with a few spots of UV cure adhesive and then cure the whole assembly in an oven at an elevated temperature. Some pipelining is envisaged in this step. (Required working time: 4 hours)

Step 4: Glue hybrids to detectors Gluing of hybrids to detectors is also a key step because the hybrid is the mechanical interface to the overall support structure. The same alignment issues discussed in Step 3 apply here as well. This step requires gluing and alignment on both sides with particular care given to maintaining good thermal contact across bond lines. Again, UV cure materials will probably be used to lock positions before oven cure of the main bonds. Some pipelining is envisaged in this step. (Required working time: 2 hours)

Step 5: Wire-bonding This step will be done on an automatic bonding machine. It is routine but can be a bottleneck in the production since the machines are expensive. (Required working time: 1 hour)

Step 6: Mechanical survey The complexity of this step may depend upon the mechanical alignment used in the detector and hybrid gluing operations. (Required working time: 1 hour)

Step 7: Electrical test and burn–in Electrical tests can include measurements of digital function, noise and charge injection response. For systems with sparse readout we would also measure thresholds. A laser or source scan may also be specified. A burn-in would typically be 100 hours but would be done on a set of modules in parallel. (Required working time: 2 hours)

The total assembly time for a module is estimated at 12 hours, exclusive of hybrid assembly. We expect to improve these estimates during the module prototyping program.

6.7 Facilities and Labour

For both the Z and Rϕ modules, the assembly sub-processes will be done in a clean environment. For the re–inspection of detectors, we will need an automated probe station to check the sensors. Other stations and microscopes are required for checking and repairing the hybrids. For the optical gluing processes, we will need either a coordinate measuring machine or a robotic work-cell with appropriate manipulators and fixtures. An adhesive dispensing or screening system will be used to apply the thin layers to silicon and hybrids. Several ovens will be needed for curing adhesive. (Required man-power: 2 technicians and 1 engineer/supervisor)

An automatic bonding machine is essential for the wire-bonding. A skilled operator is also necessary. For the final check–out, we will need a chamber for thermal cycling. We will also need a laser scan system, off–detector electronics and a data acquisition system for the static and dynamic performance checks of the module. (Required manpower: 1 wire-bonding expert, 2 technicians, 1 engineer/supervisor)

Although the assembling time of one module is 12 hours some pipelining is possible. We estimate an efficient facility can produce 3 modules per day. With one facility, it would take 1000 days to produce the total of 3000 modules; with a 100% contingency this would be 2000 days. If a year is counted as 200 days, it would take 10 years; we will need several module-assembling facilities to finish the production in less than 3 years.

6.8 Assessments: safety and failure modes

6.8.1 Assessment

CERN and local laboratory rules of safety will be strictly observed at all stages of construction and tests, in particular concerning beryllium and beryllia components.

6.8.2 Failure modes

The failure modes of the SCT module operation can be separated into the following categories listed in order of decreasing severity [81]:

A) Catastrophic hard failures: Large parts of the SCT, a whole layer or several layers, become inoperational. Examples are loss of cooling, loss of power to a detector segment, loss of a cable or fibre connection, and severe beam loss into the detector.

B) Severe hard failures: Smaller parts, of the order of a chip or detector, become inoperative. Examples are a capacitor short on one detector and a short of low voltage power on a chip.

C) Severe soft failure: The design parameters of a large part of the SCT are compromised. An example is the cooling system running at higher temperature resulting in increased noise and a potentially compromised alignment. If the alignment stability is worsened, correlated errors may increase.

6.9 Finite element simulations

As discussed, due to self heating effects in the detectors, the potential for thermal runaway exists. To study these effects a number of calculations and simulations have been run. In this section we will review these.

Dortmund/Oxford study:

A finite element program was written by A. Lahmamm and A. Reichold (A^2) to simulate the dynamic evolution of temperature profiles in two dimensional models detector models. A simplified two dimensional model of an Rϕ and "inline" type module was used. (The inline module has the electronics off the silicon, with a ceramic bridge connecting the hybrid to the detectors. The cooling is attached to the bridge) The model consists of a monolithic block of silicon of 120 by 60 by 0.3 mm^3 which is in contact with a 1 mm^2 pipe at 0°C, through a 50 μm layer of thermally conductive grease. Front–end chips were simulated on top of the silicon by adding the power dissipated in the chip to the corresponding finite elements of the silicon detector. The temperature distributions were calculated every 0.001 seconds of simulated time and could be monitored at arbitrary intervals. The effects of a gas envelope, with and without gas exchange at various velocities and inlet temperatures were also simulated. The main goal of the program was a comparison between results obtained with the full temperature dependance of the silicon power dissipation and those with a constant power density. The main results of the calculations are described here:

- A calculation assuming constant power density in the silicon detectors may be unstable under the assumption of temperature dependant power dissipation.
- The model showed that with the expected convection[1] and without any additional gas cooling a single cooling pipe across the width of the detector is not sufficient to stop thermal runaway.

Several ways to prevent thermal runaway were simulated.

- Increased convection was used to enhance the thermal conductivity of the silicon. It was found that at $\alpha = 100\frac{W}{K \cdot m^2}$ the calculations were stable but still showed an unacceptable temperature profile with temperature differences on the silicon of up to 10°C.
- Increasing the thermal conductivity of the silicon by 400% locally in an area of 20 by 120 mm^2 simulated the effect of a beryllia thermal fan in. This improved the previously simulated profile to give differences of only 7°C.
- Leaving the convection coefficient at $\alpha = 15\frac{W}{K \cdot m^2}$ and adding a second cooling pipe running along the length of 120 mm in the centre of the detector brought the temperature differences down to 4°C.
- Simulations with gas flow around the detector scanned flow velocities from 0.001 m/s to 10 m/s and inlet temperatures from -10°C to 10°C. The results are summarised in Table 23. Three conclusions can be drawn from this:
  1. Even though a sufficiently cold gas flow can stabilise the silicon temperature profile of one module and remove the temperature differences originating from the silicon self heating, a single air flow at reasonable flow velocities can not be used to cool the complete length of the detector (i.e 14 modules) without introducing temperature differences of approximately 10°C.
  2. The gas flow does not need to take away a large fraction of the power to stabilise the temperature profile. Approximately 0.25 Watts per module are sufficient to obtain an acceptable temperature profile. This is due to the fact that the cooled silicon has lower power dissipation and consequently is at lower temperature. Due to this lower temperature it will deposit less heat in the gas flow.
  3. If the gas flow is too cold, it will introduce large temperature differences between the local hot spots from the front–end electronics and the silicon detectors which are nearly at the same temperature as the air flowing past.
  4. The cooling system should be capable of compensating large variations in the silicon power density without large temperature differences building up on the detectors.

RAL study:

Studies of thermal performance of the Z module have been made at RAL using the ANSYS code. The results will be summarised in a the next version of this document.

Hytek/LBL study:

William Miller, of Hytek Inc. in Los Alamos, New Mexico, made a set of studies [82] of the KEK Rϕ and of the Z type modules. In these studies he used both finite element techniques (COSMO code) and approximate solutions to the differential equation relating silicon temperature to surface heat flux (W/mm^2) per single sided detector, to understand the parameters governing thermal runaway. (For a highly irradiated detector in ATLAS, we assume a surface heat flux of 4.2×10^{-5} W/mm^2 per single sided detector.)

A first set of calculations were applied to a module with only silicon self heating. The front–end chips were not included. For the KEK design a 0°C boundary was imposed at 9 cm from one end. For the

[1] A convection coefficient of $\alpha = 15\frac{W}{K \cdot m^2}$ was assumed

T^{in}_{gas}[°C]	T^{out}_{gas}[°C]	V_{gas} [$\frac{m}{s}$]	P_{gas} [W]	T^{min}_{Si} [°C]	T^{max}_{Si} [°C]	ΔT_{Si} [°C]	time [s]
5	≈ 5	1.0	≈ 0.13	1.186	8.478	7.292	93.3
5	≈ 5	2.0	≈ 0.13	1.185	8.461	7.276	92.4
5	≈ 5	3.0	≈ 0.13	1.185	8.456	7.271	92.1
5	≈ 5	4.0	≈ 0.13	1.184	8.452	7.268	92.0
5	≈ 5	0.5	≈ 0.13	1.188	8.509	7.321	94.9
5	≈ 5	0.2	≈ 0.13	1.194	8.605	7.411	99.9
5	≈ 5	0.1	≈ 0.13	1.203	8.759	7.556	108.0
0	0.635	0.1	0.423	0.879	3.703	2.824	79.2
5	5.002	10.0	0.133	1.184	8.449	7.265	91.7
0	0.0056	10.0	0.373	0.837	3.654	2.817	67.1
5	6.797	0.01	0.12	1.321	10.797	9.476	227.4
0	3.673	0.01	0.245	1.088	6.899	5.811	155.1
-	≈ 24	0.0	-	4.037	>53	-	94.4
5	20.42	0.001	0.103	2.29	>100	>97	316.9
0	20.115	0.001	0.167	2.27	>100	>97	418.5
0	0.0678	1.0	0.452	0.842	3.658	2.817	68.2
-5	-3.883	0.1	0.745	-0.531	3.356	3.888	40.5
-10	-8.361	0.1	1.092	-4.44	3.024	7.463	71.5
10	12.135	0.1	-	1.811	>230	-	392.0
0	0.016	4.0	0.427	0.838	3.655	2.817	67.3

Table 23: Summary of simulations with varying air velocity and temperature. Lines marked with - in the column 'stable' are divergent results due to thermal runaway. The column 'time' shows the simulated time until equilibrium or divergence was reached. For the first seven entries in the table the temperatures of the outgoing gas were not stored, but are believed to be very similar to that of line 9

Z module, the corresponding distance was approximately 6 cm. Miller observed that the length from the cooling boundary was the critical design parameter. Runaway was seen to occur at a fixed temperature difference across the silicon of 12.9°C with cooling boundary at 0°C. However the critical heat flux needed to reach this condition was seen to vary inversely with the silicon path length squared. While these results came from the finite element study, approximate solutions to the differential equations agreed within 5%. For a 9 cm wafer the critical flux was 4.91×10^{-5} W/mm^2, while for the 6 cm wafer it was 12.05×10^{-5} W/mm^2, per 300 μm thick detector.

A second set of calculations included the additional load of the front–end electronics (5 W) including conductivity through the beryllium shield plate. The critical heat flux was reduced to about 15%, corresponding to about a 2°C shift in silicon temperature for fixed surface flux. The performance was improved by the addition of a 1 mm gap in the beryllium plate below the region of the front–end chips. A larger improvement was found with the inclusion of a high thermal conductivity heat spreader between the back to back crystals. A sheet of CVD diamond was used, 12 mm wide and 300 μm thick. This material has a thermal conductivity in excess of 1300 W/m°K. In this case the heat flux range was extended by a large factor, suppressing runaway for any likely heat flux in these modules.

A third set of calculations looked at the Z module geometry. Here the effect of front–end electronics is lessened at low flux since they are located off the silicon. Some concern was raised about the efficiency of the heat paths between cooling stave, beryllia spreader, and hybrid in this design. Runaway was observed to occur around 7×10^{-5} W/mm^2 which should be safe.

The main conclusions of this study are as follows:

- At fixed coolant temperature, runaway occurs for a fixed temperature difference across the silicon.
- For the KEK Rϕ concept the margin for runaway may be small at high dose unless the thermal conductivity is enhanced.
- The basic Z module geometry is superior due to shorter cooling paths.
- Front–end chip power raises the overall module temperature and lowers the runaway surface flux value if not removed by a proper path to the cooling line.

KEK study:

Taka Kondo, of KEK, performed time dependant finite element studies of thermal runaway in an Rϕ type module using the ANSYS code. Based upon these studies, he was able to derive sets of conditions for which runaway will occur. The results are summarised in Table 24. The model is based upon Figure 80 wherein the various quantities in the table are defined.

The calculation assumes the temperature dependent leakage current,

$$I(T) = const * T^2 exp(-E_g/(2kBT))$$

where, $E_g = 1.23$ eV, $kB = 8.615 \times 10^{-5}$ eV/K, and T is in Kelvin.

The conditions for the calculations are:

1. The geometry is the same as the KEK test module (see Figure 80). Two single sided silicon crystals (60 mm x 120 mm) are glued together. The cooling channel is on one side. The "leg" is 2 mm wide and 0.20 mm high.
2. Amplifier power dissipation is 5 W per module.
3. The convection coefficient to air is 12 $Watt/m^2/K$. The temperature of ambient air is fixed to T(air) as a input parameter.
4. The temperature constraint is at the top of the beryllium tower (one side only). The constraint temperature is shown as T11 in the table.
5. The heat generation in silicon Q(Si)0 is varied as shown in the table. Q(Si)0 means the total power dissipation in the volume of silicon wafer (2 x 60 mm x 120 mm x 0.3 mm) at the temperature of 0° C. Its temperature dependence is the same as that of leakage current given by the formula above.

T1, T3, and T6 are the temperature at points #1, #3 and #6 of the KEK test module on silicon. The temperature T11=T(air) for all the cases. T11 (cooling temperature at the Be plate) is set for three cases to, -10, 0 and 10 °C.

6.10 Thermal measurements

In this section we will discuss measurements of both temperature distributions in prototype modules as well as thermally induced mechanical effects.

6.10.1 Studies of the Z module

This text is in progress and will be included in a later version of this document.

6.10.2 Studies of the Rϕ module

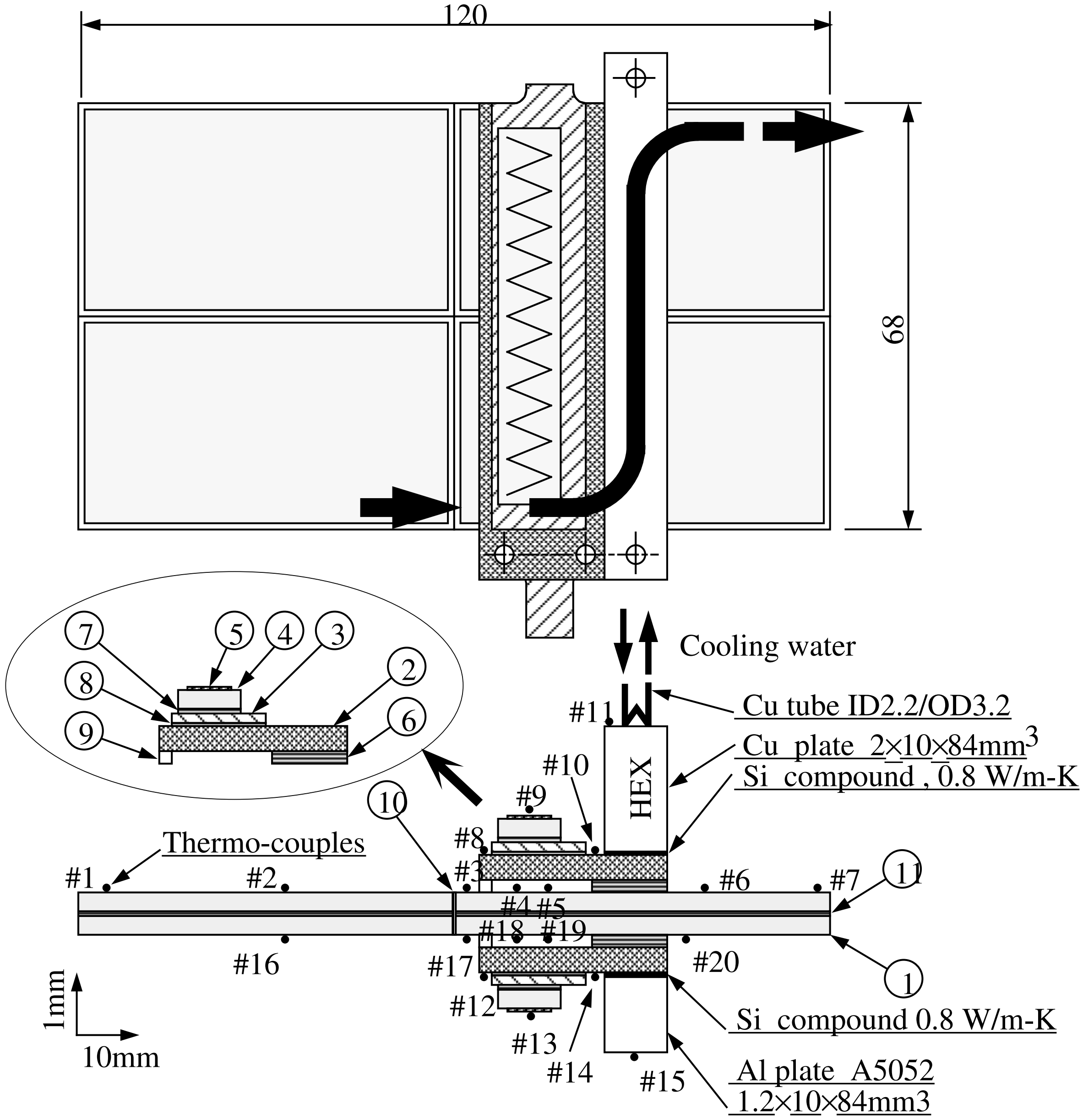
The temperature distribution within one module has been measured at KEK with a mechanical prototype. In this study, the electronics heat (5 Watts) and the detector heat (1 Watt) is simulated with thin wire heaters (see Figure 80). The module was set in a room temperature ambient and cooled by water near room temperature. A number of thermocouples measured the temperature distribution of the detector and of the hybrid. The module was set in an open space vertically and horizontally, with a plate over the module. The measurements are relevant to understanding only the relative temperature differences across the module.

The results are shown in Figure 81. They indicate that the heat leaking from the front–end electronics to the detector is largest at the point of the kapton spacer at the edge of the hybrid; underneath the electronics the temperature is lower than at the point of the kapton spacer. Solid contact through kapton is worse than air. The maximum temperature difference in the detector is 3∼5 °C depending on the relative temperature of the air and the module. The temperature difference is well within the temperature difference to cause a depletion voltage difference of 10 V over the 10 years of operation. A finite element calculation reproduces the measurement [83].

The measurements done to date do not simulate the thermal runaway of the silicon detector. This can be done by subdividing the wire heaters and using the measured temperature to regulate the voltage. This is planned for future efforts.

Table 24: Results of KEK simulation of thermal effects due to self heating of modules

run	T11	T(air)	convec	Q(Si)0	T1	T3	T6	ΔT max
1	0.0	0.0	no	0	6.7	6.7	2.2	4.5
2	"	"	no	0.1	7.7	7.0	2.3	5.4
3	"	"	no	0.2	8.9	7.4	2.4	6.5
4	"	"	no	0.3	10.8	8.1	2.5	8.3
5	"	"	no	0.4	13.0	8.8	2.7	10.3
6	"	"	no	0.5		runaway		
7	-10.0	-10.0	no	0.0	-3.3	-3.3	-7.8	4.5
8	"	"	no	0.5	-1.4	-2.7	-7.6	6.2
9	"	"	no	1	2.3	-1.5	-7.4	9.7
10	"	"	no	1.2	4.4	-0.8	-7.2	11.6
11	"	"	no	1.3	6.4	-0.2	-7.1	13.5
12	"	"	no	1.35		runaway		
13	10.0	10.0	no	0.0	16.7	16.7	12.2	4.5
14	"	"	no	0.1	19.9	17.8	12.6	7.3
15	"	"	no	0.2	27.3	20.1	13.0	14.3
16	"	"	no	0.25		runaway		
17	0.0	0.0	yes	0.0	3.0	5.2	1.7	3.5
18	"	"	yes	0.5	5.1	6.0	2.1	3.9
19	"	"	yes	1.0	9.4	7.6	2.6	6.8
20	"	"	yes	1.2	13.8	9.0	2.9	10.9
21	"	"	yes	1.25		runaway		
22	-10.0	-10.0	yes	0.0	-7.0	-4.8	-8.3	3.5
23	"	"	yes	0.5	-6.3	-4.5	-8.2	3.7
24	"	"	yes	1.0	-5.6	-4.2	-8.0	3.8
25	"	"	yes	1.5	-4.8	-3.9	-7.9	4.0
26	"	"	yes	2.0	-3.8	-3.5	-7.8	4.3
27	"	"	yes	2.5	-2.3	-3.0	-7.6	5.3
28	"	"	yes	2.75	-0.9	-2.5	-7.5	6.6
29	"	"	yes	2.8		runaway		
30	10.0	10.0	yes	0.0	13.0	15.2	11.7	3.5
31	"	"	yes	0.4	20.3	17.9	12.7	7.6
32	"	"	yes	0.5		runaway		



1: Si 0.3×34×60mm3 , 129 W/m-K
2: Al plate A5052 , 0.4×30×76mm3 , 121 W/m-K
3: Kapton 0.15×15×68mm3 , 0.3 W/m-K
4: Si 0.3×10×60mm3
5: Heter
6: Nylon mesh + Adhesive 0.2×12×68mm3 , 1.42 W/m-K
7: Adhesive 0.075×10×60mm3 , 3 W/m-K
8: Adhesive 0.05×15×68mm3 , 1.42 W/m-K
9: Kapton + Adhesive 0.2×2×68mm3 , 0.3 W/m-K
10: Adhesive 0.05×0.3×68mm3 , 0.84 W/m-K
11: Adhesive 0.075×68×120mm3 , 1.5 W/m-K

Figure 80: Cooling test of the KEK mechanical module.

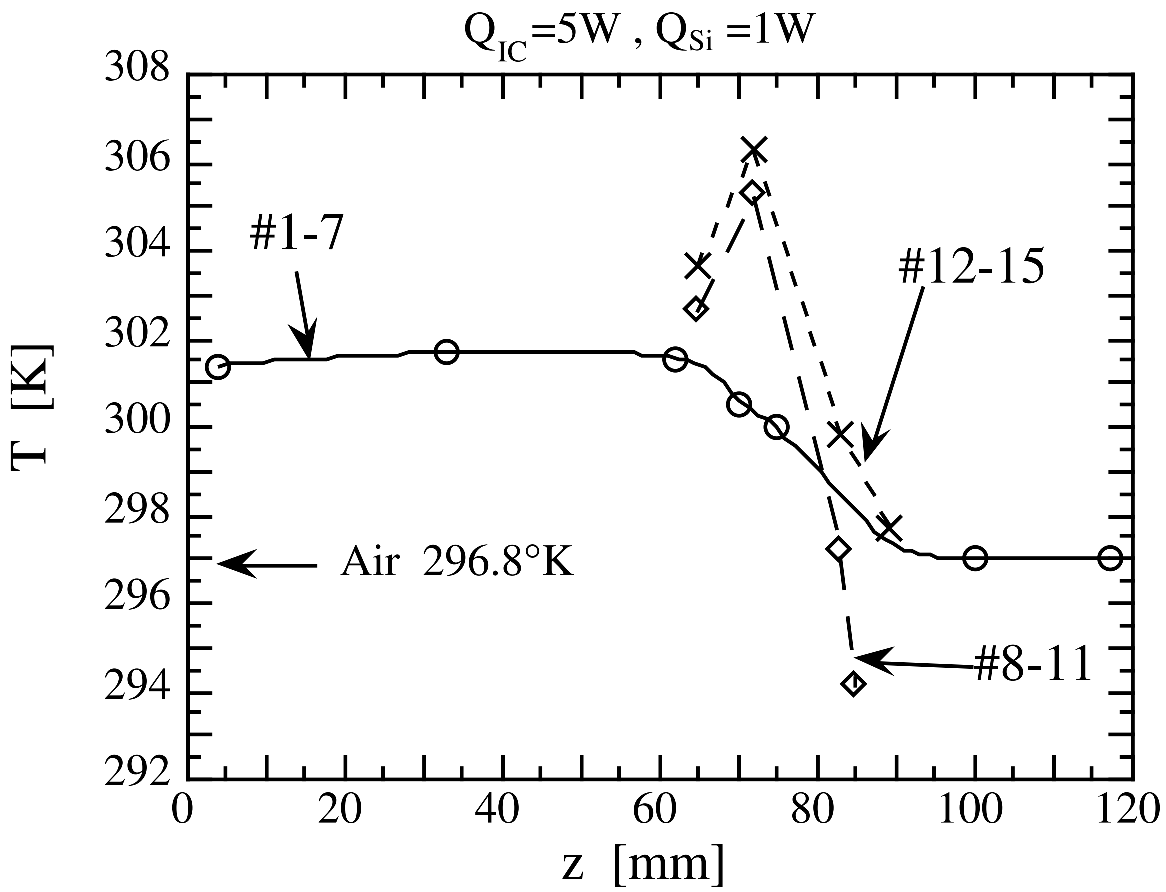


Figure 81: Temperature distribution along z of the module. The cooling water temperature determined the relative temperature difference between the module and the air.

7 Mechanical Engineering

In this section we discuss mechanical engineering aspects of the detector support, cooling, and construction. Some mechanical engineering issues were dealt with in Section 6 but those refer to specifics of module design. The Inner Tracker layout is shown in Figure 13. The Inner Tracker occupies a volume 6.9 metres long by 2.3 metres diameter and has a mass of 1500 kilograms. It is composed of three sub-assemblies, a central barrel region which has most detectors arranged in concentric cylindrical arrays, and two forward regions with radial arrays. Stability of the low mass silicon structure to the 10 μm level is critical. Instability may result from temperature changes, humidity, and micro-creep effects.

7.1 Overview

7.1.1 Technical Description: Barrel

The central barrel region is composed of the TRT barrel, four silicon detector barrels, two pixel barrels, and four pixel discs. It is the four silicon detector barrels which will be discussed in this section. The barrel layout is given in Table 25.

Table 25: Numbers of detector modules per barrel.

Name	Radius in mm	Detectors in Phi	Number of modules	Mass kg
Barrel 1	300	34	476	16
Barrel 2	400	46	644	20
Barrel 3	500	56	784	26
Barrel 4	600	68	952	31

The mechanical engineering requirements for the silicon detector barrel are:

- a) support of 41 m^2 of pixel and silicon detectors, stable at the ± 10 μmlevel in rϕ, at constant temperature.
- b) removal of heat dissipated by the front end electronics (20 KWatt). This heat load will increase over time as the silicon detectors degrade with radiation damage.
- c) operation at a stable temperature of 0° C.
- d) ensure that services and beam pipe supports have no strain effects on the structure positioning and stability.
- e) minimisation of dead space between the barrel and forward regions.

The Inner Tracker project is at an early stage regarding detailed structural design and materials choices. At present two basic support structures are under consideration for the barrel. We envision choosing a baseline structure by the end of 1995. In this document we will describe both schemes in some detail. One structure is based upon carbon composite cylinders (the Cylindrical Design). The other is based upon longitudinal beryllium supports containing integrated cooling channels tied together with metal matrix rings (the Stave Design). Each scheme has certain advantages in performance, cost, assembly etc. Also, with a particular choice made, certain silicon strip module designs are more or less natural. A basic layout for one radial layer is shown in Figure 82 and Figure 83 for each design. An end view of the SCT barrel for a stave design is also shown in Figure 84

7.1.2 Technical Description: Forward

Currently the only element of the SCT which is part of the forward system are the GaAs wheels. These will be supported off the MSGC wheels. The mechanical discussion is deferred to that backup document for the present. Some discussions are in progress concerning a forward silicon backup. Some ideas relevant to that system might be included here in a future version of this document.

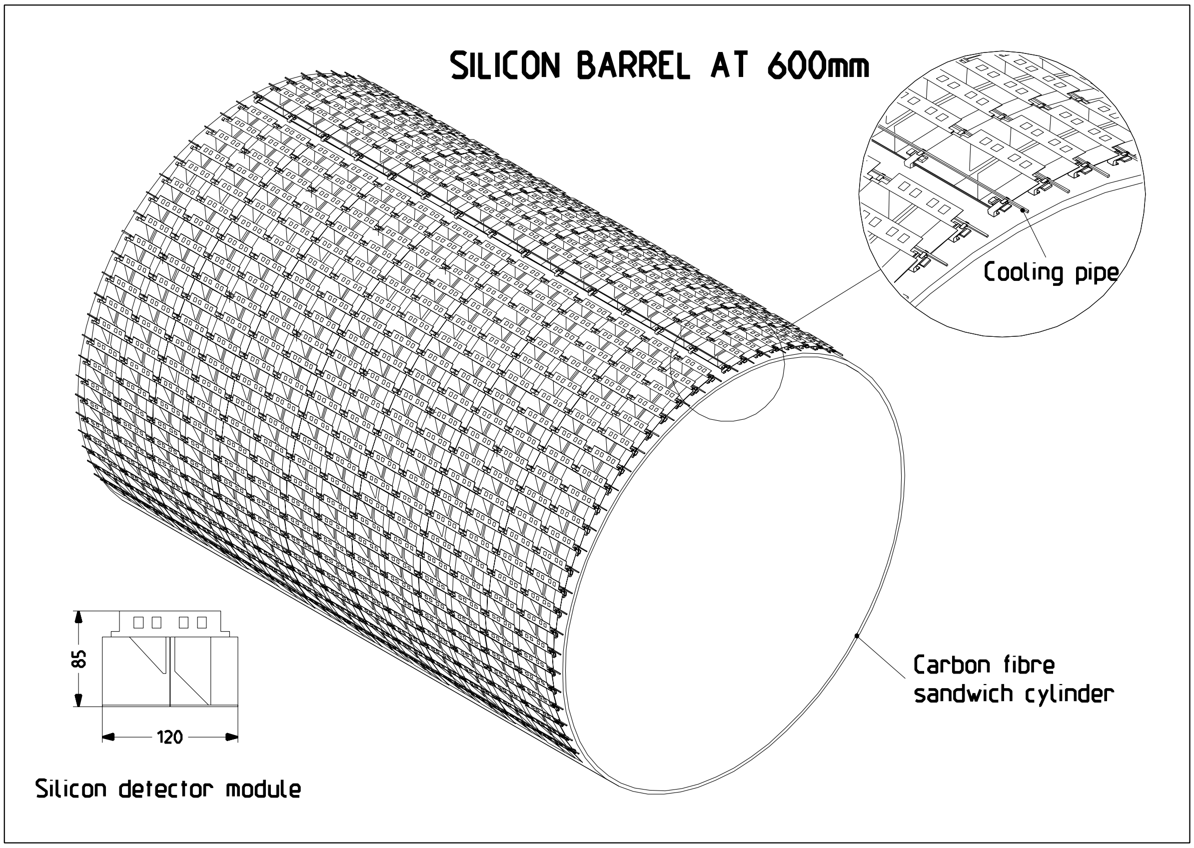


Figure 82: The 600 mm radial layer of the Cylindrical Design.

7.2 Materials

A survey has been carried out of lightweight materials with high specific stiffness. Among these are materials which are, or were proposed for use in other tracking systems for high energy physics [84]. These include boron, fibre-glass, and metallic foams, Rohacell, Nomex, whisker aluminum, aluminum alloys and beryllium. After study, the materials available for use in the support structure have been optimised to carbon-fibres, Nomex, metal matrix composite materials, and beryllium. Table 26 gives typical values for materials used in lightweight construction, and Figures 85 and 86 show elastic modules versus coefficient of thermal expansion (CTE) and radiation length for some of them.

As a pure single component metal, the properties of beryllium are rather simple to understand and well documented. The material is widely used where a high strength to weight ratio is required. Applications include guidance system components for missiles and optical platforms for spacecraft. There are a number of issues of concern for the tracker application. The primary issue is the large coefficient of thermal expansion. A beryllium structure will change considerably between room and operating temperatures in ATLAS. A design incorporating this material will have to account for and monitor these changes. Other concerns come from cost and health and safety. The impact of cost will emerge fully over the upcoming prototyping process. Though beryllium oxide dust can be toxic we expect that procedures are possible to allow the material to be used safely in ATLAS. Particularly strict requirements will be needed on engineering design to avoid the need for any modifications to the structure during assembly.

Carbon fibre composites are attractive examples of advanced materials for the tracker application. Ultra high modulus (UHM) carbon fibres with appropriate epoxy or cyanate ester matrix are employed for stable structure applications in spacecraft manufacture [85]. In those applications, the temperature changes are generally larger than in high energy physics applications. Another positive feature of carbon fibre for stable structure applications is that its CTE can be tuned to essentially zero (in the laminate plane), or even to

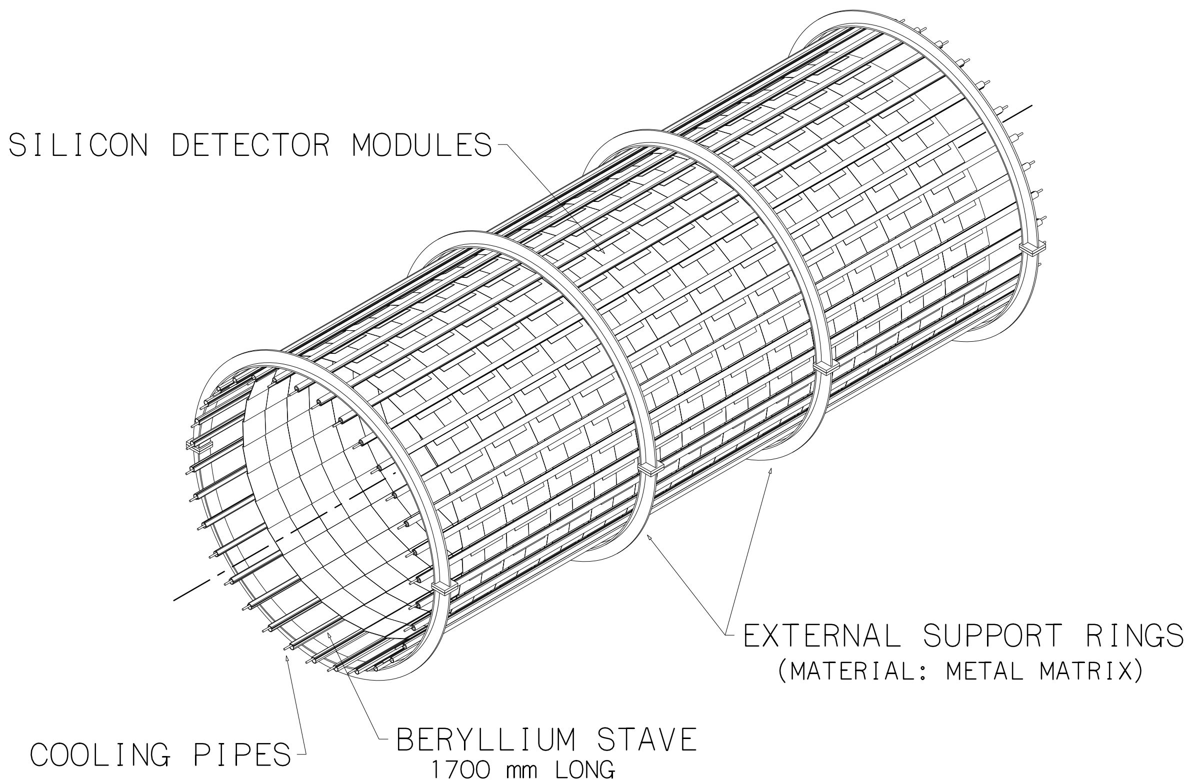


Figure 83: One radial layer of the Stave Design.

match the CTE of silicon. In the latter case there is considerable loss in Young's modulus because high modulus carbon fibre cannot be used. Refer to Figure 85 showing the Elastic modulus–CTE dependence for several quasi–isotropic carbon fibre laminates [86].

Even considering manufacturing costs, a carbon fibre design is thought to be less expensive than a beryllium design.

There is a concern regarding humidity and other factors affecting the stability of carbon fibre components [87]. Here, the use of low moisture intake resins, for example Cyanate ester or Polyimide will make the laminate considerably less prone to swelling. Expected maximum shrinkage of an RS-31XN50A laminate submitted to an environmental change from 55% RH to 0% RH (dry nitrogen) may be evaluated [87]. The coefficient of moisture expansion (CME) of this laminate in this range of RH is 170 ppm/% and the maximum moisture content of 0.18% gives a shrinkage of 30 ppm or 30 μm per metre. The use of a moisture barrier coating is another method applied to reduce humidity effects.

7.3 Material Tests

Particularly in the area of carbon fibre composites, a wide range of parameters exist for selecting an optimal material. For the material tests of carbon fibre composites, several sandwich plates, using different skin–core combinations, ranging from general use to sophisticated aerospace formulations, were manufactured. Four point bending tests were carried out to ASTM specification C393 to determine the flexural properties of these test sandwich plates. Comparisons made with calculations were in good accordance.

With the high stability requirements for the tracker structures, dimensional changes induced by moisture absorption or desorption is also a concern as discussed above. Moisture absorption tests have been carried out to ASTM specification C272 on samples cut from the same plates used for bedding tests. The results are presented in Table 27.

These moisture absorption tests combine high temperatures and wet atmospheres and are sufficiently severe to induce moisture saturation in the resin matrix. This will normally never be the case with the real

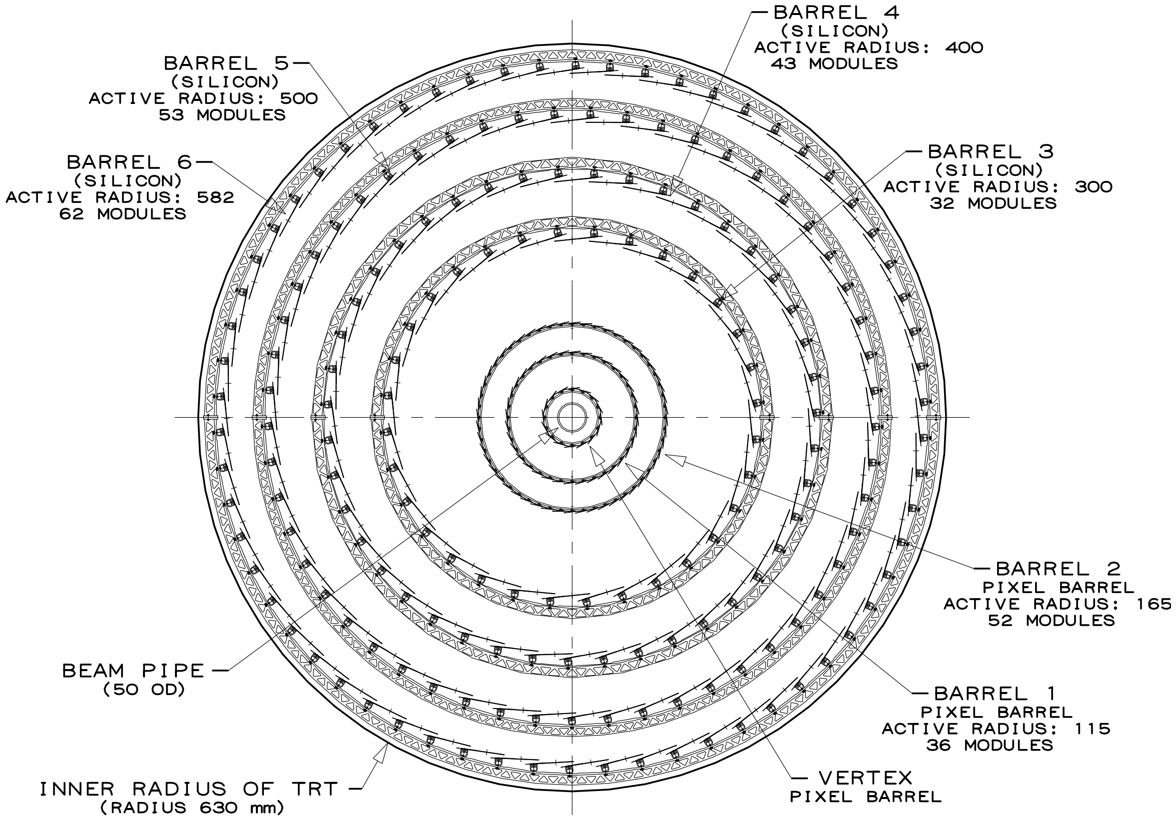


Figure 84: End view of the SCT barrel for the Stave Design.

structure. Movements induced by change from ambient to working conditions are expected to be smaller than the ones observed in these tests. New tests will be carried out in 1995 to measure the dimensional changes of samples submitted to room temperature environmental changes from 50% RH to 0% (dry nitrogen) and vice-versa.

Regarding beryllium structural parts which have cooling pipes integrated, corrosion tests have also been carried out to check any possible effect.

Heat sink compound DC340 samples were exposed to radiation up to fluence of 6×10^{13} neutrons/cm^2 without showing significant change in texture and thermal conductivity.

7.4 Structural Studies

7.4.1 Barrel

In this section we will give engineering details related to the two structural designs under consideration

An engineering drawing of a beryllium stave is shown in Figure 87. The stave both supports the silicon strip modules and incorporates the cooling channel. Consequently good thermal contact is made to the front end electronics and detectors. A stave design works naturally well with the Z module described in Section 6.4. Since each stave is an independent unit, the stave design allows the assembly and testing of modules on staves to occur at multiple establishments. The completed stave assemblies can then be transported to the final barrel assembly area. The dimensional stability of beryllium was discussed in the previous section. The stave will be manufactured by a filled billet extrusion technique [88] followed by final machining at the thermal contact regions. All surfaces other than those used for thermal contact will have anodised or similar coatings. Discussions have been held with industry regarding manufacture including full machining of these staves and vendor quotes have been obtained. It is planned to order a beryllium stave in

	ELASTIC MOD. (N/mm^2) E	DENS. (g/cm^3) ρ	CTE (ppm/K)	THERM. COND. (W/mK)	RAD. LEN. (cm)	REL. SPEC. STIFF. E'/ρ'	$E' = E/E_{Al}$ $\rho' = \rho/\rho_{Al}$ E'^n/ρ'		
							$n = \frac{2}{3}$	$n = \frac{1}{2}$	$n = \frac{1}{3}$
St. Steel	$2.1 \cdot 10^5$	7.85	11	15	1.76	1.02	0.71	0.59	0.49
Titanium alloy	$1.1 \cdot 10^5$	4.5	9.4	7–16	3.56	0.89	0.78	0.73	0.68
Aluminium alloy	$7.1 \cdot 10^4$	2.7	23	100–200	8.9	1.00	1.00	1.00	1.00
Magnesium alloy	$4.5 \cdot 10^4$	1.8	25	60–130	14	0.95	1.11	1.18	1.29
Beryllium	$2.9 \cdot 10^5$	1.85	11.6	145	35.4	6.0	3.75	2.96	2.34
C-fibre/Epoxy/M601	$3.5 \cdot 10^5$	1.7	–1.1		25	7.8	4.60	3.52	2.7
C-fiber/Epoxy/UHM	$9.6 \cdot 10^4$	1.6	0.2		25	2.28	2.06	1.96	1.87
Fibreglass EPGII	$2.1 \cdot 10^4$	1.9	10–15		19.4	0.42	0.63	0.77	0.94
Silicon	$1.3 \cdot 10^5$	2.3	2.6	129	9.37	2.15	1.76	1.59	1.43
Mg-MMC	$1.9 \cdot 10^5$	2.0	–0.3	412	16.8	3.6	2.60	2.21	1.87
Al-MMC	$2.0 \cdot 10^5$	24	0	409	13.2	3.08	2.20	1.86	1.58

Table 26: Typical materials properties and selecting chart for elastostatic stability criteria

Table 27: Typical dimension change due to humidity. The measured changes are over a 70 mm sample.

Core	Skin	μm
Rohacell 31	EP104 2 plies 0/90	25
Nomex 3.2-48	UHM/3501 4 plies	5
Hexcel HFTG	EP104 2 plies 0/90	15
Hexcel HFTG	UHM/3501 4 plies	5
Hexcel HFTG	YLA RS-3 XN-70A	¡4

early 1995 for prototype cooling and alignment/survey studies.

Finite Element Analysis has been carried out on a typical structure. The design aim was keep deflections within 60 μm with low stresses. A study is being carried out of the fault condition of Binary-Ice cooling (see Section 7.5.1) being switched off in one or more staves. The goal is to understand the possible impact this local "warming up" will have on the overall structural stability.

The carbon fibre cylindrical support concept offers the advantage of a single, stable and relatively lower cost structure for preassembled and tested silicon modules. A sandwich type of construction, exhibiting a zero CTE both longitudinally and circumferentially has been chosen. With this approach, a very high stiffness to weight ratio is obtained and the structural material is evenly spread on the whole barrel surface. This surface may also act as an electrical shield to reduce external pick-up and radiation. Optimization of the four sandwich cylinders is underway. This includes the understanding of the structural impact of different skin lay-up and core combinations and the influence of the end rings on the overall performance of the cylinder.

Finite Element Analysis using Algor composite software has been carried out on Barrel 1 (600 mm diameter and 1700 mm length). Loaded with its 476 modules and supported at four points, the structure shows a maximum deflection of less than 20 μm and a first eigen–frequency value above 50 Hz for an 8 mm thick sandwich with 0.2 quasi isotropic XN50A/RS-3 skins [89]. Similar analyses will be carried out in early 1995 for the other three barrels.

Manufacturing techniques for the composite cylinders are being developed in contact with industry. Typically, the inner skin plies (prepregs) are laid up on a mandrel with the required fibre orientation, then the core material is added, and the premachined end–rings are inserted. Finally the outer skin plies are laid up in a similar way as the inner skin. To apply pressure during the curing cycle, a vacuum bag technique is well suited to the proposed thin skin sandwich. If more pressure is required for curing, an autoclave could be used instead. If needed, moisture barrier coatings may be integrated during fabrication or added later. A prototype composite cylinder will be built in early 1995 and loaded to compare its behaviour with

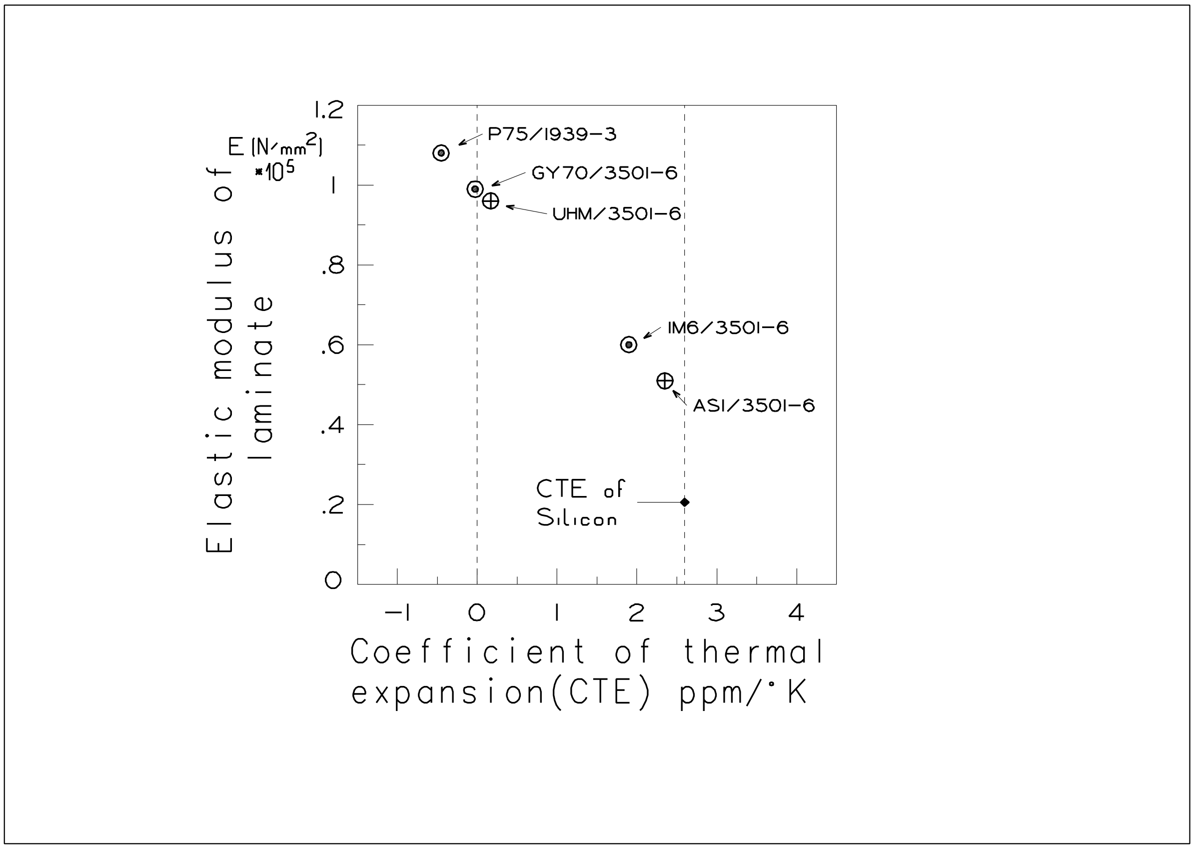


Figure 85: Correlation of elastic modulus and coefficient of thermal expansion for various materials considered for use in the tracker.

calculations.

In a cylindrical design, separate cooling pipes and module fixation points must be mounted on the barrel surface. A cylindrical design works naturally with both rϕ and Z module designs (see Section XX). Preliminary mounting studies were made for both module designs [90],[91]. Detailed studies will be pursued in 1995 and a prototype fixation, including cooling pipes, built and tested.

In the interface between modules and barrels, for all designs, there are several materials which do not match in CTE. This CTE difference could introduce stresses at these interfaces. For example, in the stave design one difference is in the detector (silicon) and the beryllium components. This results in a differential length change of 9 ppm/° K. A thermal excursion of 35° K (-5 to 30° C) generates 315 ppm of difference. This spatial difference is 20 μm over a 60 mm module. A flexible adhesive can absorb this difference. The beryllium stave itself, of 1.7 m length, will experience approximately 0.7 mm of difference over the same temperature range. The CFRP cylinder is to have zero CTE. However, cooling pipes and module fixation points will be mounted upon it. Over the span of 120 mm between the fixing point of the module/heat exchanger, cooling pipes of beryllium, aluminium, stainless steel generate 40, 83, and 58 μm spatial difference for the temperature excursion of 35° K respectively. The cooling pipe will either require a bend or loop to absorb this shrinkage or be mounted in a way which allows for differential movements. As basic scheme for the cooling pipe layout is shown in Figure 88. Finite element stress analysis will be carried out to check these effects.

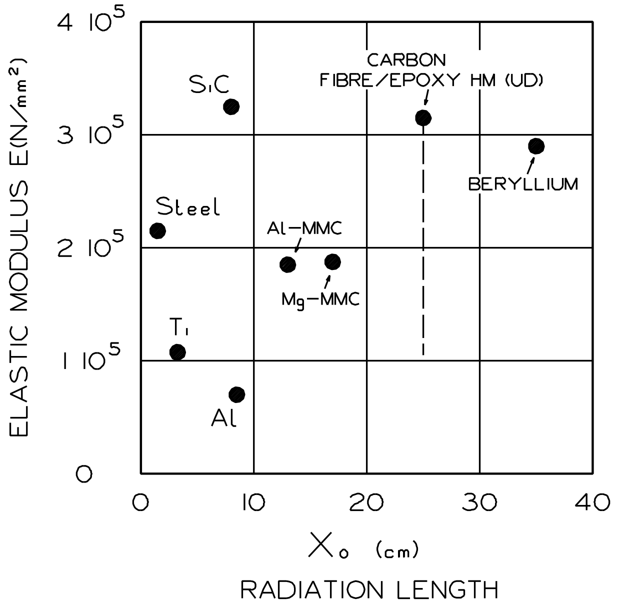


Figure 86: Correlation of elastic modulus and radiation length for various materials considered for use in the tracker.

Not yet available

Figure 87: Detail of a single stave.

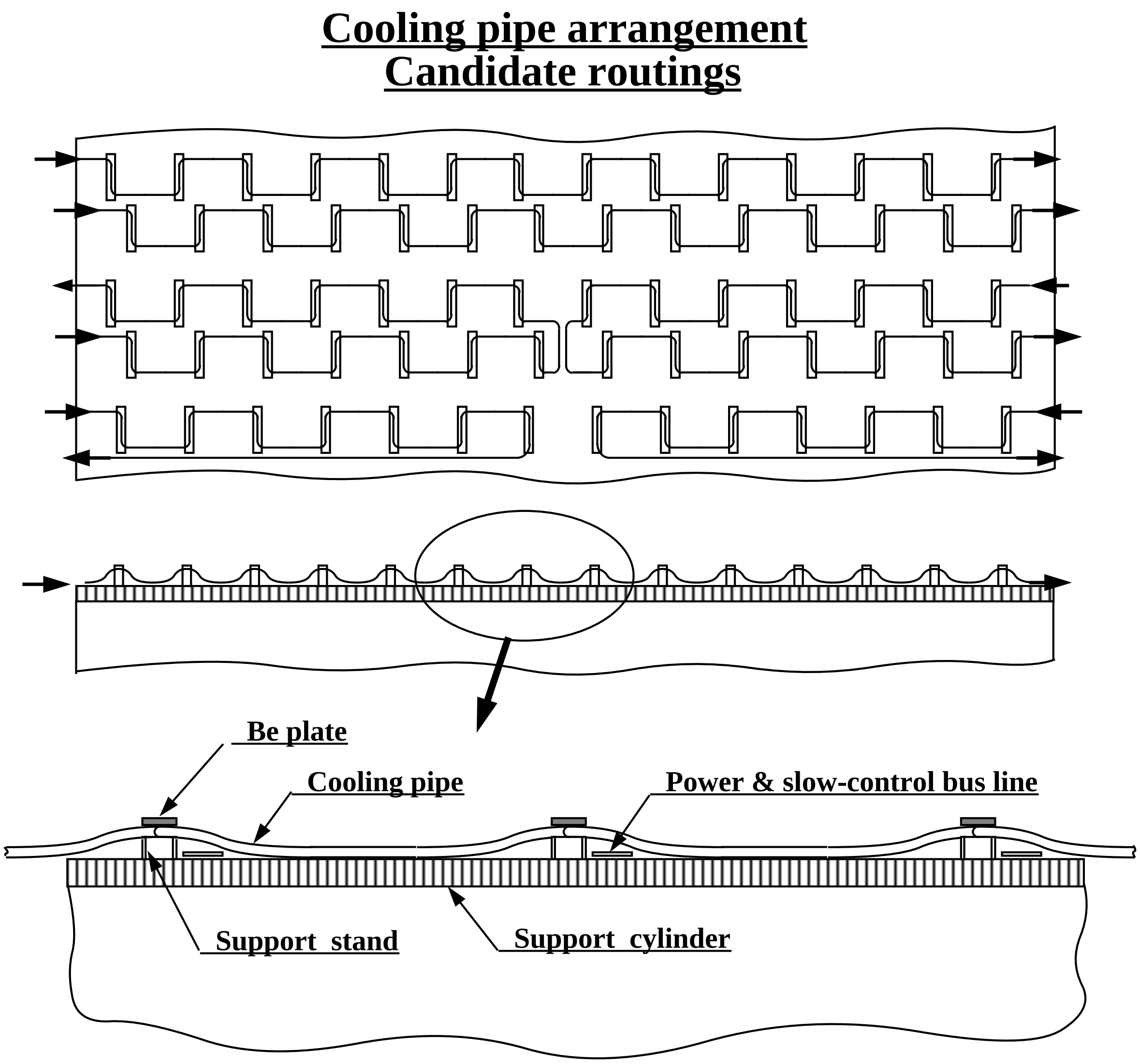


Figure 88: Cooling pipe arrangement. Three routing candidates are shown in the top figure. The radial bending absorbs the difference of the thermal shrinkage of the support cylinder and the cooling pipe.

Regardless of which support structure is chosen, an insulating enclosure is required around the detectors operating at around 0° C. Inside this enclosure there will be a flow of dry nitrogen gas at approximately 0° Cto prevent condensation. Sealing this gas envelope will not be simple considering the amount of services which have to pass through the enclosure wall. As there will inevitably be some small leaks, there will be a flow of dry nitrogen in the rest of the Tracker volume.

Finite Element Analyses Data is to be presented here describing the results of finite element analyses of the support structures. These studies are currently in progress and the results will be included in a future version of this document.

7.4.2 Forward

Currently the only element of the SCT which is part of the forward system are the GaAs wheels. These will be supported off the MSGC wheels. The mechanical discussion is deferred to that backup document for the present. Some discussions are in progress concerning a forward silicon backup. Some ideas relevant to that system might be included here in a future version of this document.

7.5 Cooling

A fail–safe, leak free system with near uniform temperature is of is a basic requirement for the Tracker. Fluid cooling, for example Binary ice or water, is being studied for use in all detectors as a prime candidate technology. In addition, a flow of cold (approximately 0° C), dry nitrogen is anticipated in the silicon barrel to eliminate the risk of condensation. Alternative schemes of evaporative and bi–phase fluid cooling are also under investigation for both silicon and pixel sub-detectors. The susceptibility of the materials and fluids to radiation damage will be studied.

Temperature monitoring will be in place on all detector modules. The information provided will be used to control the cooling system. Successful tests have been performed on cooling single sided silicon detector modules using ultralight 1 mm^2 square aluminium tubes [92]. Thermal modelling to the parameters shown in Figure XX have been produced and are shown in Figure XX. The cooling must cope, not only with the basic power load of the front–end electronics, but also with the heat dissipated in the power cables and in the detector elements. This last component will rise as a function of integrated dose.

Since the two suggested versions for the overall mechanical support structures of the SCT barrel layers feature materials of largely different CTE they require different restrictions on the global temperature profiles. The carbon fibre cylinders are supposed to show a CTE of zero in the relevant directions and therefore will suffer very little in positioning accuracy and stability if the temperature varies along the length of a barrel. The beryllium staves show a large CTE and are more prone to torsion due to their construction principle. For these, the temperature must be more homogeneous.

Below we give details on specific approaches to cooling the SCT.

7.5.1 Binary Ice

Binary-Ice has been proposed as a coolant because it offers a realistic method of removing many kilowatts of heat from a low mass structure whilst maintaining a stable temperature and imposing the minimum of stress due to thermal gradients. If, for example, it were necessary to remove 10 kW using a water flow of 1000 gm/sec, the associated temperature rise would be around 2.4°C. The addition of 3% ice to this coolant would result in virtually no temperature rise across the detector.

The term “Binary-Ice” is used to describe a suspension of microscopic ice crystals in an aqueous solution containing a freezing point depressant. When this liquid comes into contact with a source of heat, an energy flow from the hot source to the liquid results in a melting of some of the ice crystals. Since the latent heat of fusion of ice is approximately 334 Jg^-1, the addition of a small percentage of ice crystals will boost the effective heat capacity of the suspension to levels that compare favorably with some evaporative cooling systems. It is therefore possible to design a realistic system in which an ice/water mixture is flowing through the entire length of the tracker, resulting in near zero temperature changes. The potential benefits of a Binary-Ice system have been discussed elsewhere [Refs].

Binary-Ice is produced by circulating a solution of anti-freeze and water through an ice generator. The resulting suspension is then pumped through the cooling channels of the detector and heat generated within the electronics and detectors flows into the fluid and results in the melting of ice crystals. The fluid temperature will not significantly increase until all the ice has been melted. A system can therefore be designed to ensure that ice is always present in the return flow and that the silicon detectors operate in a near isothermal environment. The ice generator is a high tolerance, double skin vessel, fabricated in stainless steel, whose inner wall is chilled by a conventional freon evaporative refrigerator. Inside the generator is a motor driven scraper which is rotated, in close proximity to the vessel's inner wall, by an external motor and gearbox. The solution of antifreeze and water is pumped into the vessel and ice crystals are formed on the chilled wall. These are removed by the rotating scraper soon after nucleation. The ice crystals have typical diameters in the 50 – 500 μm range and these form a stable suspension in the cooled liquid. The suspension is then pumped into the external apparatus from which it removes heat.

The cooling capacity of Binary-Ice compares favorably with other systems. For a total heat load of 8.4 kW, a water flow of 500 gram/s would result in a 4°C temperature rise. A flow of Fluorinert liquid under the same conditions would exhibit a 16.4 degree increase. An identical flow of water with 5% ice crystals would remove all the heat with negligible temperature increase whilst a flow of only 400 g/s of Fluorinert would be required if an evaporative systems was used thereby permitting utilisation of the full latent heat of vapourisation.

A test system has been constructed in collaboration with Integral Technologie, Flensburg, Germany. This consists of a 1.7 kW ice generator and all necessary controls, monitoring and interlocks to ensure reliable operation. The fluid used is an 8% solution of propylene glycol in water with a freezing point of approximately −3.4°C.

The output of the Binary-Ice machine enters a unit that allows flows to be split, combined and measured. This section also contains heaters and control valves that allow the ice concentration in the supply to the test enclosure to be measured. The test enclosure consists of a large, insulated "Celotex" box approximately 2.0 x 1.2 x 1.0 m^3 in volume, mounted on a surface table. Inside this box is an array of ten copper tubes equipped with 30 watt, distributed heaters. The inside diameter of each pipe is chosen to be close to the anticipated bore of a beryllium stave. Each pipe is equipped with a flowmeter and several temperature sensors and a flow control valve. This array is used to simulate the heat load and distribution of ten staves, each fitted with fourteen "Z" modules.

Clearly, if a device is to operate below 0°C, considerable care must be taken to ensure that there is no risk of condensation or ice forming on the cooled detector elements or their associated electronics and mechanics. To mimic the environment inside the SCT tracker volume, a specially designed dehumidifier and air cooler has been constructed and installed below the Binary-Ice test enclosure. Air is drawn from the surrounding area into this unit. It is then chilled to reduce the dew point to around −10°C and the water vapour released by this cooling is collected, as ice, on one of two sets of refrigerated, regenerating baffles.

The cold air from the dehumidifier is then passed over a heater element which heats the gas to the required operating temperature for the current test. In this way, a gas flow is maintained throughout the test enclosure and can be used to either provide extra cooling for the detectors or to impose an additional heat load onto the binary-ice system.

A "LabView" control system is used to provide monitoring, data aquisition and slow control features. All sections of the apparatus (ice generator, flow distribution, control and thermal load power supplies, air dehumidification) are linked via an IEEE bus to multi-channel data loggers and a Macintosh computer.

Results The nominal capacity of the ice generator was chosen as 1.7 kW since this represented a worthwhile fraction of the total anticipated heat load in the final experiment and could be obtained at a reasonable price and operated from a single phase 220/240 V mains supply. To balance this cooling capacity, six 250 W balancing heaters were factory installed in the Binary-Ice unit and a further 600 W load was added to the external pipe-work. For the mixture used in the tests performed to date, the final freezing point was found to be −3.4°C.

A measurement of the ice concentration as a function of applied heat load is clearly of fundamental importance. During initial trials at Flensburg, a simple heated pipe was used to obtain an estimate of this quantity. A fraction of the total flow, in either the flow or return line, was directed through this pipe and the temperature rise along the pipe recorded as a function of applied load. Provided that sufficient heat was applied to change the composition of the fluid from a suspension of ice crystals to a homogeneous solution of antifreeze in water, an estimate could be made of the additional heat removed compared to that which would be removed by an identical flow of the liquid alone. This additional heat removal was attributed to the melting of ice in the fluid. This technique is still employed on the prototype system, albeit with some improvement.

Temperatures, pressures and flow rates are measured and recorded at points around the fluid circuit. The majority of temperature measurements are made with k-type thermocouples although several Pt100 measuring heads are also used. Flow measurements are made with rotating vane flowmeters using magnetic pick-ups.

Future Prototyping Plans and Timescales The pipe array inside the test enclosure will be augmented by the addition of extruded aluminium staves, manufactured to the current beryllium stave design. These staves will permit the inclusion of thermally realistic dummy silicon modules and later, working modules, onto the final mechanics. In this way, it is planned to study the mechanical and thermal stability of this design.

The final operating temperature of the SCT silicon has still to be specified. For this reason, it is essential that the Binary-Ice system should be able to operate over as wide a range of "sub-zero" temperatures as possible. Further unknowns are the cooling channel diameter and the radiation hardness of the coolant mixture. At present, propylene glycol is used as the freezing point depressant but this will shortly be changed to ethyl alcohol as a result of research carried out by Integral Technologie [ref: private communication, Sept. 1994]. They have reported successful ice production over the full temperature range from $-1°$C to $-40°$C. This is accomplished by varying the percentage of alcohol used. A further result of this change is a reduction in the mean ice crystal diameter which should permit the use of smaller pipe cross sections. The radiolysis of aqueous solutions of ethyl alcohol is better understood than those containing propylene glycol.

Much work will concentrate on the development of a stand-alone control programme that fully integrates the data aquisition and run control requirements of the cooling system. The chief goal of this work is to demonstrate a cooling scheme that can run unattended for long periods of time under varying conditions. It is anticipated that this will be completed by the Summer of 1995.

The operating conditions within ATLAS will be simulated to allow the integration of prototype modules onto final design mechanics. This will also allow the integration of an alignment system and a gas cooling/purging system with the Binary-Ice cooling plant. A full study of potential failure modes will be undertaken and possible ways of coping with these included in the final design.

Risk assessment Any cooling system relying on a change of state has higher risks associated with it than one using a homogeneous coolant. A principle worry with ice is clearly its capacity to block a cooling channel. This is somewhat offset by the fact that the change of state produces a very much smaller volume change than say with a liquid/gas interchange. At the ice concentrations we propose (3 - 6%) the risk of blockage is very small. Provided there is no additional source of cooling adjacent to the cooling channel, a total freezing of the coolant is impossible. Any blockage will only occur due to a physical agglomeration of discrete ice crystals. These can be cleared by powering down the electronics on that cooling path and either (a) waiting for the blockage to clear or (b) reversing the flow through that channel. Provided each cooling path is designed with only one small diameter section, this approach should prove to be effective. In all the tests performed so far, no blockages have occurred in pipes with diameters greater than 2.5 mm. Binary-Ice has been flowed through short lengths of 1 mm bore pipe but this would not be recommended for use in a final system. It is anticipated that the risk of blockage will be further reduced by employing ethyl alcohol in the coolant. The reported reduction in ice crystal diameter should add a further safety margin.

The radiolysis of aqueous coolants is a potential risk. It is known that hydrogen peroxide is a potential radiation induced by-product from solutions of ethyl alcohol in water. Hydrogen peroxide has been reported to increase the corrosion rate on beryllium pipe. This, in turn, can be overcome by the correct surface treatment of the beryllium.

7.5.2 Evaporative Cooling

There are three very good reasons for using evaporative cooling of a silicon microstrip detector. First of all the phase change process is isothermal. Secondly only a very low mass flow is needed, due to the large latent heat of evaporation (in the order of 90 kJ/kg). Thirdly, with a proper design, very high heat transfer coefficients can be achieved. Measurements have shown heat transfer coefficients larger than 45 kW/m^{2}K.

The family of fully fluorinated liquids (Fluorinert) are available with boiling points from $-2°$ C and upwards. (Lower boiling point liquids are available from other sources than 3M). The boiling point can be tailored by mixing different liquids. The liquid is non-flammable and non-reactive chemically and has a dielectric strength in the order of 9 kV/mm.

Due to the very large difference in density between liquid and vapour, the total radiation length will be minimised by separating the liquid and the vapour at the point of evaporation.

When using evaporative cooling, and liquids compatible with the electronics and the detectors, the cooling can be done on the module itself, and not through the mechanical support structure. In this way the cooling agent is in direct contact with the heat sources, reducing the thermal gradients. At the same time, thermal gradients in the support structure can be reduced which is important for the mechanical stability.

With an "in module cooling", there is no need for thermal contact between the module and the mechanical support, and the mechanical coupling is dramatically reduced. The whole mechanical fixation of the module can be done using only one fixed and one flexible point along the z axis. In this way, problems with thermal mismatch between the module and the mechanical support structure is eliminated. For evaporative cooling, module design is, according to Goessling's classification [93], an rϕ module, with the electronics mounted at the end of the strips.

Using single sided detectors, the detectors in the z-plane are separated from the u, v plane by a fixed distance (in the order of 1 mm) to accommodate the in-module cooling. The front–end amplifier will be placed directly on the detector silicon, which must be extended beyond the active area, with the cooling underneath. The cooling is done by evaporating the liquid in a thin permeable membrane, which at the same time separates the liquid and the vapour.

The channel formed between the two active silicon planes is used for the exhaust of the vapour, and thereby also controls the evaporation atmosphere. When evaporating into a pure vapour atmosphere the problems related to the Onset of Nucleate Boiling is eliminated, and the temperature of the evaporation is well defined.

The vapour flow will participate in the cooling of the detector area, together with the thermal conductivity of the silicon itself. The spacers between the two silicon planes are made of high thermal conductivity carbon fibres which contribute to the cooling of the detector area.

The rest of the electronics is placed onto a double sided hybrid substrate with a small extension. This extension is glued between the two detector planes, and the rest of the gap is closed by a high thermal conductivity spacer. Wire bonding is used for electrical connections. The hybrid is proposed to be made of a material which is closely matched to the silicon with respect to thermal expansion, and has also a good thermal conductivity.

In addition to the electrical connections needed, the module also needs connections to a liquid supply. One supply line is used for each row of modules. This could be the one square millimetre aluminium tube used by RD-2. From this tube, a flexible liquid supply line is needed to each of the modules in the row. Due to the volatile and chemically inert liquid, a number of small leaks will not affect the operation of the detector. The vapour produced will be pushed through the detector volume using an inert gas. The vapour will then be collected, condensed and degassed, before it anew is fed into the detector.

7.5.3 Liquid System

Early RD2 developments of silicon detectors included a fluid cooling optimized for a tight material budget. Coolant circulates in ultralight 1 mm^2 square aluminium tubes. Thermal contact with modules is made with heat sink compound, thus allowing the pipe to expand or retract without transmitting strains to the silicon modules. Several successful tests were made with Fluorinert as coolant and dummy modules equipped with heat foils. The system was able to remove heat from single sided modules [92]. Development is needed to further engineer this ultralight system.

The Leakless (Registered) cooling system The Leakless(R) cooling system, consists of a secondary liquid cooling circuit, below atmospheric pressure, removing the heat produced by the detector to transfer it into a primary circuit situated outside the detector. It was developed by Michel Bosteels of CERN. A study was recently made to investigate the application of this system to the barrel silicon sub-detector. Using water as a coolant, calculations show that a pipe 3×2 mm^2 and 1.7 m long is able to remove the 60 watts produced by the 14 silicon modules (fixed on a stave or in a row on the carbon-fibre cylinder) with a difference in pressure of 200 millibars and a difference of temperature of 2° C. This is considered small enough to be manageable. Naturally, all other manifolds and pipes should be carefully designed to produce a very small drop of pressure to be compatible with the overall pressure available within the Leakless (R) system. The main disadvantage induced by this limited difference of pressure is that it prevents the use of radial cooling pipes between the inner detector cavity and the experimental area where most of the cooling system components will be located. The intrinsic operational safety of the Leakless system (liquid being under atmospheric pressure, air will enter the circuit, instead of liquid escaping, in case of leak), coupled with the experience gained in operation on several CERN experiments during many years, are important criteria to be considered in the choice of a fail–safe cooling system.

7.6 Assembly and Construction

Silicon strip modules will be assembled and tested on dedicated production lines. Following this they will be mounted, in groups, either on staves or cylinders. In the former case this can be done at the module assembly site. In this section we discuss the assembly and construction process beginning with completed modules. Further discussion can be found in Section 7.7.1 which specifically concerns alignment and survey.

7.6.1 Requirements on Module Mounting

In order to obtain maximum coverage in the tracker, minimize systematic degradations in performance, and allow for ease of repair, certain requirements are placed on the module mounting. These were introduced in Section 3.6 and are reproduced below.

- The modules are to be tilted at 6°to allow overlapping in the ϕ direction (Figure 89). The tilt angle also allows a better charge sharing together with Hall effect as discussed in the bellow. The z overlapping can be done by changing the radii of the adjacent modules in the z direction.
- The modules must be placed to an accuracy of (± 50 μm, ± 250 μm, ± 100 μm) in the ϕ, r, z directions, respectively.
- The overlap in z will be 3 mm. The overlap in ϕ will be 8 strips.
- Each module must be easily dismounted mechanically, electrically, and optically.

7.6.2 Stave design

This option is based on a modular component and assembly concept and has the cooling channel incorporated in the extrusion. Modules are first mounted on the beryllium stave (supported by a jig also serving as a transport container). This is shown in Figure 90. Opto-electronic boards are added and all services connected. The fourteen modules are assembled respecting the specified z overlap. They are tested as they are assembled to the stave. The completed stave is then tested and transported to the barrel assembly area. To assemble each barrel, the four support rings will be positioned in a jig and staves will be mounted on the rings. Testing can be carried out as assembly proceeds. Survey of the completed barrel can be carried out through the open structure.

7.6.3 Carbon fibre cylinder design

The four carbon fibre cylinders at radii of 300, 400, 500, and 600 mm and length 1700 mm will be held in specifically designed jigs. Service cables will be laid and fastened on the cylinder (connections, not connectors, are foreseen at the end of each cylinder). Cooling pipes may then be clipped in place and tested (separately or in batches depending upon how they are connected together).

Assembly is carried out by mounting modules in longitudinal rows. This is shown in Figure 91. Modules are staggered to give the specified overlap in z. Two module fastening methods are under study. One requires an external microscope to align the modules which are pinned in position with UV cured adhesive filled through small holes. The other to a mechanically precise location with two fixation points for each module, one precise and one slotted to avoid unwanted strains or stresses to be transmitted from the structure to the modules. Thermal contact is improved with a heat sink compound which is applied just prior to module mounting. After curing, fixation points are permanently bonded to the cylinder shell. This method was prototyped and may compensate radially up to 500 μm thus not requiring high precision in the dimension of the sandwich itself. This technique is also compatible if a moisture barrier coating is used and introduces

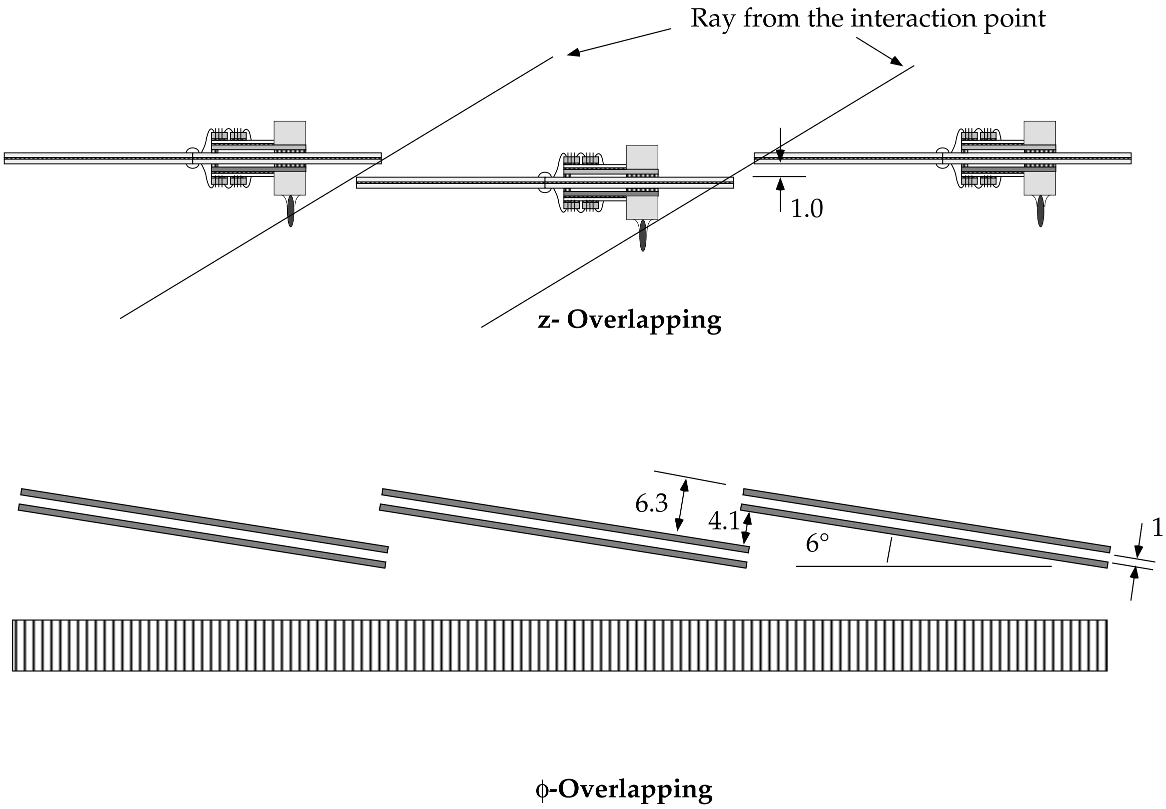


Figure 89: Overlaps in z and ϕ. The double layer in ϕ represents the alternating radii of the modules along z.

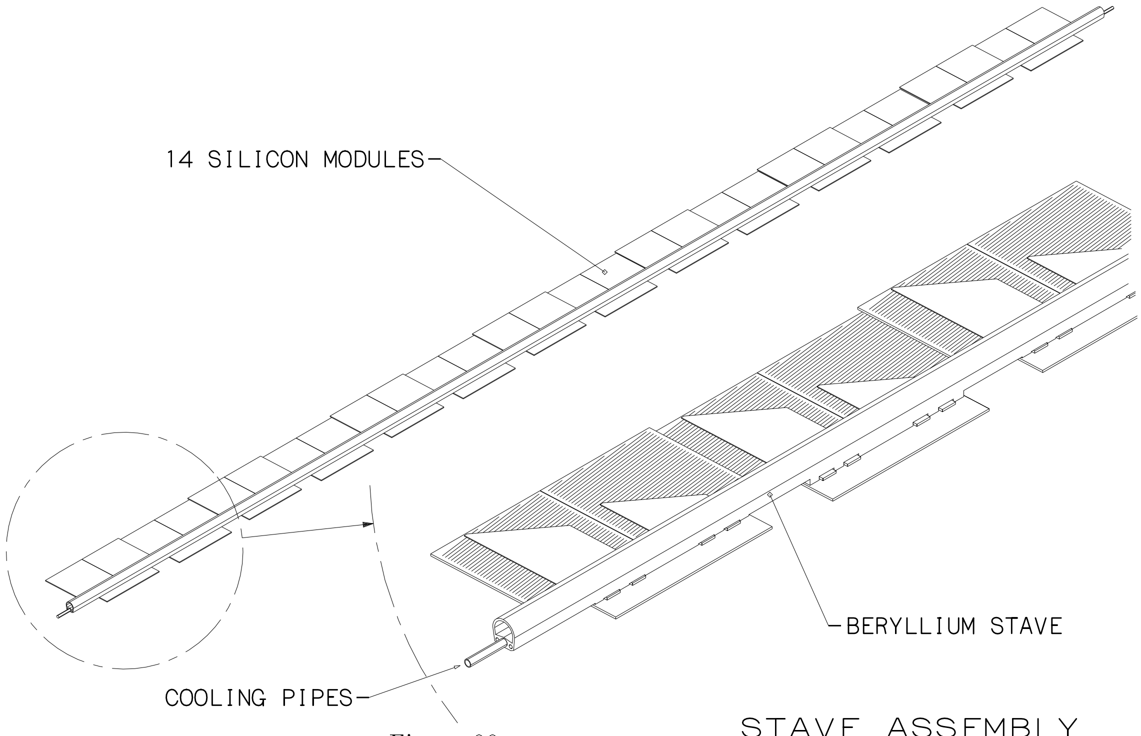


Figure 90: Assembly of modules on a stave.

no break in the barrier. The fasteners themselves will be mass produced mouldings in thermoplastic resin reinforced with short fibres. They may also serve as cooling pipe support. The aluminium alloy pipe will be clipped into place but will be allowed to move longitudinally. This arrangement will prevent the transfer of thermal strain to the silicon modules. A good thermal contact is assured using a heat sink compound such as Dow Corning DC 340 [94].

Once the modules are mounted, connections will be made with services already in place on the cylinder.

Tests can either be performed at the level of individual or groups of modules. For these, jump connections to the test electronics are made at the end of the cylinder and optical fibres are connected to the opto-electronic boards. After completion of modules assembly on the four barrels final tests and survey can be performed before assembling the four barrels together.

7.6.4 Silicon barrel assembly

After completion of assembly and tests of the four silicon strip barrels, they can be assembled together to form the full barrel SCT. Initially, barrel 4 will be supported from the outside. Barrel 3 will be inserted in Barrel 4 with the appropriate handling jig, temporary links mounted and handling jig removed. Barrels 2 and 1 will be similarly inserted. End structures, providing the final interlinking will be fastened and temporary interlinks removed. In the case of composite cylinders, services have to be connected at the end of each cylinder (and temporarily supported before assembly with TRT) and a full test performed.

7.6.5 Inner detector assembly and installation

In a CERN surface building, pixel, silicon strips and TRT detectors will be assembled and tested separately before being interlinked to form the barrel region. This requires special tooling be developed in order to perform the assembly in a safe, precise and controlled way. Once in the transport cradle (which will contain the same rail system as in the cryostat) cooling pipes, cables and optical fibres will be laid on the barrel end faces up to the outer TRT radius and then supported by a lightweight structure up to the patch panel to

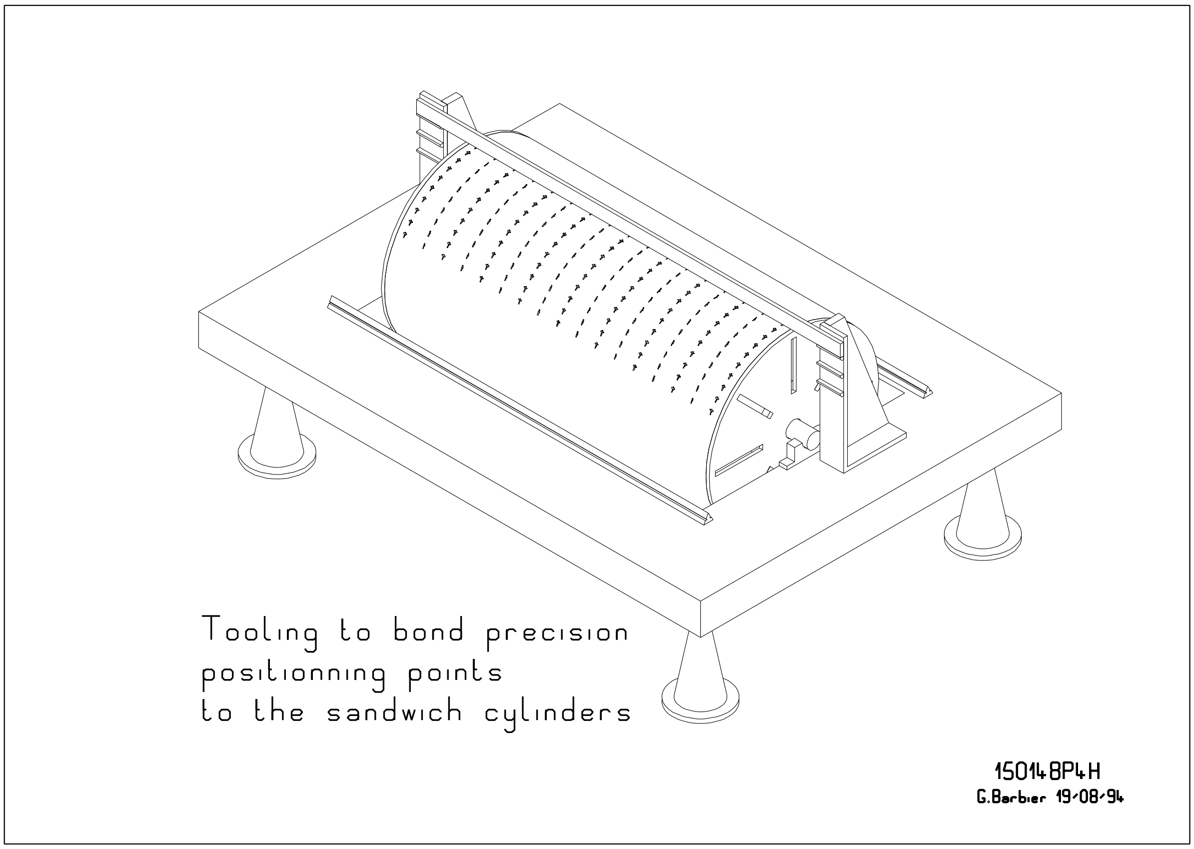


Figure 91: Assembly and fastening of modules on a cylindrical assembly.

allow for the introduction of the forward regions. In this configuration, the barrel region including services at both ends is a 6.8 m long device.

TRT wheels, and MSGC and GaAs disks will be assembled together to form the two forward regions. Cables, cooling and gas pipes will be arranged on the outer shell of the TRT. The forward regions can then be transferred to the transport cradle, rolled into position and linked with the barrel.

The complete inner detector, together with the central part of the beam pipe, will then be lowered into the pit. The cradle will be positioned in line with the rail system in the cryostat and the inner detector moved into its final position, prior alignment and survey.

The maintenance scenario consists of three types of access:

- Short term: Typically lasting one week which will allow access to both ends of the Inner Detector and the patch panels.
- Medium term: Typically lasting three weeks which will allow for the Forward Region to be pulled out giving access to the Barrel ends.
- Long term: Occurring at most once a year, which would involve the removal of the complete Inner Detector to the surface for major maintenance or upgrade. It is a requirement that the pixel and silicon detectors be kept at their operational temperature during removal or installation at maintenance times.

Due to the induced radioactivity after several years of LHC operation, the installation, including all the services connections, will be studied with particular attention to reduce hazardous exposure to personnel.

7.6.6 Services

Services from the barrel region will be routed up the crack between barrel and forward sections and then along the inside of the liquid argon cryostat inner cylinder wall. All services from the forward region will be routed between the outside diameter of the forward TRT and the inside of the barrel services. There have been two rounds of requests for service information from detector elements and the presently know data is summarized in Table 28. Service quantity and routing will be closely watched since this is always an underestimated area in both quantity and impact on design for installation and access.

Pixel 1 at 115 mm	80 cm^2
Pixel 2 at 165 mm	110
Barrel 3 at 300 mm	62
Barrel 4 at 400 mm	84
Barrel 5 at 500 mm	103
Barrel 6 at 600mm	125
TRT Barrel	227
MSGC	350
GaAs	200 (sealed by channel counts)
Dry nitrogen	100
TRT Forward	663
TOTAL	2104 cm^2
Packing factor 40%	842
Total	2946 cm^2 (Barrel = 1107, Forward = 1839).

Table 28: Estimate of services cross sectional area at z=3500 mm.

7.7 Alignment

7.7.1 Global Alignment

In this section we discuss specifications and implementation of a global alignment and monitoring system. The present proposal for system can be found in an ATLAS note [95]. The refer that document rather than repeat the discussion here.

7.7.2 Local Alignment

Requirements The final alignment of the modules within the SCT will be done using tracks. The installed modules must however be surveyed prior to installation to provide a set of reasonable optical constants from which to begin the track alignment. The specified survey and assembly accuracies were presented in Section 3.6. After track alignment, the accuracy with which the strip locations must be known is set by the position resolution in ϕ and is of order 10 μm. The impact of radial displacements is less severe and the required precision in the alignment constant is of order 50 μm.

During assembly the precision with which the elements must be aligned is determined by a number of factors. Experience shows that difficulties arise with the final survey if the assembly precision is more than a few times the final accuracy required. It is desirable that the mechanical precision is sufficient that the trigger rates at LEVEL-2 are not seriously increased in the absence of survey or track fitted calibration constants (studies have shown this to require precision in rϕ of order 200 μm). Assembly precision is also required to avoid mechanical interferences between modules themselves and the with the support structure. Another factor which determines the required accuracy of assembly is the final mass of the system; a less accurately aligned system requires a greater overlap of detectors to ensure no gaps. In particular if the alignment is performed between neighbouring modules cumulative errors can build up. Taken together these demand that the detector elements should be placed with the specified accuracies.

The ATLAS silicon tracker may suffer from survey problems that previous silicon detectors have not. Part of the power dissipated in the system is proportional to data rate and will therefore depend upon the machine luminosity. During a fill this may result in changes in the temperature profiles in the support mechanics. Since alignment using tracks is a process that averages over a substantial period of time (at least historically), it may be necessary to track the location of the detector with an alignment system that is independent of tracks.

Assembly The details of how the detector is assembled is dependent on the final choices of module and support structure design. As an example we outline the assembly procedure associated with the beryllium stave design. Use of Frequency Scan Interferometry is assumed; this is described briefly following this section. With the possible exception of the module assembly, all of the following steps are conducted on optical tables in a clean room with precision temperature control (0.5°C). Humidity is also controlled and in the conservative approach the assembly takes place at the nominal operating temperature.

- The detectors and hybrids are assembled into modules. These are built with an internal accuracy consistent with the final survey precision desired. Each module has three survey marks on it which define where it must go in the detector.
- The staves are assembled, complete with power, optics and cooling looms.
- The modules are attached to the staves using a conventional travelling microscope assembly jig. The stave/module assembly is then internally accurate at the nominal operating temperature. This stage of assembly is carried out at the nominal operating temperature. Each stave/module assembly has three corner cubes attached for use in the next stage of assembly.
- Stave/module assemblies are then attached to the end rings to form barrels. The end rings are supported in space by a dynamic assembly jig, which maintains the two rings relative position to 2 μm accuracy using a small set of FSI distance measurements; the heads are attached directly to the end rings. The stave/module assemblies are presented (properly tensioned) to the rings in the correct orientation; the rings do not rotate. The location of the stave/module assemblies in space is measured using the corner cubes attached to them. The measurement of the stave/module assembly position and the two ring position is relative to a set of four floating FSI monitor heads. These measure the position of each other and also the rings and stave/modules. A computer can then reconstruct the relative position of stave/modules and rings without relying on any large scale mechanical structure to be precise a priori.
- The barrels are nested and the end-rings connected together to form the drum. The end rings are equipped with their final monitoring system prior to this stage and it is used to locate the barrels correctly.

- The drum is inserted into the outer space frame and tensioned to stretch the staves into their operating position.

Frequency Scan Interferometry (FSI) Interferometric methods for measuring movements are well known. By forming the interference between a fraction of the light from a coherent light source and another fraction which has bounced of a target it is possible to measure the motion by counting the number of dark/light fringes that occur during the motion. The same principle is invoked if the distance is left constant and the frequency of the coherent light source is scanned. The number of light/dark fringes and the frequency scanned give a measure of the absolute distance between source and target. The precision possible in principle is sub–micron, but in practice, for the applications in question, limited by a number of factors to the micron range. A detailed description of the system under development can be found in [96, 97, 98, 99, 100, 101, 102]. The salient features of the system are

- Measurement of the absolute distance between the end face of an optical fibre and the point in at the geometric corner of a corner cube. Continuous operation is not required.
- Very low mass system. In the minimal case it consists of two 125 μm diameter fibres at one end and a 3 mm diameter air corner cube at the other. The bulk of the system is external to the detector.
- Radiation hard. Provided that the fibre optics and the coating on the corner cube are radiation hard, so will the system be. It is known that fibres can be radiation hard to the extent required and it would seem a reasonable supposition that the mirroring can be.
- Low cost per measurement. The lasers and associated equipment are expensive, but can be multiplexed so that the incremental cost per measurement is that of the fibre and the corner cube.

  The potential flexibility of FSI has lead to it being actively developed within ATLAS. FSI has been shown to work in other fields, however it is not yet demonstrated to be viable in a form suited to the applications intended. The methods used for assembly and for survey would be different if FSI was not available. If expectations are incorrect and FSI will not work, this will be known very early and an alternative assembly/alignment strategy will be developed using more conventional techniques.

7.8 Program

In this section we present a schedule of milestones to be met in the development of the mechanical design of the SCT supports and installation procedure.

8 Vertexing Layer

Two designs for a vertexing layer at small radius are being considered. The first is a pixel device at a radius of 4 cm. This is discussed in detail in [1]. The second design, described in this section, uses silicon micro-strips and is based on existing technologies and on experience from the DELPHI Vertex Detector. It envisages crossed strips on double sided silicon at a radius of 6 cm, with $\sigma(R\phi) = 10\ \mu$m and $\sigma(z) = 20\ \mu$m and only 0.5% X_0. More details can be found in [103].

8.1 Motivation

The impact parameter resolution given by a cylindrical detector layer can be parameterized in the following way:

$$\sigma_{\mathrm{IP}} = A \oplus \frac{B}{p_{\mathrm{T}}} \frac{1}{\sqrt{|\sin\theta|}} \tag{8}$$

$$\sigma_z = A_z \oplus \frac{B_z}{p_{\mathrm{T}}} \frac{1}{\sqrt{|\sin\theta|}^3} \tag{9}$$

where A is the asymptotic resolution for high transverse momentum charged particles and B is the multiple scattering term. The dependence of both terms on the first layer radius is shown in Fig.92 for different resolutions and radiation lengths.

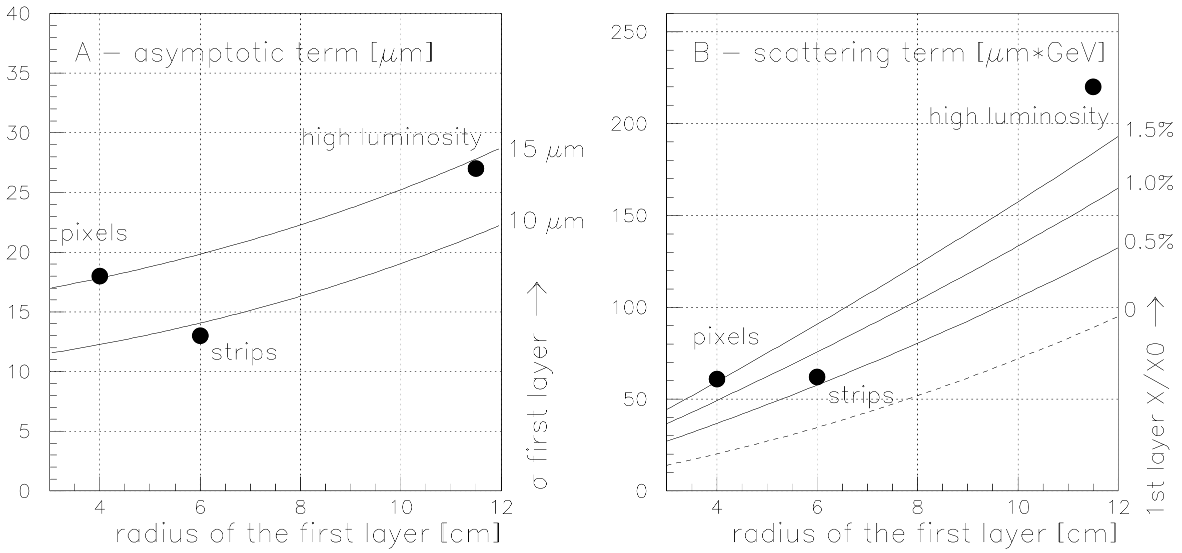


Figure 92: Asymptotic and multiple scattering terms of the impact parameter resolution versus the innermost layer radius. Lines represent approximate analytical calculations. Points represent Monte-Carlo calculations for the different variants of vertexing layers of ATLAS Inner Detector.

For B-physics studies, where typical track momenta are of the order of a few GeV, it is essential to minimize the multiple scattering contribution to the measurement error. The radiation length of the innermost layer should be minimal. For other physics, involving heavier objects (e.g. t-quark), where track momenta are higher, it is more important to improve the resolution in order to reduce the asymptotic term. For both, it is advantageous to reduce the innermost layer radius.

We propose to build the vertexing layer using double sided silicon strip technology. Silicon strips offer the best performances in the critical aspects discussed above: spatial resolution below 10 μm for $R\phi$ and below 20 μm for z readout with $< 0.5\%$ X_0. The layout is similar to the one used in the DELPHI Vertex Detector, having no electronics and almost no support structures in the active region.

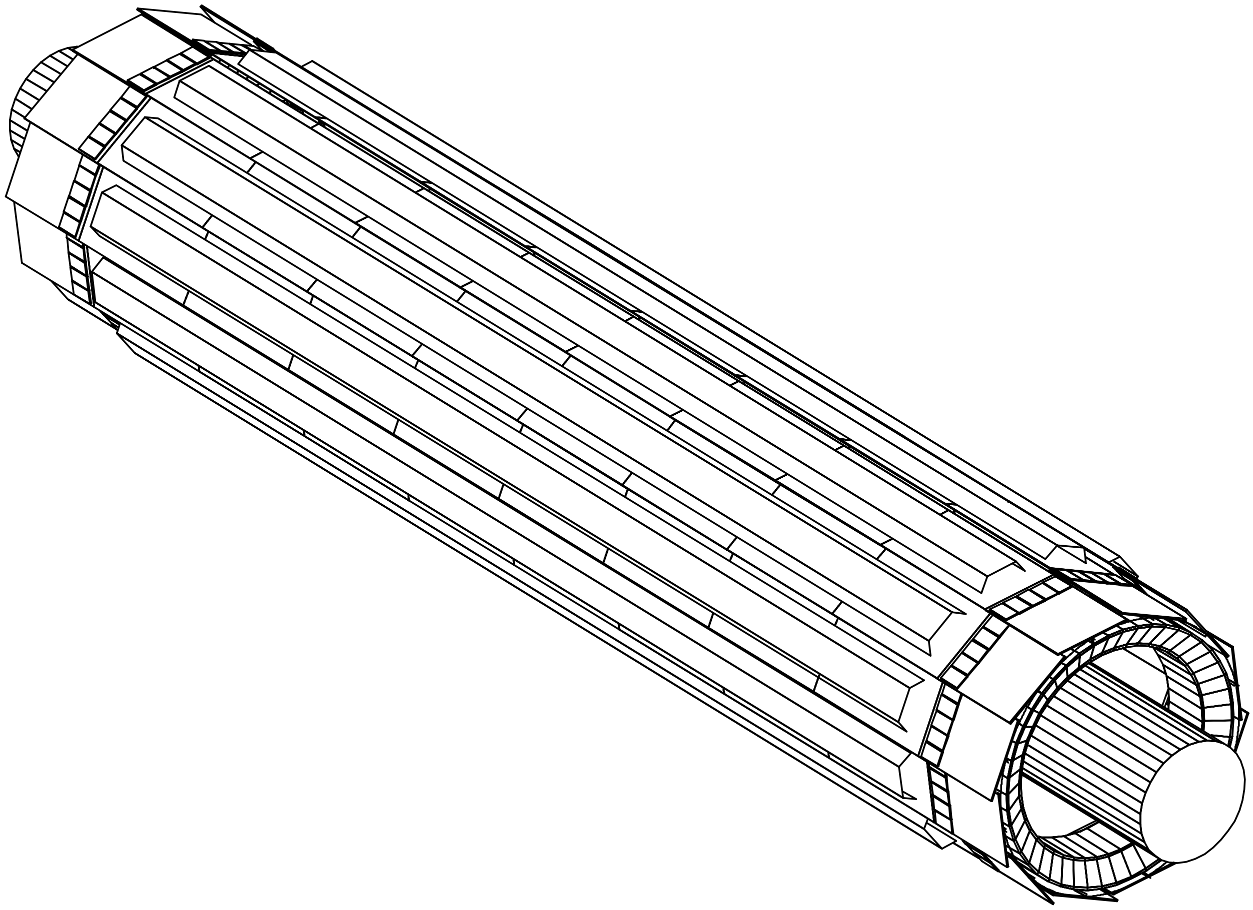


Figure 93: A view of the silicon strip vertex detector.

The radius of the silicon micro-strip vertex detector was tentatively set to 6 cm at the time this detector was proposed [104]. According to present radiation damage predictions this would allow 5 to 6 years of low luminosity operation. It should also be possible to operate such detector for one year at the intermediate luminosity of $5 \cdot 10^{33}$ cm^{-2}s^{-1}. Both radiation damage predictions and occupancy calculations indicate that the radius can be decreased to ~ 4 cm if only low luminosity running is foreseen. The final decision on this parameter will be taken when more precise radiation damage predictions are available and a more complete study of pattern recognition has been carried out. An artist's view of the layer is shown in figure 93, with a 6 cm layer radius and a 5 cm beam-pipe diameter.

The readout scheme proposed here is based on existing bipolar amplifiers and allows 52 cm long detector modules. Assuming a radius of 6 cm, this gives a rapidity coverage of $|\eta| < 2.2$ (or $|\eta| < 1.5$, subtracting 2σ of the primary vertex distribution from the length).

The impact parameter resolution degrades for small polar angles, and is not useful beyond $|\eta| \sim 2$. Therefore, any extension of a vertex detector does not justify sacrifices in resolution of the central part, resulting from increased capacitance.

The impact parameter resolution given by the layer, in the present layout of ATLAS ID, is

$$\sigma_{\mathrm{IP}} = 13 \oplus \frac{62}{p_{\mathrm{T}}} \frac{1}{\sqrt{|\sin\theta|}} \quad [\mu\mathrm{m}] \tag{10}$$

$$\sigma_z = 39 \oplus \frac{90}{p_{\mathrm{T}}} \frac{1}{\sqrt{|\sin\theta|}^3} \quad [\mu\mathrm{m}] \tag{11}$$

(p_{T} in GeV), for 6 cm radius.

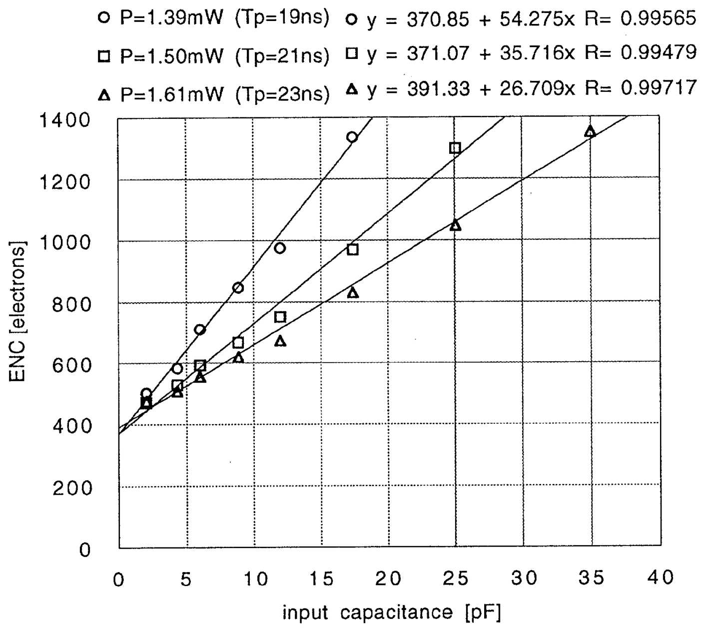


Figure 94: Measured noise slope of the bipolar amplifier designed for NA50 experiment

8.2 Detectors

The basic detectors are assumed to be double sided, double metal silicon strip detectors 6.5 cm by 3.3 cm, with strips crossed at 90°. The $R\phi$ readout will be done using p-side strips with 50 μm readout pitch and 25 μm diode pitch. Four units will be bonded together, forming 26 cm long half-modules. This geometry should achieve a resolution of about 10 μm, taking into account the signal-to-noise ratio which is expected due to the big capacitive load (see next section).

For the z readout, we propose to use the n-side strips with 200 μm readout pitch. Capacitive charge division should allow us to reach a resolution of $\sim$ 20 μm. The widths and the number of intermediate strips will be optimized to minimize the charge loss due to back-plane capacitance.

Since the readout electronics will be placed only at the end of a module, the baseline option for the detectors is double metal technology. Alternatively we have at our disposal another well proven technique, namely the fanout using kapton cables, which introduces only a small additional amount of material.

Because the vertexing layer will be placed at a radius of 6 cm or lower, the radiation damage to the detectors is a critical issue. The increase of depletion voltage will be the limiting factor determining the detector lifetime. Estimates of radiation damage predict lifetimes of 5 to 6 years at 10^{33} cm^{-2}s^{-1} [104]. An expected leakage current is $\sim 2\mu$A per strip after 5 years, assuming detector cooling to $\sim 0^{\circ}$C.

Occupancy from minimum bias events is about 1% at 10^{33} cm^{-2}s^{-1}, assuming 26 cm strip length.

8.3 Front-end electronics

The layout of the vertexing layer results in a big capacitive load for the front-end electronics. The deconvolution CMOS front end electronics would have to be run with a high drain current in the input stage in order to achieve tolerable noise levels with this high capacitance. An alternative solution would be to use front-end electronics based on a bipolar input transistor. Three versions of fast bipolar front-end electronics have been developed; two for the SDC silicon tracker and one for the NA50 silicon multiplicity detector. The noise of the NA50 chip has been measured at different currents used in the input stage. Very promising results which have been obtained are shown in figure 94. For a capacitive load of 35 pF, a noise $<$ 1400 electrons has been measured with a total power dissipation of 1.6 mW/channel, including the preamplifier,

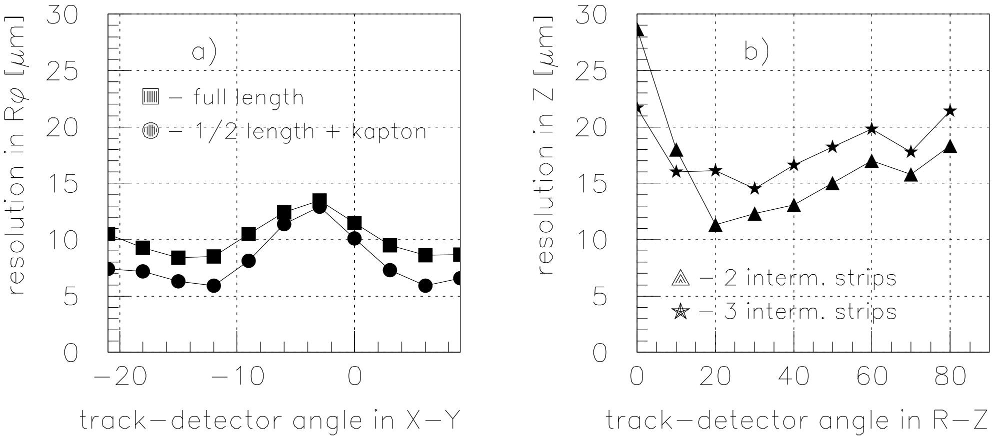


Figure 95: Spatial resolutions in $R\varphi$ and in z as a function of track-detector angle (degrees). In both cases the angle is in the plane orthogonal to strip direction.

the shaper and the discriminator.

2μA of leakage current per strip introduces some extra parallel noise - 700 electrons for the bipolar amplifier with 20 ns peaking time and 480 electrons for the deconvolution CMOS front-end. The shaping given by deconvolution reduces the parallel noise [105], but in both cases the series noise due to detector capacitance is dominant.

To obtain the required spatial resolution, capacitive charge division will be used and therefore the pulse height information is required. The bipolar preamplifier-shaper chip will be followed by an analog data buffer and sparse data scan circuitry. The possible readout schemes are those proposed for the main silicon tracker and no extra development will be needed, except for the adjustment of the preamplifier.

8.4 Spatial resolution

A Monte-Carlo program [106] was used to calculate the spatial resolution from the detectors and readout electronics discussed above. The program simulates charge collection in a silicon detector and generation of the current pulses on readout strips. The response of the front-end electronics is then calculated using SPICE. The full equivalent circuit of the strip detector, as well as models of front-end electronics including noise parameters, was implemented. Finally realistic cluster search and position calculation algorithms are used to determine spatial resolution and efficiency. The program agrees well with existing experimental data for different strip pitches.

We assume strip capacitance values of 1.2 pF/cm for 50 μm readout pitch and 25 μm strip pitch on p-side, 0.5 pF/cm for kapton fanout, 1.5 pF for 200 μm readout pitch and 67 μm strip pitch on n-side. Results for the 26 cm long half-module are shown in figure 95 for $R\phi$ and z strips. In both cases the scanned angle is in the plane perpendicular to strips. For a 3.3 cm wide detector placed at a radius of 6 cm and rotated by $6°$, angles between -21 and $+9°$ are of interest in x-y plane.

In the case of $R\phi$, even with 26 cm long strips, the average resolution is $\sim 10\mu$m. This could be improved by using 13 cm long strips and a kapton fanout. For tracks at $-3°$, the Lorentz angle compensates the track angle and the charge is collected almost completely on one strip, deteriorating the resolution.

In the case of z-strips, the resolution given by 200 μm pitch with one intermediate strip is adequate for all tracks except for a small fraction of tracks which are almost perpendicular to the detectors. The benefits of varying the strip pitch along z will be investigated.

The studies of different readout pitches and connection schemes will continue in order to optimize these parameters.

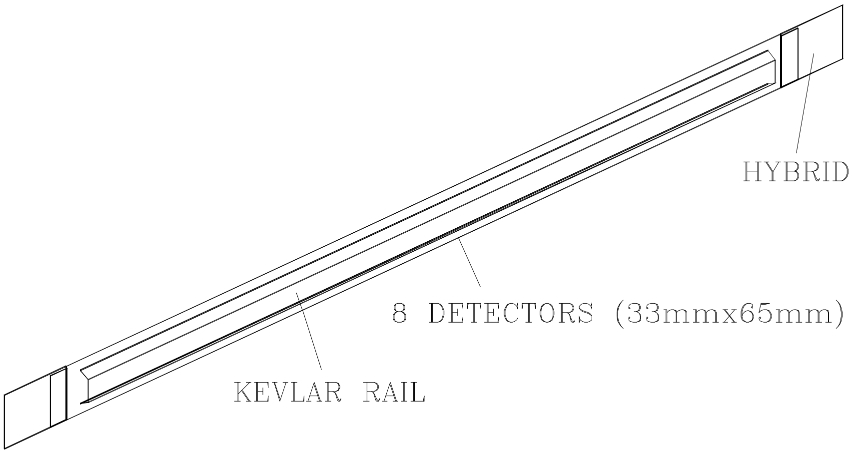


Figure 96: A view of the detector module

8.5 Mechanical structure and cooling

The mechanical design of the vertexing layer is based on experience from the DELPHI Vertex Detector. An 8 detector long module will have all the readout electronics and its cooling placed near the end-rings. Essentially only silicon detectors remain in the active area of the layer. Mechanical stability is given by a light-weight kevlar fiber structure glued to the detectors. This is shown schematically in Fig.96. The chosen solution minimizes multiple scattering.

Similar modules of 48 cm length have recently been constructed for DELPHI. Mechanical stability measurements have been done and the modules fulfill the requirements for the DELPHI Vertex Detector.

Arrangement of the detector modules is shown in Fig. 97. 12 modules of 3.3 cm width are needed to form a layer of 5.7 cm average radius. Small overlaps in $R\phi$ plane are include to allow the alignment with tracks.

The cross-section of the end-ring, made from carbon fiber is also shown in Fig. 97. Aluminum, as used in the DELPHI Vertex Detector, is a possible backup. More studies and construction of prototypes are needed before a decision on this point can be taken.

A cooling channel of $\sim$ 12 mm^2 cross-section will be sufficient to extract the heat produced by the readout electronics. The maximum radiation length of the end-ring support structure including cooling pipe is $\sim 4\%$ X_0, localized in the $2.2 < |\eta| < 2.5$ rapidity range. The ceramic hybrid and readout electronics contribute another 4% X_0 in the narrow $|\eta|$ region of 2.2 to 2.3 because of the shallow angle.

It is expected that the temperature of detectors will have to be maintained at the level of 0 to 5 °C and the cooling of the vertex detector will be integrated with other SCT components.

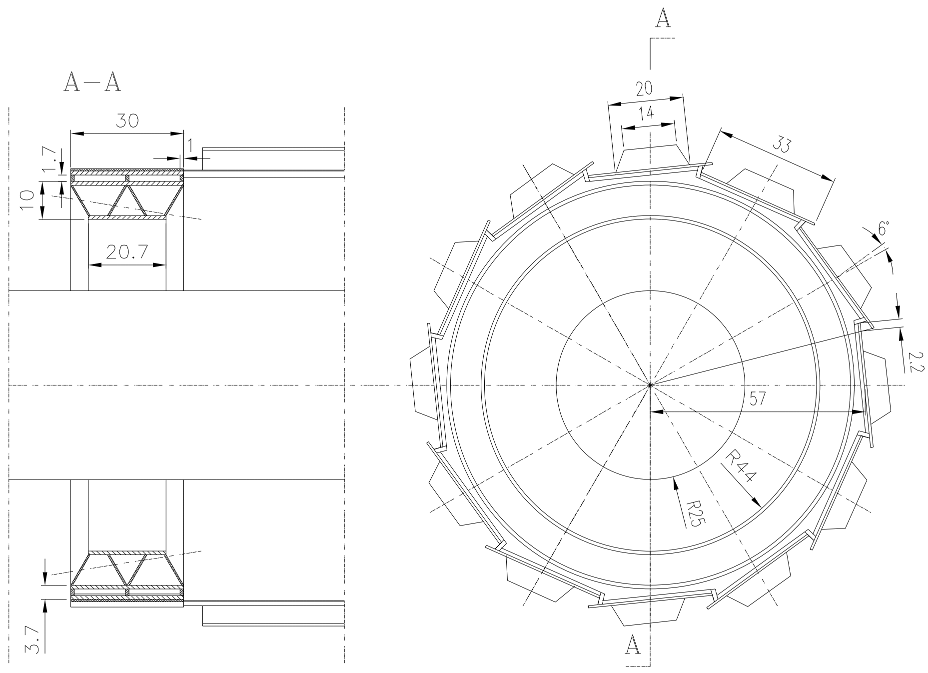


Figure 97: Cross-section of the end-ring support structure (left) and a view of the layer seen in X-Y plane (right)

9 Silicon backup for the forward tracker

A design for the forward region based on silicon devices is also under consideration [107] as an alternative to the MSGC one. This has the advantage of using a single technology for the strip detectors, and identical readout electronics in the barrel and forward regions.

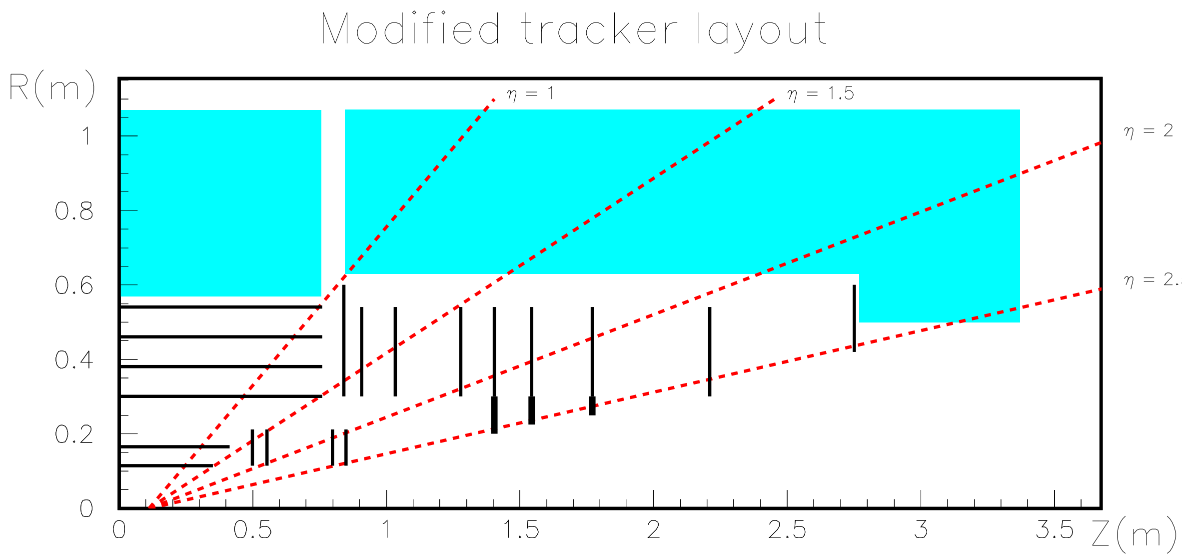


Figure 98: Conceptual layout of the silicon option for the forward precision tracking. Thin lines represent silicon detectors, while thicker lines represent GaAs detectors.

Figure 98 shows the conceptual design for the silicon option. It also incorporates improvements to the transition region which could be implemented with either technology option. All the dimensions of this conceptual layout will be optimized following a full engineering and performance study. The performance in the region of the transition from barrel to disk geometry is improved by reducing the outer radius of the barrel SCT and the inner radius of the barrel TRT by 6 cm. The first eight forward disks consist of silicon detectors in the radial range 30–54 cm, with some GaAs extensions to continue the coverage down to $|\eta| \sim 2.5$. The first disk also has an extension outwards to 60 cm, completing the hermeticity of the transition. The last disk covers the radial range 40–60 cm at $z \sim 275$ cm, replacing the 'full wheel' foreseen in the baseline layout at the end of the tracker volume. This solution has only slightly reduced performance in momentum resolution, and offers considerable savings in the area of precision detectors required, both for the GaAs substrates (saving ~ 1 m^2) and the barrel SCT (saving ~ 8 m^2). The GaAs detectors can be integrated into the silicon mechanical support and the readout architecture for the silicon.

It is proposed to use only two stereo layers per disk in the silicon option, as in the barrel design, rather than the three foreseen for the MSGCs. The extra silicon area needed would be 20.2 m^2 of single-sided detectors. The modules would be 12 cm long, as in the barrel, with two rings of modules per disk. A variety of detector designs and readout solutions can be considered, including 'keystone' geometries for the strips which simplify the use of the detectors in the LVL2 trigger. The material per point would be $\sim 1.6\%$ X_0. A combined development programme is under way to find an optimal solution, taking into account the needs of both barrel and forward regions.

In the forward region, the radiation damage expected at 30 cm radius increases only slowly with z (see Section 2.3), and therefore the detector lifetime is expected to be similar to that in the barrel. However, the radiation varies rapidly with radius, and so the exposure will not be uniform along the strip length. The effect of this both during and after inversion of the detector substrate type will be investigated.

Milestones have been set to establish the viability of this design by the end of 1995. These include the choice of detector type and geometry, and the study of radiation damage as noted above.

10 Performance of Proposed Layout

10.1 Introduction

The SCT is a crucial component of the ATLAS Inner Detector. The required tracking performance may only be obtained by the integration of the TRT in the Barrel and the MSGCs and TRT in the Forward region. Nevertheless, the SCT should have significant standalone capabilities in the Barrel, where it constitutes the only precision tracking. In the Forward region, an equivalent role will be undertaken by the MSGC (or silicon) disks in combination with the Forward Pixels and GaAs disks.

Some reasons for ensuring a level of standalone performance are:

- *Alignment within the SCT:* At the beginning of operation, it will be helpful to be able to concentrate on the alignment of one subsystem without the added complications arising from problems inherent in detectors with very different technologies and precision.
- *Inner Detector alignment:* Having aligned the SCT, it will provide a good reference by which to align the TRT and the Forward disks.
- *Efficiency determination of other systems:* If good track-finding efficiency can be achieved in the SCT, this will enable the efficiency and two-track separation of the TRT to be determined.
- *Complementarity of efficiency:* It is likely that because of the good pattern-recognition capabilities of the TRT, it will be used to initiate track-finding. However, in situations with high occupancy (high luminosity or in the core of narrow jets), it will be essential to search for tracks in the SCT alone.

In this chapter, the emphasis will be on the Barrel tracker where the SCT technologies can be considered to provide a coherent detector. Nevertheless, some of the general principles will also apply to precision tracking in the Forward region.

Unless otherwise stated, the layout considered consists of pixels at 11.5 and 16.5 cm, and four silicon stereo superlayers (u-ϕ, v-ϕ, u-ϕ, v-ϕ) at 30, 40, 50 and 60 cm. The pixel resolutions are taken as 14 and 87 μm in $r\phi$ and z respectively, while each silicon strip detector has an intrinsic resolution of 20 μm and a stereo angle of 40 mrad. The systematic error used is 10 μm. When a beam constraint is used, the position of the beam is assumed to be known with a precision of 15 μm. More details can be found in Ref. [12].

No consideration is given here to:

- *Effects of non-uniform magnetic field:* Fringing effects of the magnetic field arise from the design of the short solenoid. The effect in the Barrel is small.
- *Electron recovery after bremsstrahlung:* Electrons which undergo radiation of hard bremsstrahlung in material can be reconstructed, but this requires the use of the TRT and calorimeter, and so is beyond the scope of the SCT.
- K^0_s *reconstruction:* The reconstruction of V^0's relies heavily on the use of the TRT.

10.2 Basic features

10.2.1 Space-point resolution

For a stereo superlayer (u-ϕ or v-ϕ) where each silicon strip detector has an intrinsic resolution of σ_1 and a stereo angle between the strips of $\alpha \ll 1$ rad, the coordinate error in the superlayer is described by an error ellipse whose major and minor axes are $\sigma_\perp = \frac{\sigma_1}{\sqrt{2}\cos\alpha/2} \approx \sigma_1/\sqrt{2}$ and $\sigma_\parallel = \frac{\sigma_1}{\sqrt{2}\sin\alpha/2} \approx \sqrt{2}\sigma_1/\alpha$ and which is rotated with respect to the ϕ-axis by $\alpha/2$.

The effective $r\phi$ and z errors cannot be given in a simple form and the correct treatment involves using the complete error matrix in a helix fit. However, in the presence of other detectors which do measure ϕ, the effective errors become $\sigma_1/\sqrt{2}$ and σ_1/α respectively. For the current parameters of the stereo layers, the intrinsic resolutions are 14 μm and 500 μm respectively.

10.2.2 Parameter measurements

Momentum resolution The momentum resolution is described by $\sigma(1/p_T) = A_p \oplus B_p/p_T$ (at $\eta = 0$), where the first term describes the sagitta measurement, while the second results from multiple scattering. Values for the coefficients A_p and B_p are shown in Table 29. These numbers yield fractional resolutions $\sigma(p_T)/p_T$ at 500 GeV/c of 46% and 30% respectively. These numbers do not allow for the degradation coming from pattern-recognition and inefficiencies.

Table 29: Coefficients for momentum resolution with p_T measured in GeV/c.

	A_p	B_p
No beam-constraint	$0.92 \cdot 10^{-3}$	0.012
With beam-constraint	$0.59 \cdot 10^{-3}$	0.011

Angular resolution In the barrel, the dip angle λ is only measured by the SCT and the multiple scattering contribution is small compared to the measurement errors. The resolution is 0.9 mrad, which increases to 3.5 mrad in the absence of pixels. For high momentum tracks, the measurement of azimuthal track direction has a resolution of 0.20 mrad (no beam-constraint).

Impact parameter resolution The impact parameter d_0 is defined as the distance of closest approach of a reconstructed track to the beam-spot in the transverse plane. The parameter z_0 is defined as the z position of the track at the same point. The resolutions of these parameters in the Barrel in the presence of multiple-scattering (allowing for detectors and support structure) are described approximately by Equations 12 and 13.

$$\sigma(d_0) = A_d \oplus \frac{B_d}{p_T} \frac{1}{\sqrt{\cos\lambda}} \tag{12}$$

$$\sigma(z_0) = A_z \oplus \frac{B_z}{p_T} \frac{1}{\sqrt{\cos^3\lambda}} \tag{13}$$

Values for the coefficients A_d, B_d, A_z and B_z are given in Table 30 for four situations: (a) a default layout, corresponding to high luminosity, (b) the default layout with a layer of crossed-strips at 6 cm giving resolutions of 10 μm by 20 μm, (c) the default layout with a third (binary) pixel layer at 4 cm and (d) the default layout, with the pixel layers removed (just stereo layers). Layouts (b) and (c) incorporate vertexing layers which would be used for low luminosity physics. More details on layout (b) can be found in Ref. [103]. The resolutions as a function of rapidity over the complete range are shown in Figure 99. These include the use of the MSGCs in the Forward region.

Table 30: Coefficients for impact parameter resolution with p_T measured in GeV/c. See text for explanation of layouts.

Layout	A_d (μm)	B_d (μm)
(a) Default	27	220
(b) Default + crossed-strips	13	62
(c) Default + pixels	18	61
(d) Just stereo	205	
	A_z (μm)	B_z (μm)
(a) Default	130	240
(b) Default + crossed-strips	39	90
(c) Default + pixels	84	130
(d) Just stereo	1600	

10.2.3 Hits in detectors

Occupancy

Occupancies at the strip level from pile-up events depend on the length of strips and their spacing. In Table 31 information is given for the SCT strip detectors. The *logical*[2] wafers are approximately 6×12 cm^2 and 2.5×10 cm^2 for the silicon barrels and GaAs disks respectively. The numbers given include magnetic field effects, ejection of delta rays from the substrate and pair conversions. The numbers for the barrels are averaged, but since the wafers at the highest rapidity in a given barrel subtend a smaller rapidity interval (and the distribution of tracks is fairly flat in rapidity), the number of clusters per wafer is peaked towards $|\eta| = 0$. Clusters in stereo superlayers are only counted once. The occupancies given are derived by dividing the number of clusters in a wafer for 20 pile-up events by the number of strips. In practice, the charge deposited by an ionising particle will give signals on more than one adjacent strips due to diffusion, angular effects, creation of delta rays and capacitive coupling. This group of strips is referred to as a *cluster* and from it a *hit* can be constructed for use by the pattern-recognition. Due to these effects, the actual number of fired strips is greater than the the number of particles crossing the detectors. This factor (the *cluster width*) has not been included in the numbers for the occupancy because it overlaps with the definition of the two-hit separation required to resolve clusters. In the subsequent work, it is assumed that only tracks separated by at least three strips can be resolved in the silicon detectors. This means that the fraction of strips 'used' is approximately six times the occupancy given in the table.

Table 31: Occupancies of SCT strip detectors. Each particle crossing a detector typically produces one *cluster* of fired strips. *c/w* is the number of clusters per wafer for one minimum bias event; $\times 20$ indicates the clusters per wafer with 20 pile-up events; *occ.* is the corresponding occupancy, determined as clusters per strip; in brackets is the result of a simpler calculation; *width* is the width of a cluster in terms of the number of strips fired. See text for important comment concerning fraction of 'used' strips.

r (cm)	c/w	$\times 20$	occ. (%)	width
30	0.115	2.3	0.30 (0.18)	2.9
40	0.075	1.5	0.20 (0.11)	3.2
50	0.046	0.9	0.12 (0.07)	3.4
60	0.035	0.7	0.09 (0.05)	3.7
GaAs			(0.12)	

In addition, occupancies corresponding to a simpler model are also given in brackets in Table 31. This model, used extensively for the studies of section 10.3, does not include 'material effects'. Delta rays and asymmetric photon conversions can give rise to low energy electrons which spiral axially along the wafers in the Barrel giving rise to many correlated hits. This tends to make pattern-recognition very difficult locally, but seen over the length of a moderate p_T track, the effect probably manifests itself more like an efficiency loss rather than many hits distributed randomly. Due to the robustness of the SCT, this efficiency loss is probably less problematic than the corresponding random distribution of hits would be.

Ghosts

The number of ghost hits which can be formed in a stereo (u-ϕ) layer depends on the proximity of tracks and the stereo angle. If the hits on a wafer are independent, then the fraction of ghosts is well described by $F_\alpha(N_{part} - 1)$ (for small N_{part}) with $F_\alpha \equiv \frac{\alpha L}{W}$, where N_{part} is the number of particles crossing the wafer, L and W are the wafer dimensions and α is the stereo angle. This is illustrated by Figure 100 for a GaAs wafer.

Ghost hits are particularly prolific in the core of jets. Further, things are complicated by the tendency for ghost hits to be correlated from one layer to another, enhancing the probability of finding fake tracks. For a b-jet in the absence of pileup, for every hit in the 30 cm layer coming from a particle in the jet, there are 0.4 hits additional ghost hits on the same logical wafer. This number falls like r^{-1}. While this is not a

[2] A *logical* wafer may consist of two pieces of silicon or GaAs bonded together.

big problem for the outer stereo layers, it could be for a 6 cm vertexing layer. Even with a 40 mrad stereo angle, there are five ghost hits per real hit.

Two-hit resolution

The ability to resolve two nearby hits on a strip detector depends on their separation, the strip pitch and the readout scheme used. If the hits on a wafer are independent, then the efficiency for hits to be resolved by ν strips in a stereo detector with N_{strip} strips is approximately $1-4\frac{\nu}{N_{strip}}(1-\frac{\nu}{F_\alpha N_{strip}})(N_{part}-1)$ where the quantities were defined in the previous section. One factor of two comes from the fact that there are two sets of strips in the superlayer (for just ϕ-strips, replace $2(1-\frac{\nu}{F_\alpha N_{strip}})$ by 1), while the other comes from the band required to have a good measurement of one ordinate extends $\pm\nu \times pitch$ either side. When the N_{part} particles are considered to come from pile-up, $(N_{part}-1)/N_{strip}$ can be replaced by the strip occupancy. This is illustrated by Figure 101 for a GaAs wafer.

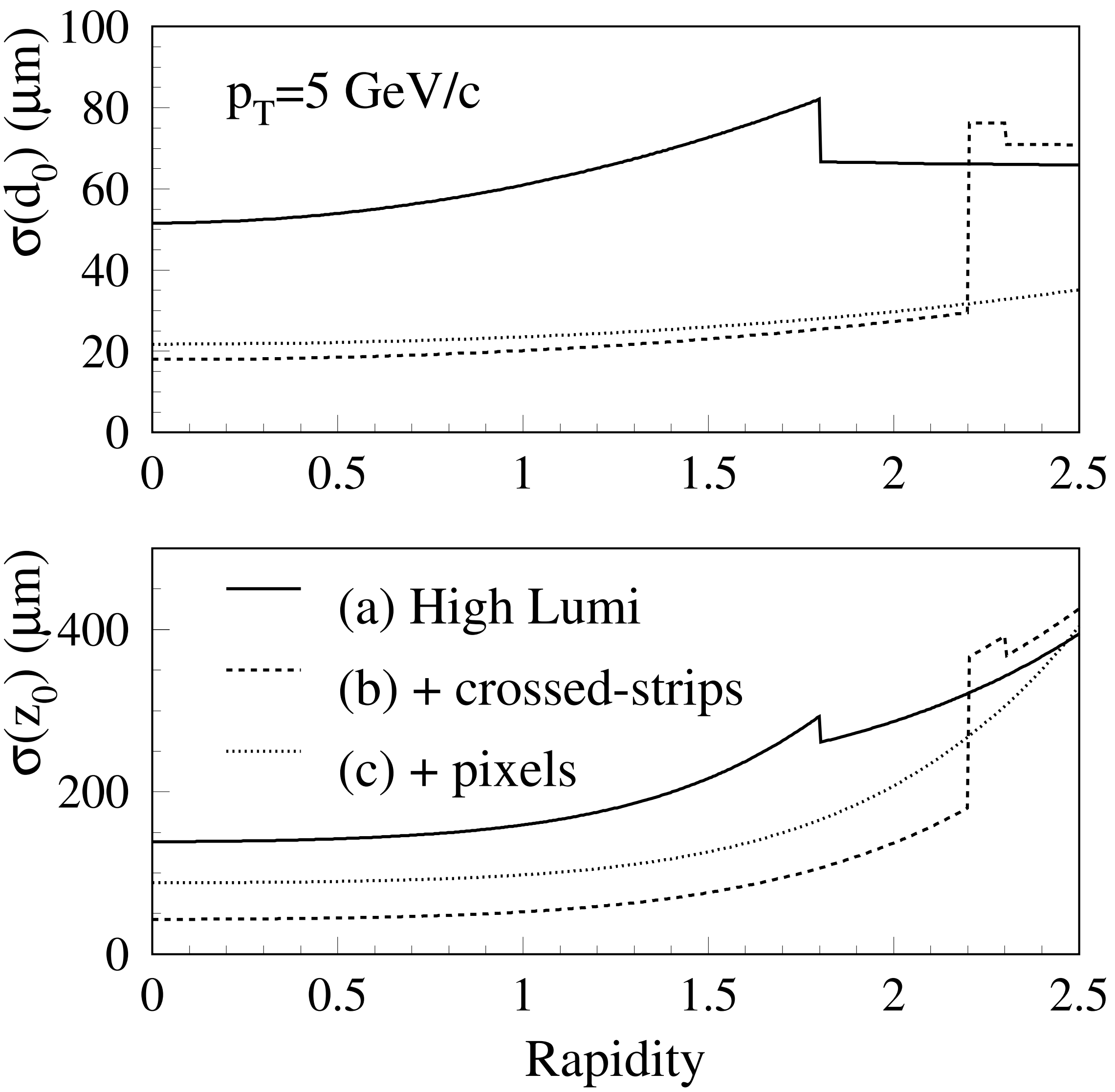


Figure 99: Impact parameter resolutions in both r-ϕ and r-z for a 5 GeV/c particle. The different layouts are described in the text.

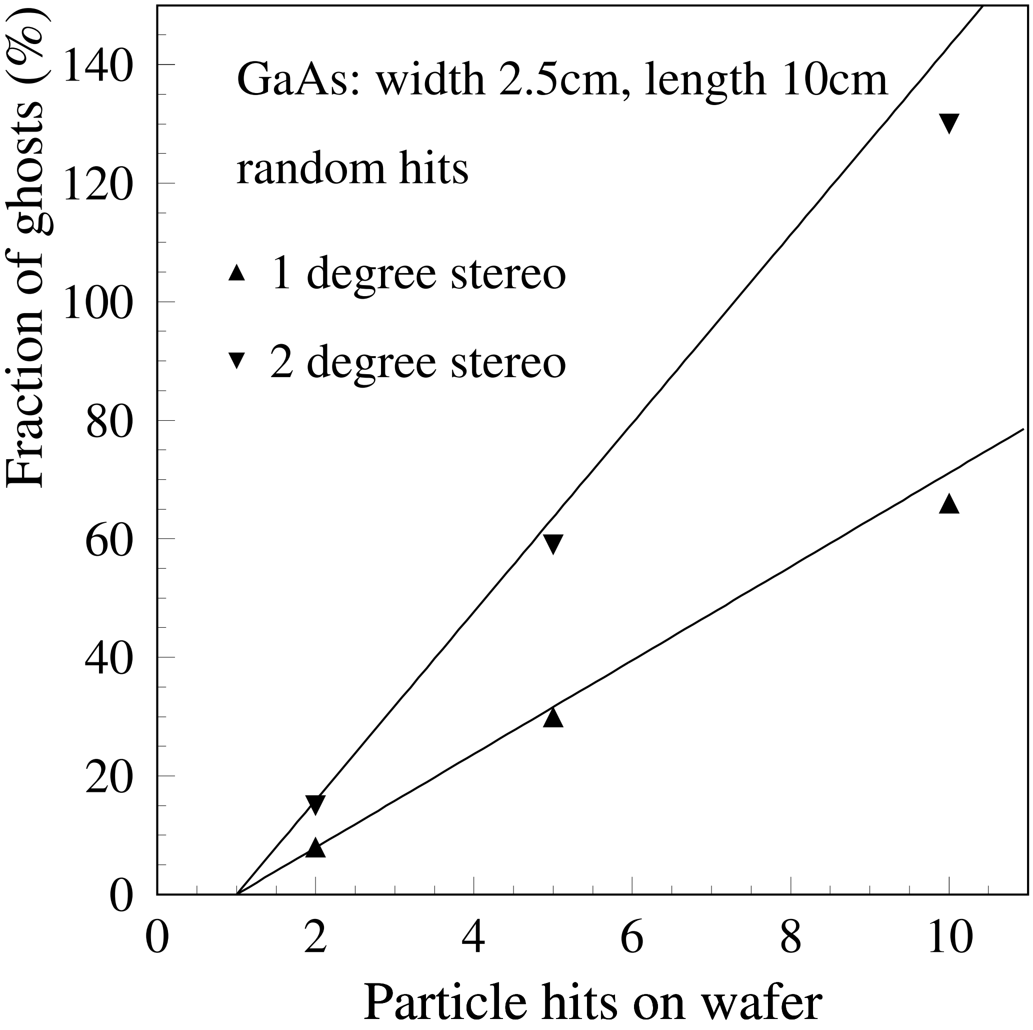


Figure 100: Fraction of ghosts as a function of number of particle hits N_{part} on a GaAs wafer. Stereo angle α is measured between u and v strips. The lines correspond to $F_\alpha(N_{part}-1)$ - see text.

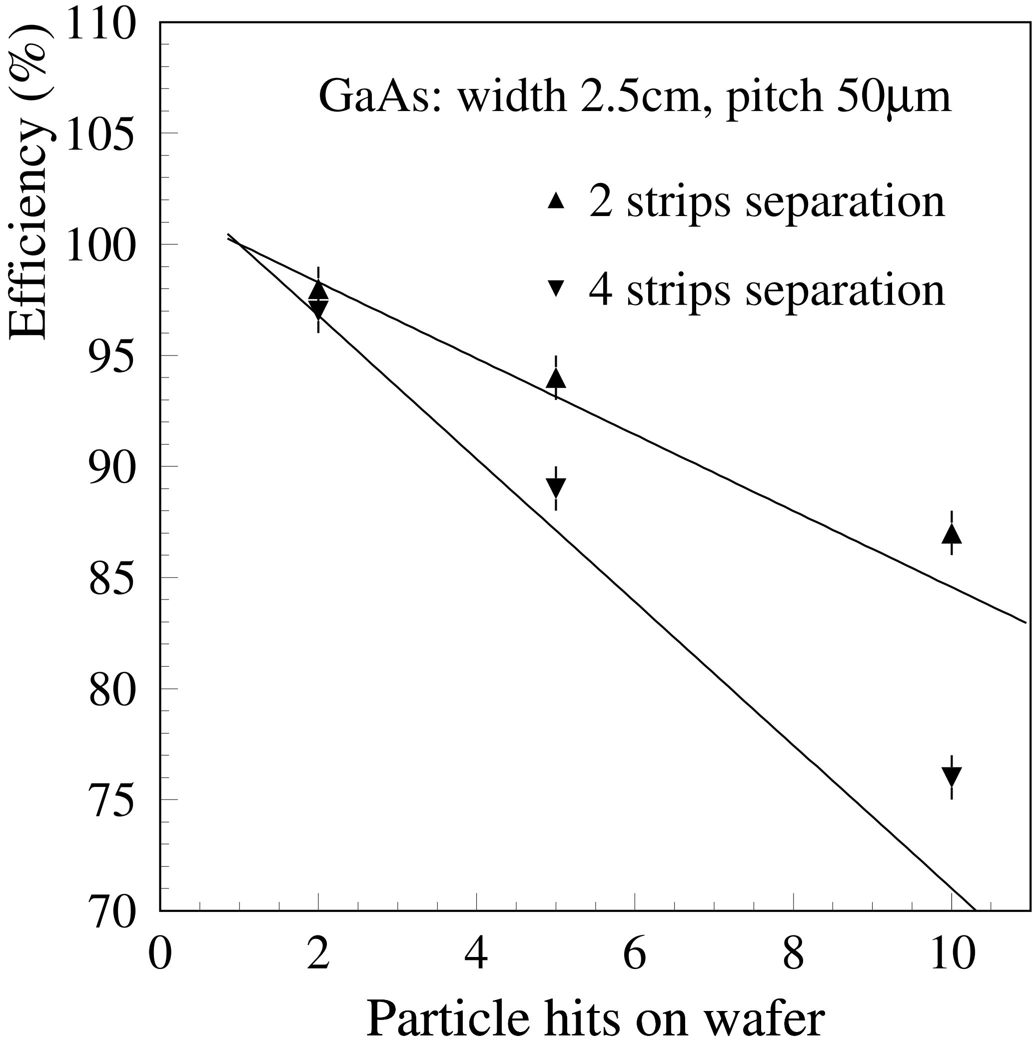


Figure 101: Efficiency for hit reconstruction as a function of number of particle hits N_{part} on a GaAs wafer. Data is shown for two different resolutions, measured in number of strips ν. The lines correspond to $1 - 4\frac{\nu}{N_{strip}}(1 - \frac{\nu}{F_\alpha N_{strip}})(N_{part} - 1)$ - see text.

As usual, the problem is most severe in dense jet environments. Figure 102 shows the separation between a hit from a particle from a b-jet (produced in a $t\bar{t}$ event) and the nearest neighbour in $r\phi$ on the same wafer in the $r = 30$ cm layer. Only 5% of the time is a hit the only one on the wafer.

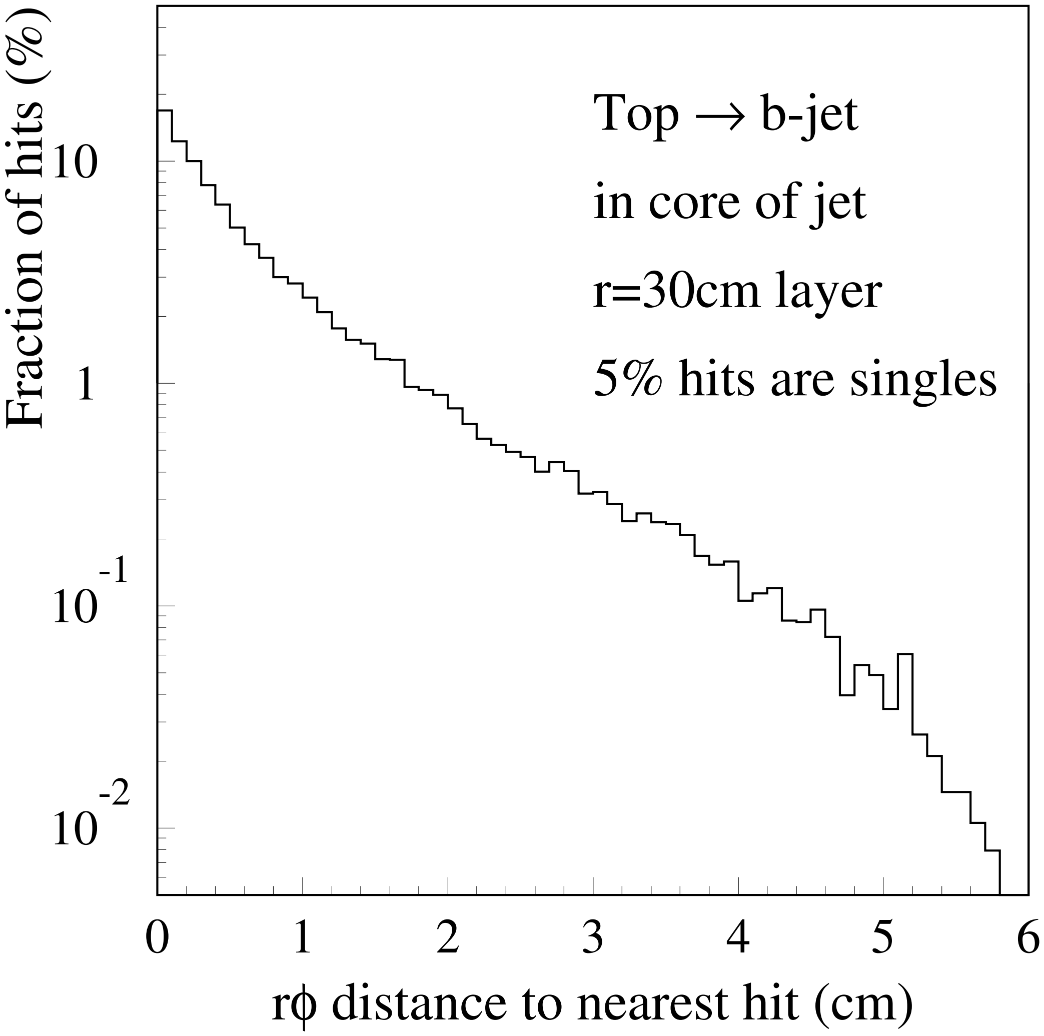


Figure 102: $r\phi$ separation between a hit and its nearest neighbour on the same 6×12 cm^2 wafer.

Distributions such as that in Figure 102 can be used to determine the distances as a function of radius which contain 1, 5 and 10% of hits. These distributions are shown in Figure 103. Using an estimate of the two-hit separation required to resolve hits, an estimate can be made of the hit loss as a function of radius. The figure shows the separations corresponding to three readout strips. The nominal 20 μm intrinsic resolution of the silicon strips can be achieved by having readout strips spaced by $70 = \sqrt{12} \times 20$ μm if direct readout is used or approximately 1.5 times this if charge division is used. In the studies which follow, direct readout is assumed.

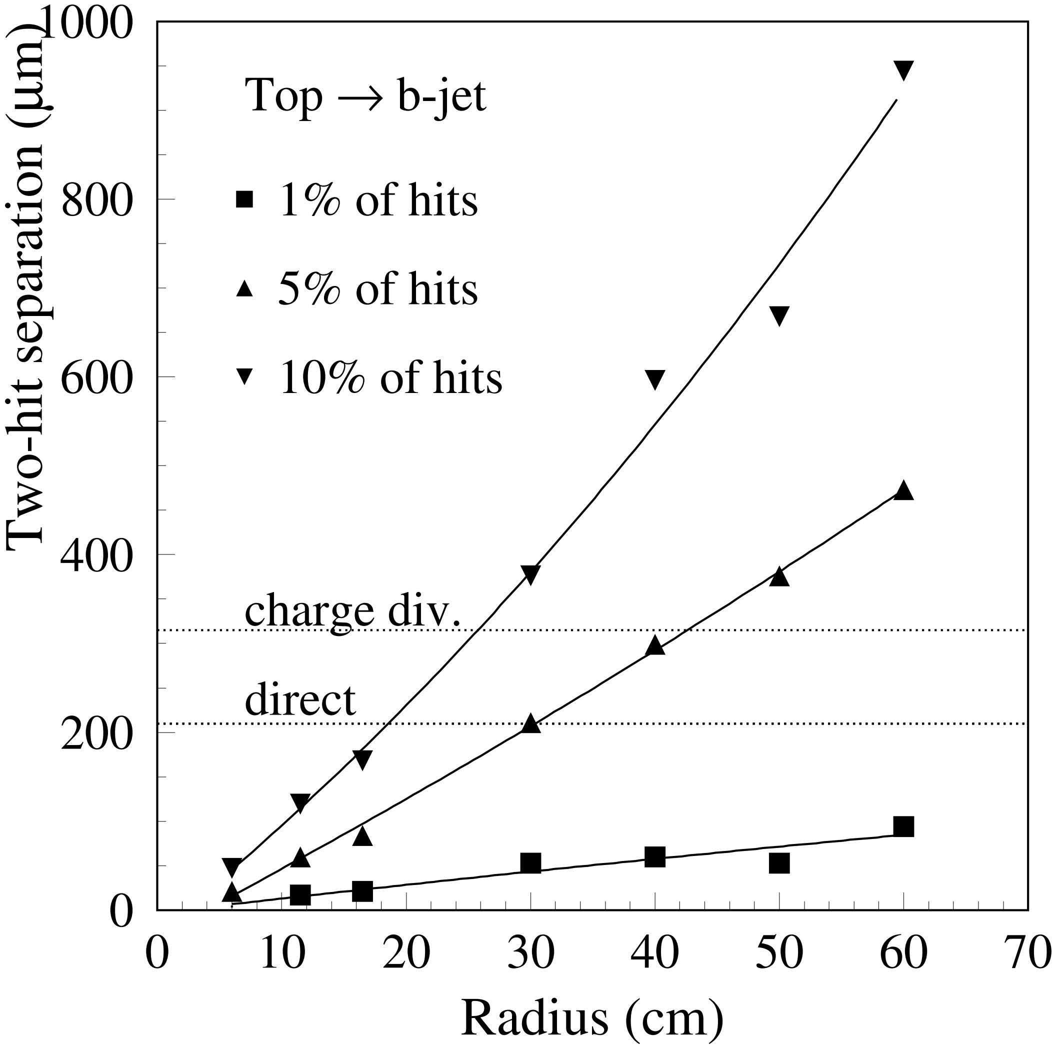


Figure 103: Two-hit separation as a function of radius with plausible separations corresponding to two different readout schemes, both designed to achieve 20 μm intrinsic resolution. For example, at 30 cm, with direct readout, 5% of hits are not resolved.

Within a b-jet, there are correlations between the separations of hits on subsequent superlayers. However, this is reduced by the divergence of the tracks and the dispersion produced by the magnetic field. It is found that approximately 78% of tracks have four resolved space-points in the four stereo superlayers, while 15% have only three.

10.2.4 Track-finding

The track-finding efficiency which can be obtained for a given physics channel depends not only on the detector but also on the analysis used. In particular, it depends on the track-finding algorithm, the values of the algorithm parameters which are used and the definition of efficiency. Parameters should be tuned to obtain a compromise between high efficiency and low numbers of fakes. In principle, every time the layout or detector parameters are modified, the algorithm parameters should be retuned. Due to the changes in layout and the evolution of algorithms, little work has been done with the full ATLAS simulation (DICE) and corresponding reconstruction software.

Tau decays The most complete study performed so far with DICE and using the IPATREC algorithm [108] has been made for tau decays using the *Panel layout.*[3] Tau decays are important for investigating the MSSM Higgs sector, and the three-prong decays are particularly challenging. The mean energy of the tau's is 150 GeV, giving rise to a mean track separation of 4 mm at a radius of 60 cm. Single three-prong tau decays have been generated in the absence of underlying events or pile-up and in a p_T range of 40 to 300 GeV/c. In 90% of cases, exactly three tracks were reconstructed, with only two tracks in 7% of cases. The number of events reconstructed with ghost tracks was less than 1%, as was the number of events where the charge of at least one track was mismeasured.

A second approach

For the systematic studies presented in the next section 10.3, a simpler and faster approach has been taken, using standalone code with a uniform magnetic field. The precision tracking is simulated as a set of planes tessellated by detector modules. These have by default the nominal resolutions; random inefficiencies corresponding to the nominal geometric gaps; noise at the level of one per mille on strip detectors only. Multiple scattering is included, but there is no bremsstrahlung, delta rays or nuclear effects. Hits are required to be separated by three strips in order to be resolved; hits closer than this are flagged and not used for the pattern-recognition.

Reconstructed tracks are required to have at least four $r\phi$ measurements and at least four z measurements, and are allowed up to two missing hits. A *good* track is defined as having at least eight correct measurements (in $r\phi$ or z) and no more than two incorrect ones. Tracks are searched for in a cone of $|\Delta\phi| < 0.30$ rad and $|\Delta\theta| < 0.15$ rad around a *trigger* defined by Monte Carlo truth information. A 'first come, first served' algorithm is used and as soon as a plausible track candidate is obtained, the corresponding hits are flagged as 'used'. No attempt is made to resolve possible conflicts over the association of individual coordinates or to retain a large number of possible candidates, possibly sharing coordinates. One consequence of this is that one track candidate may incorporate a hit which comes from another track. Because the hit has been stolen, it may be difficult to complete the pattern-recognition for the second track.

Efficiency is defined by the ratio of *interesting*[4] tracks found to the total number of *interesting* tracks which could be found in the cone. The *fake rate* is defined as the ratio of bad tracks to good tracks, and is given for $p_T > 2$ GeV/c. This algorithm has the advantage of being fairly robust to changes in layout or detector parameters, which facilitates systematic studies. At some level, the efficiency and fake rate are complementary: tracks which are lost often are spoilt and appear as fakes.

The software described above has been used to study a) isolated leptons from the decay $Z \rightarrow \mu\mu$ and b) particles in b-jets arising from the decay of $t\bar{t}$ systems. In both cases, complete events, including the appropriate underlying event, are generated with PYTHIA. m_t is taken as 160 GeV/c^2. Generally pile-up has not been included. All errors are statistical, corresponding to the analysis of a finite number of events, and no estimate of systematic errors is made. However, the systematic uncertainties arising from variations of certain key parameters can be understood from sections 10.3.1 to 10.3.3.

[3] The *Panel layout* was the forerunner to the current layout with one pixel layer and five sets of silicon crossed-strips.

[4] See next paragraph for definition.

Isolated leptons The reconstruction of isolated leptons is important for alignment considerations and the physics provided by decays such as $H \to ZZ \to 4\ell$. The efficiency for reconstruction of the leptons is $99.6 \pm 0.3\%$ with a fake rate in the cone of $0.5 \pm 0.3\%$. Low momentum tracks in the cone are reconstructed with an efficiency of $> 95\%$. With pile-up at luminosities of 10^{34} cm^{-2}s^{-1}, the efficiency falls to $99.4 \pm 0.4\%$ with a fake rate of $1.7 \pm 0.5\%$.

b-jets from $t\bar{t}$ High momentum jets provide a good test of pattern-recognition because of the problems arising from the presence of ghost hits (formed when combining strip detector hits to create space-points). The reconstruction of tracks in three dimensions is important for good association to external points, calculation of invariant masses and application of alignment corrections. The b-jets have a mean p_T of 70 GeV/c (in the absence of any trigger cuts) with charged multiplicities of about five. The mean angular separation between particles (at the decay vertex) is about 50 mrad while the separation as a function of $r\phi$ within a module was presented in section 10.2.3. The tracks from b-jets are reconstructed with an efficiency of $96.6 \pm 0.1\%$ for $p_T > 2$ GeV/c, while the fake rate is $3.2 \pm 0.1\%$. The *cumulative*[5] efficiencies and fake rates as a function of the p_T of the particle are shown in Figures 104 and 105. The efficiency is also shown as a function of the jet p_T in Figure 106.

[5] For a given p_T, the efficiencies are shown for tracks whose transverse momentum exceeds that p_T.

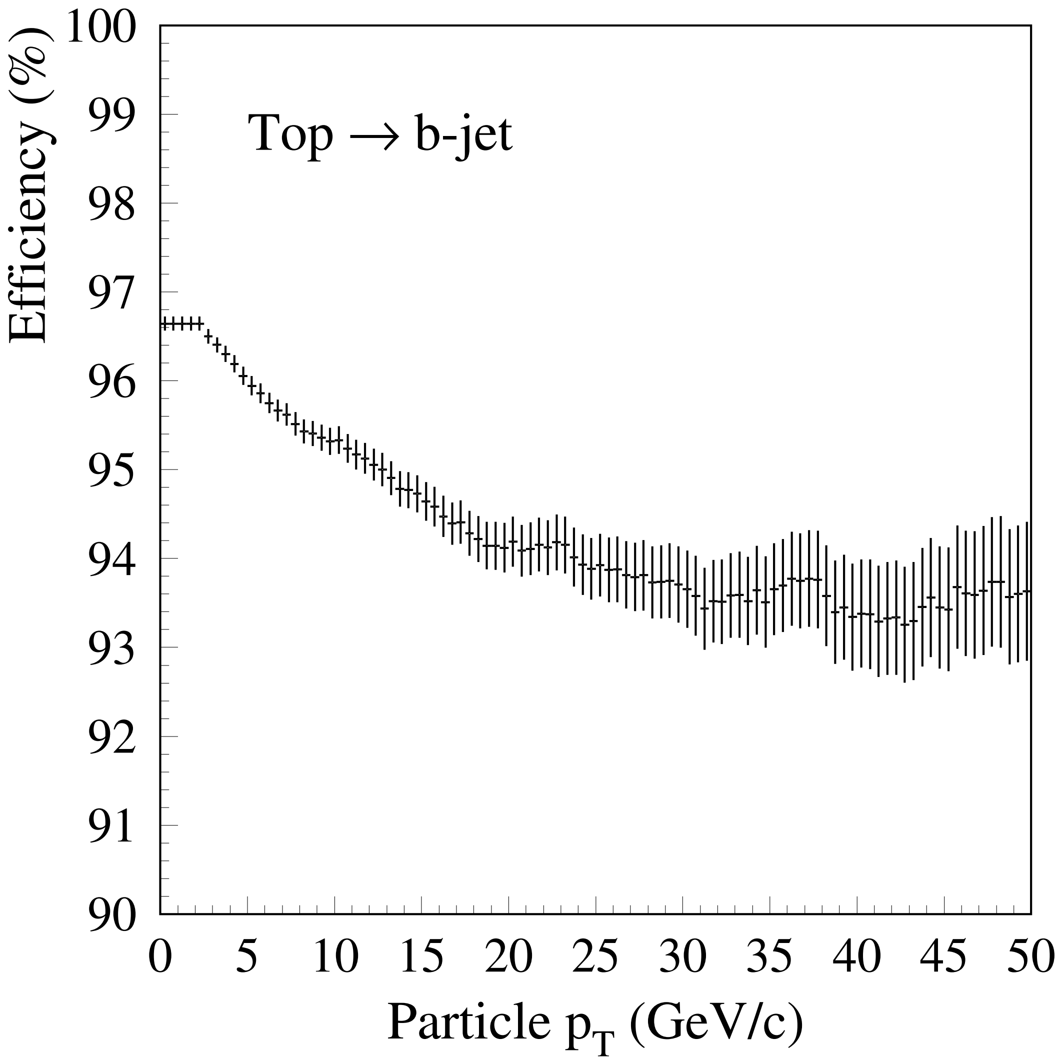


Figure 104: Track-finding efficiency for p_T in excess of the abscissa value.

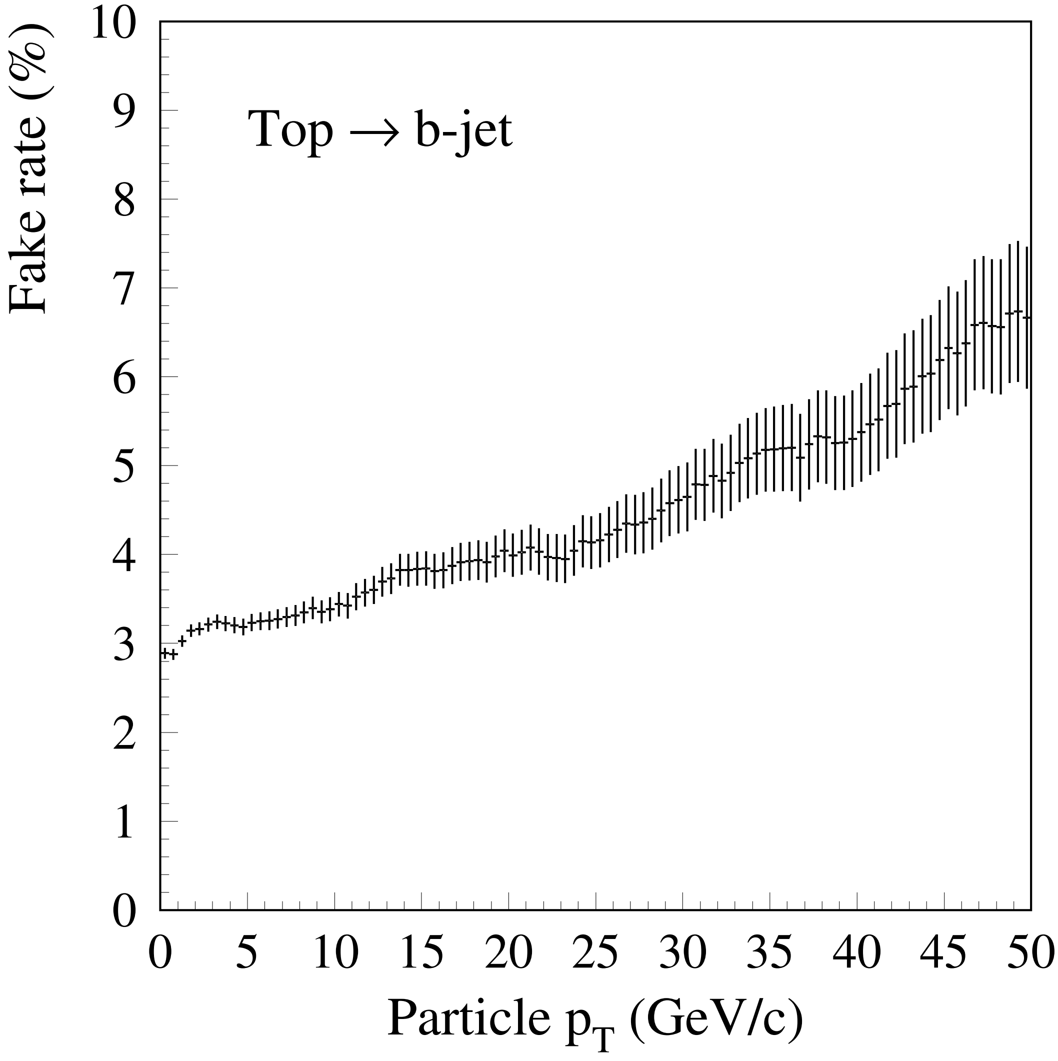


Figure 105: Fake rate for p_T in excess of the abscissa value.

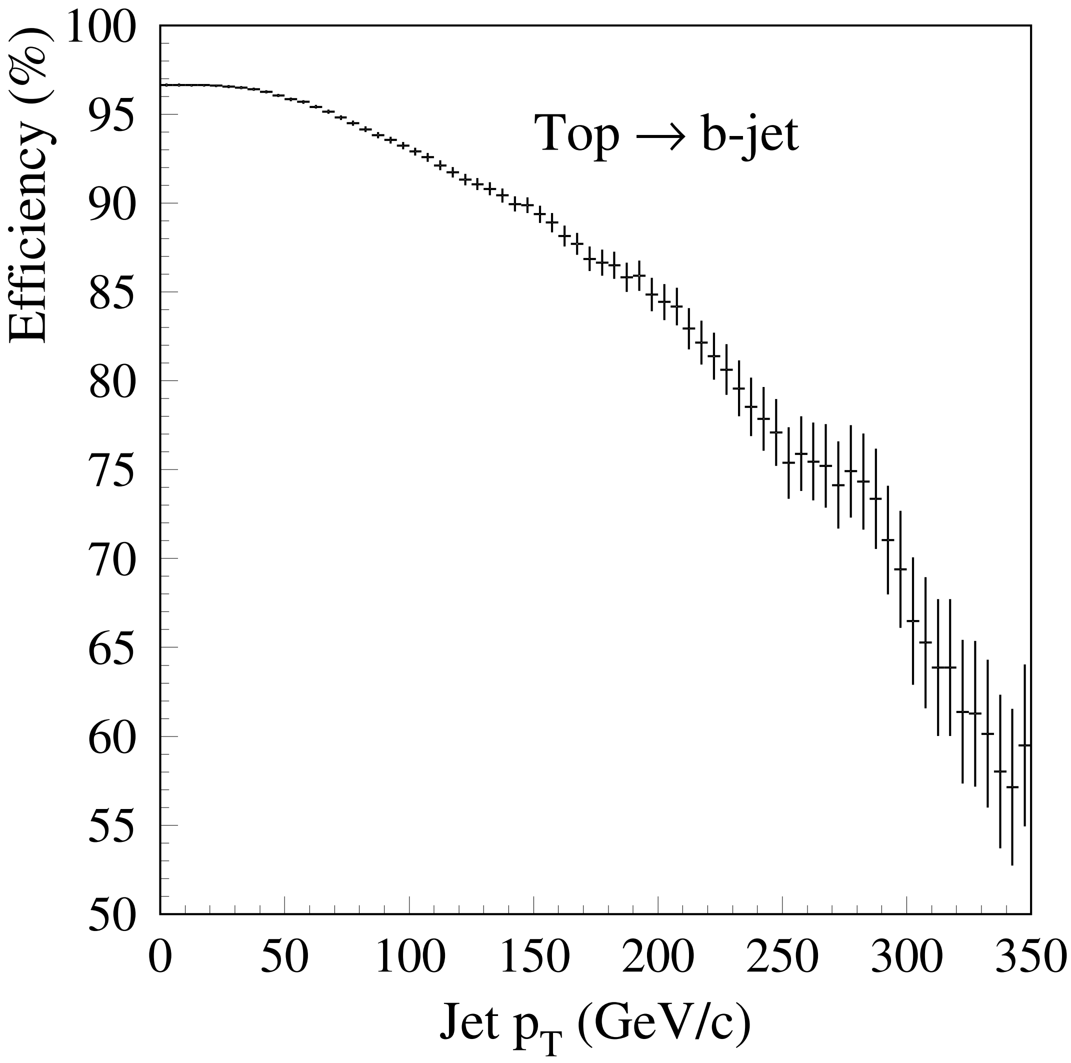


Figure 106: Track-finding efficiency for jet p_T in excess of the abscissa value.

10.2.5 b-tagging

The second approach to track-finding, described in section 10.2.4, has been used to consider the b-tagging potential of the SCT. This manifestly depends on the algorithm used, but not on the definition of the correctness of the reconstruction. It is conceivable that the definition used in this work to define a *good* track is pessimistic. Decays of the form $t \to bl\nu$ have been considered. Jets formed from light-quarks with similar kinematic distribution have been simulated by forcing the particles in b-jets to come from the interaction point.[6] Tracks were reconstructed in the cones around the jet directions and only those tracks with $p_T > 1$ GeV/c and $|d_0| < 2$ mm were used to form a discriminating function. In Figure 107, the *rejection* (defined as the reciprocal of the efficiency) against light-quark jets is shown as a function of the efficiency to select b-jets. This gives no information about the rejection of c-jets, which inevitably will dilute the b-jet purity. Curves are shown for the high luminosity and the pixel vertexing layouts of section 10.2.2. The poorer performance at high luminosity is primarily due to the absence of a vertexing layer, rather than the effect of pile-up. In the low luminosity scenario, the steep rise for small losses of efficiency between 80% and 100% is a function of the impact resolution of the detector and the b lifetime. The plateau below 70% efficiency comes from the tales in the impact parameter distribution due to pattern-recognition problems. More details can be found in Ref. [109].

10.2.6 Extrapolation errors

The extrapolation errors on a fitted track at $\eta = 0$ are shown in Figure 108. While the errors in $r\phi$ are shown with their statistical and systematic components, the implications of the systematic when extrapolating to the TRT will depend on the exact nature of the errors.

10.2.7 Material distribution

Within the Barrel, the description of the material distribution is fairly trivial. The radiation lengths and corresponding sums are given in Table 32. Within the Barrel, there is an increase with dip angle λ which varies as $\sec\lambda$, with an additional small increase of the order of 10% at the ends of the barrels coming from the cabling of the silicon strip modules. For more information, see Refs. [12] and [110].

Table 32: Radiation lengths X in the SCT at $\eta = 0$.

Detector	radius (cm)	X (%)	ΣX (%)
Beam-pipe	2.5	0.3	0.3
Pixels	11.5	1.2	1.5
Pixels	16.5	1.2	2.7
Si stereo	30.0	1.6	4.3
Si stereo	40.0	1.6	5.9
Si stereo	50.0	1.6	7.5
Si stereo	60.0	1.6	9.1

10.2.8 Level-2 trigger

The Level-2 track trigger searches in regions of interest (ROI) defined by the Level-1 calorimeter trigger for hits in the silicon strip detectors consistent with high p_T tracks. All combinations of $r\phi$ hits on the four layers within the road are made in r-ϕ. At least three hits with residuals from a straight line of $\leq 400\ \mu$m are required; the slope of the track being a measure of the p_T.

[6] Subsequent work shows that this corresponds to the results which would be obtained from real u-,d-, and s-jets after all the V^0's have been identified and removed.

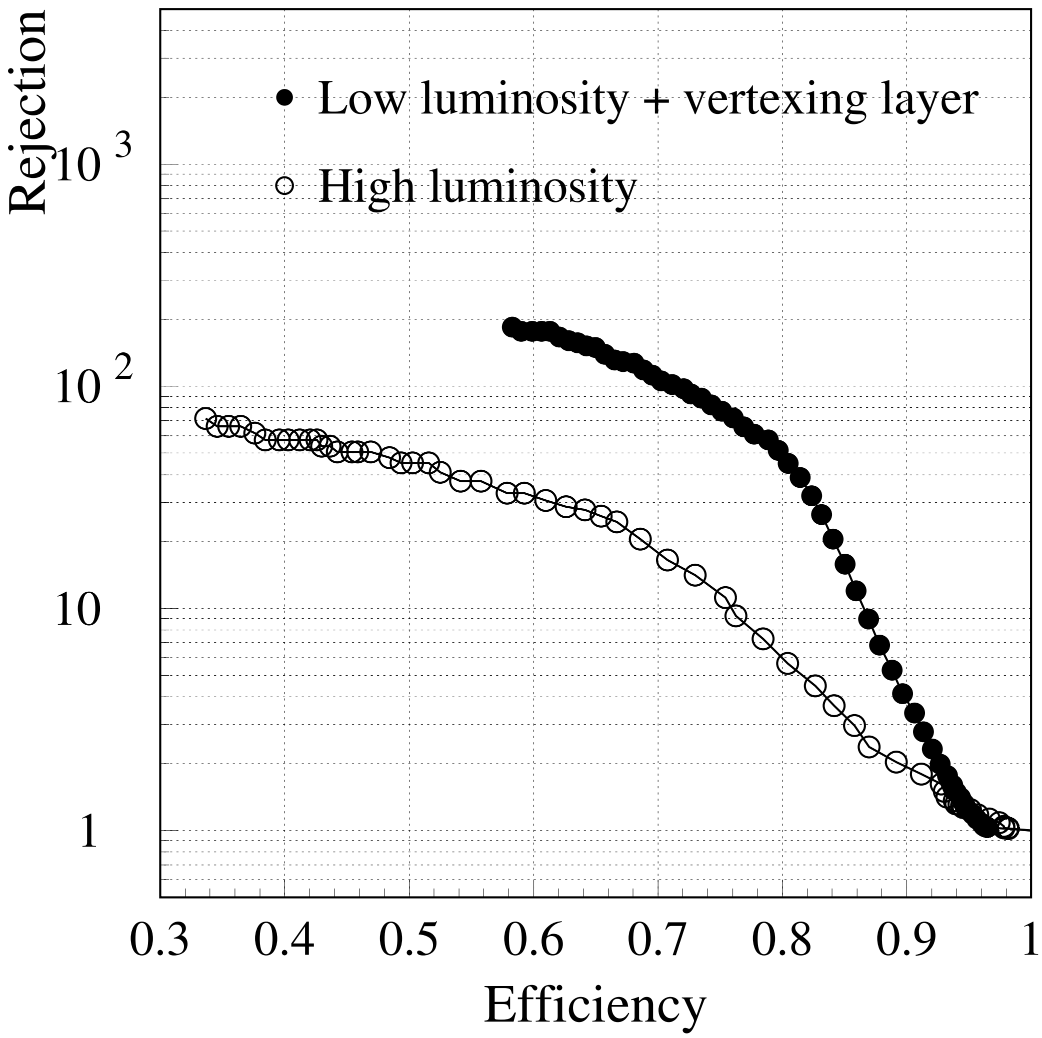


Figure 107: Rejection of light-quark jets as a function of efficiency for tagging b-jets in the Barrel. Curves show results expected for low luminosity with a pixel vertexing detector at 4 cm radius and for high luminosity with no vertexing detector.

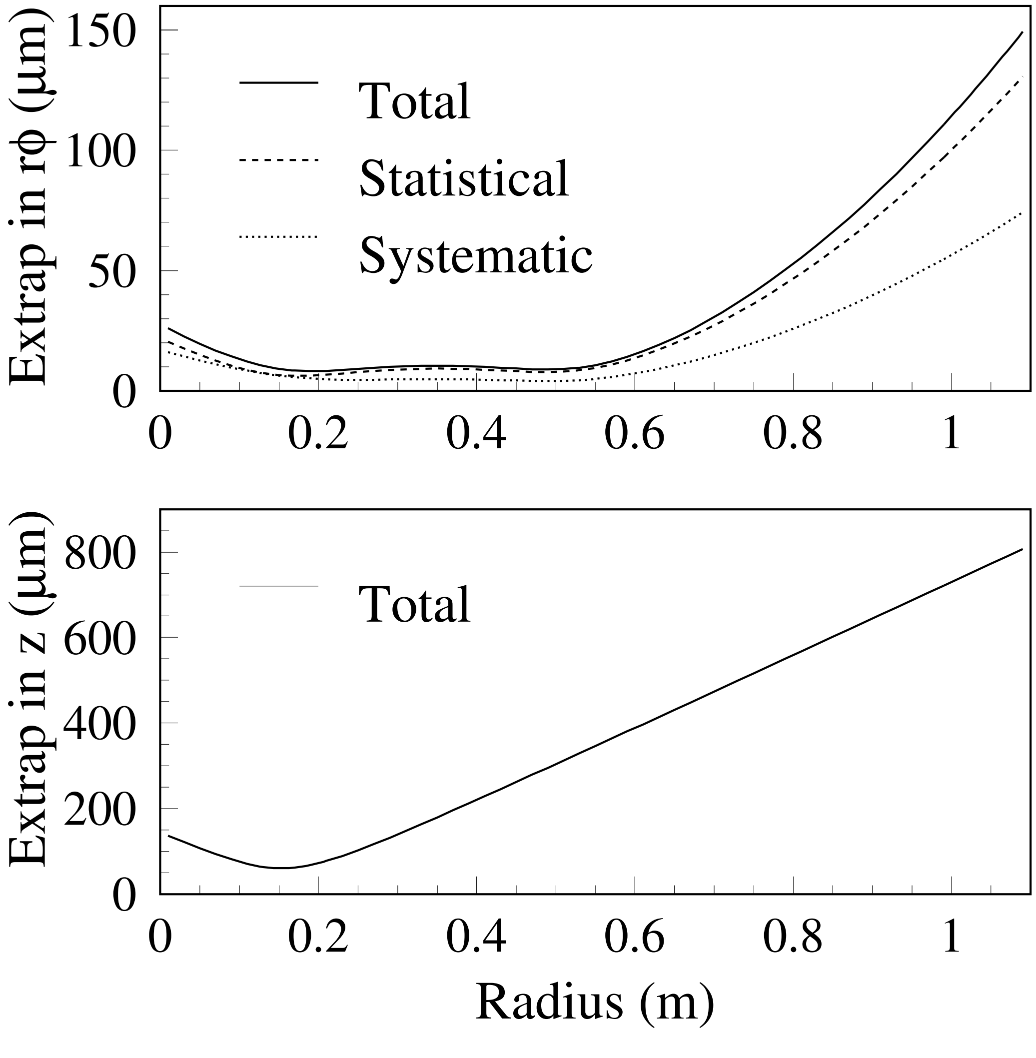


Figure 108: Extrapolation errors in $r\phi$ and z as a function of radius (for high p_T particles).

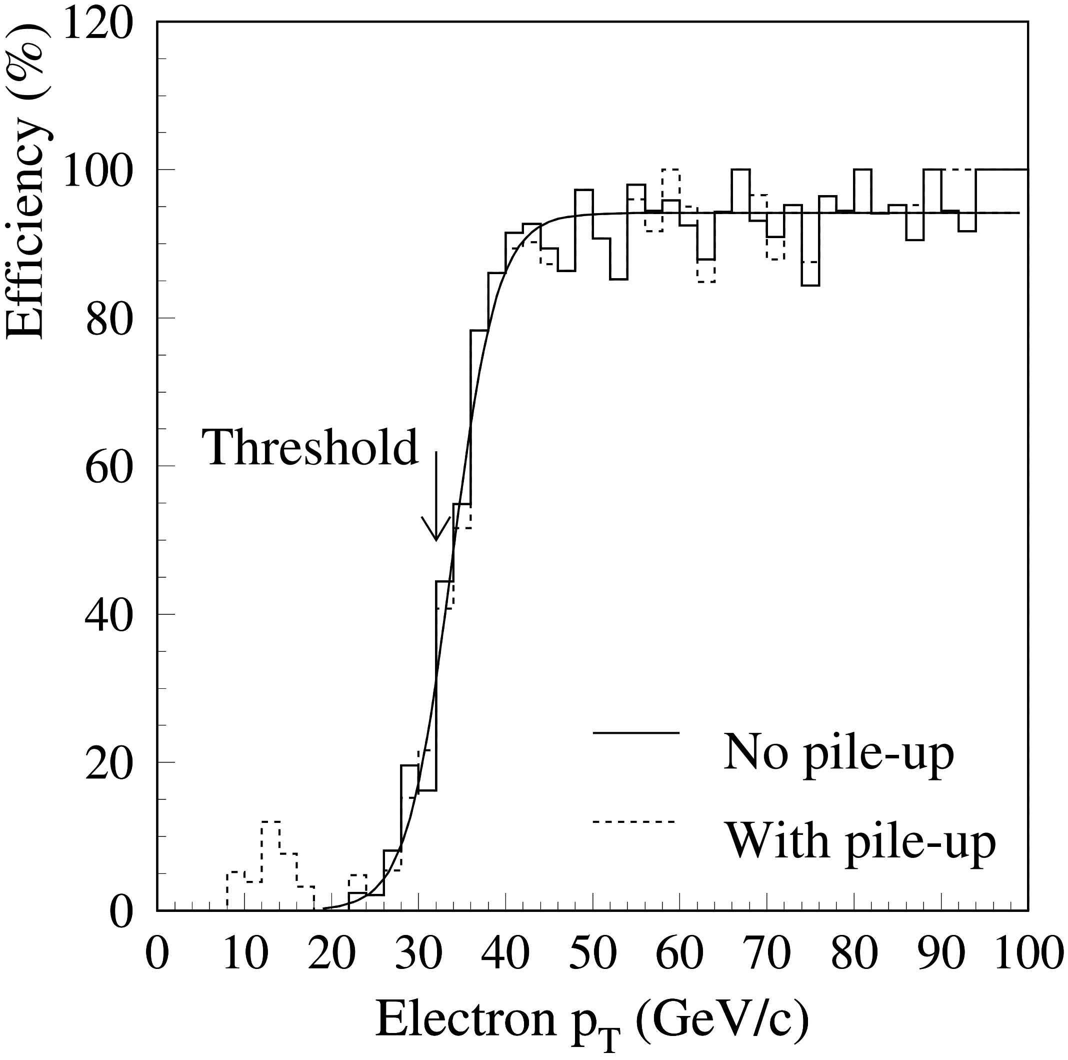


Figure 109: Trigger efficiency for electrons as a function of the true p_T. The sigmoid curve is a fit to the histogram with no pile-up.

Figure 109 shows the efficiency for accepting electrons as a function of their true p_T for a cut on the (Level-2) reconstructed p_T of 32 GeV/c. A detector inefficiency of 2% was included. Pile-up does not significantly affect the results, nor do noise hits on up to 1% of strips. The trigger is 94% efficient at the asymptotic value, and attains 90% of this value at 40 GeV/c - the desired *nominal* threshold. The corresponding plots for muons have an asymptotic efficiency of 100%, and have a sharper threshold behaviour because of the absence of bremsstrahlung. A rejection against jets passing Level-1 of 14 (7) is achieved at low (high) luminosity. Combining with the Level-2 calorimeter trigger and requiring good angular matching between the calorimeter and the tracker, rejections of 50 (37) can be obtained. With a nominal threshold of 20 GeV/c, similar efficiencies are achieved, but the jet rejection degrades by a factor of two. More details can be found in Ref. [111].

10.2.9 Lorentz effect

In the presence of an axial magnetic field, charge carriers in the silicon wafers of the Barrel will move at an angle ω with respect to the normal to the wafers due to the Lorentz effect. $\tan\omega$ is given by μB where μ is the carrier mobility and B is the magnetic field. Ionisation liberated within the bulk of the substrate of thickness t by a particle crossing at normal incidence (parallel to the electric field) will drift in a band of width $\Delta = t\tan\omega$ perpendicular to the normal. Using the mobilities given in Ref. [112] and taking $t = 300$ μm, values for ω and Δ are given in Table 33.

Table 33: Lorentz effects in 300 μm silicon wafer with a 2 T field.

Carrier	ω (deg)	Δ (μm)
e^-	15	81
h^+	5	27

Depending on the type of carrier collected, it may be possible to tilt the modules so as to reduce the spread of the charge for radial tracks (high p_T).

10.3 Systematic studies

For these studies, the second of the two approaches described in section 10.2.4 has been used. Generally, pattern-recognition parameters have not been reoptimised for different scenarios. In most cases, calculations have been performed on a set of 400 events - this tends to mean that measurements made under different conditions are correlated. Except were stated, no pile-up has been used. Efficiencies and fake rates are quoted for $p_T > 2$ GeV/c unless otherwise stated.

In this section, the first part considers the effect of parameters which are given and cannot be changed. However the *a priori* value which they have may not be obvious. The second part considers the effect of variations of the design parameters which can be chosen to achieve a required performance. While multiple scattering noticeably increases the errors on track parameters, its effect on pattern-recognition for $p_T > 2$ GeV/c seems to be small.

10.3.1 Two-hit resolution

In section 10.2.3, it was seen that the proximity of hits within a strip detector caused a non-negligible loss of reconstructed hits which could be resolved. The two-hit separation is a key parameter in the model and Figure 110 shows the variation of the efficiency. The choice which has been made is to require hits be separated by at least three strips for them to be resolved. DELPHI [113] were able to resolve hits 50% of the time when separated by 2.4 strips.

The study was performed for $p_T > 6$ GeV/c, rather than 2 GeV/c, since at lower momentum, multiple scattering dominates the $r\phi$ resolution. A strip resolution of 50 μm was used, corresponding to a pitch of about 170 μm (assuming direct readout). With the preferred resolution of 20 μm, changes in the efficiency are less apparent, since close hits which might not have been resolved with coarser detectors tend to lead to confusions and the possibility of picking up wrong hits. This highlights the inadequacy of non-global algorithms.[7]

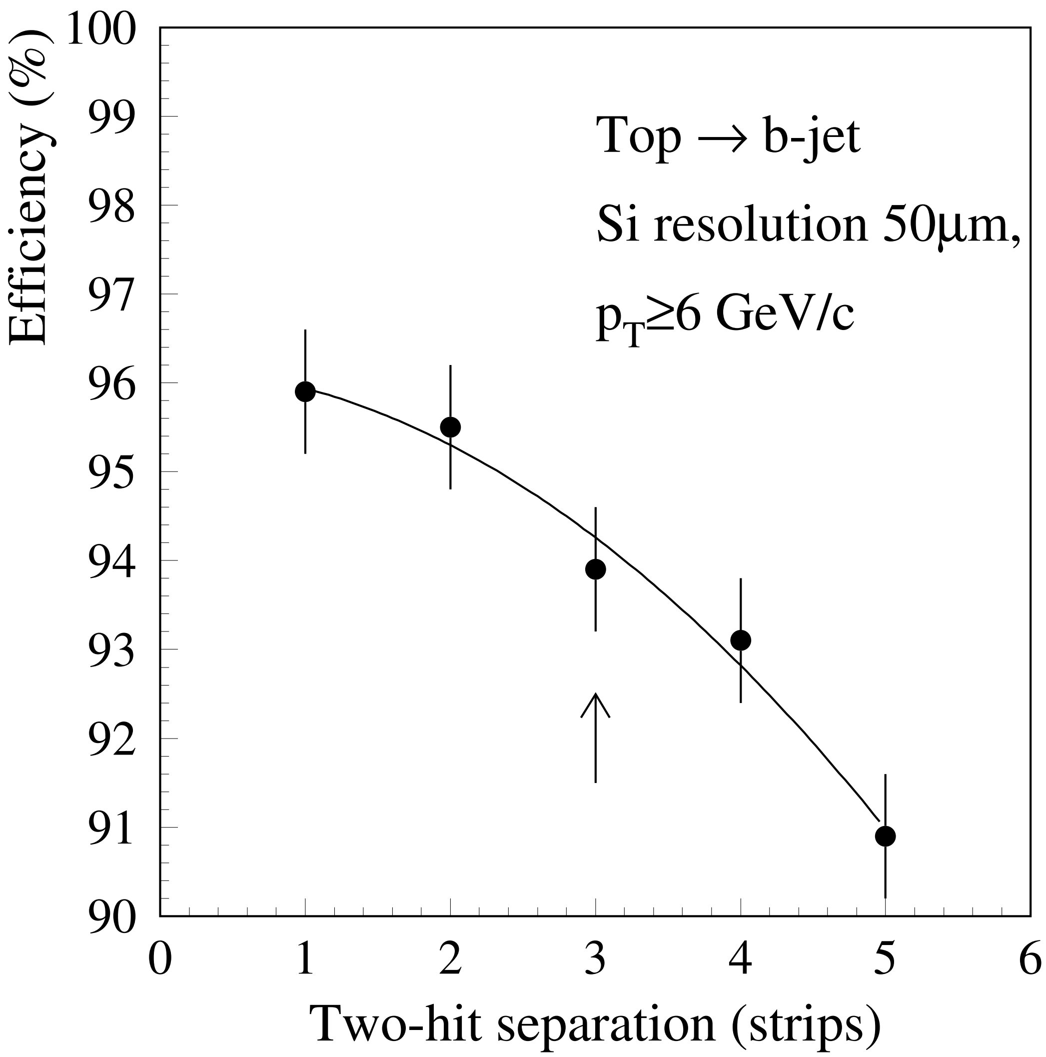


Figure 110: Efficiency as a function of two-hit separation.

10.3.2 Wafer efficiency

Figure 111 shows the track-finding efficiency as a function of the hit finding efficiency on a wafer (either silicon strip or pixel). The efficiency of each of the six superlayers in the Barrel was set to the given value, and hits were discarded randomly. The default representation is to consider an efficiency coming from geometrical acceptance of 98% for the strips (and 97% for the pixels) - strictly, this is not randomly distributed. Other sources of efficiency loss might come from the masking of noisy channels. The fact that the variation has

[7] *Global* algorithms consider 'simultaneously' all possible track combinations, rather than sequentially identifying candidates without further reference to other candidates.

a slope of less than unity arises from the robustness of having more layers than are required for a minimal track candidate.

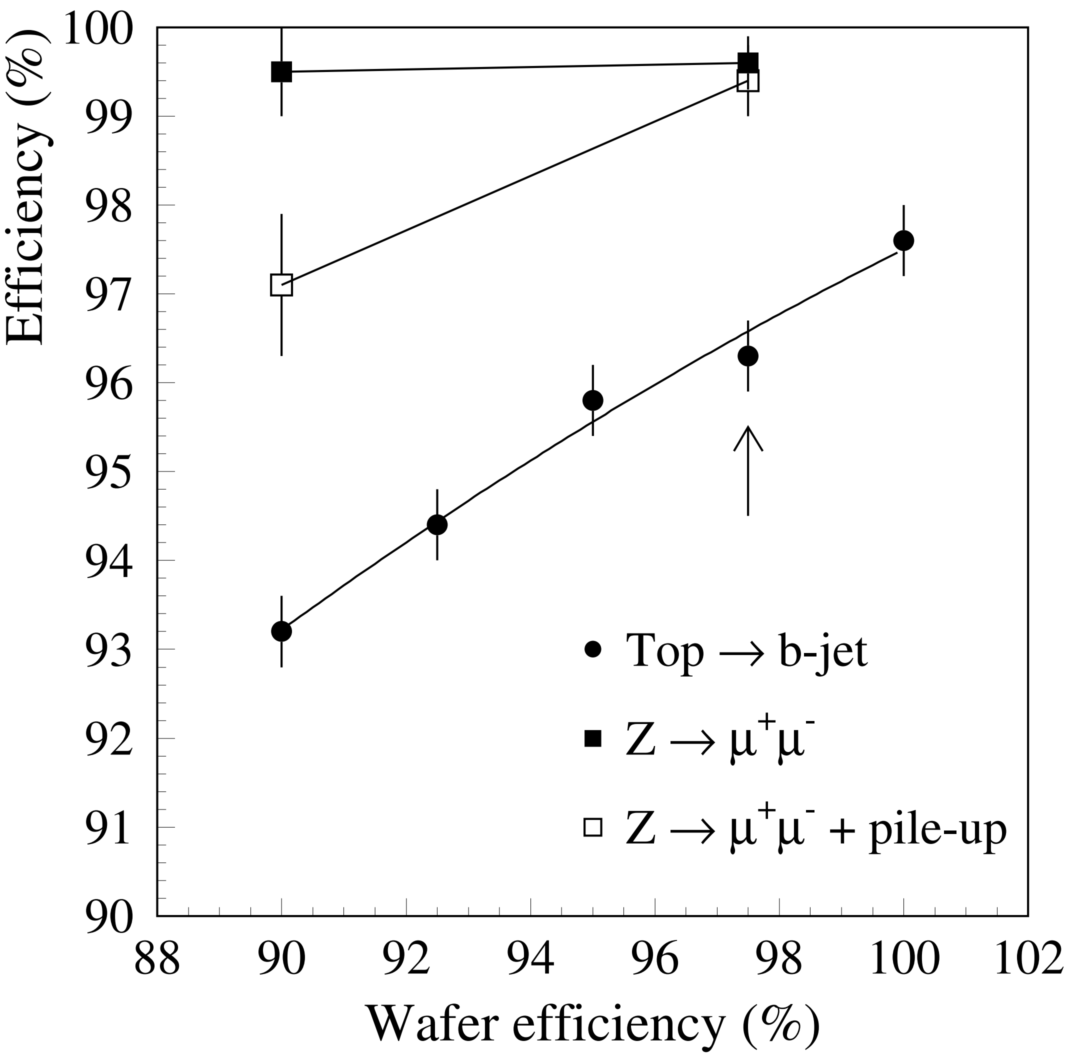


Figure 111: Efficiency as a function of hit efficiency on a wafer.

10.3.3 Noise

The efficiencies and fake rates as a function of noise are shown in Figures 112 and 113. In the first three curves, noise is only generated in the silicon strip detectors. The silicon detectors have of the order of 900 strips of 70 μm pitch. For 1% noise, noise hits are generated separately in both ϕ and $u(v)$ strips with a mean of nine hits per projection per wafer. Only combinations forming valid space-points within the module are retained.

'Back of the envelope' calculations suggest that the number of noise hits seen at the time of reconstruction (after reasonable threshold cuts) might be of the order on one per mille. This assumes that noisy channels are flagged as such. It can be seen that once the noise starts to rise above five per mille, the fake rate rapidly increases.

The rise in the fake rate associated with isolated high-p_T leptons compared to b-jets is a little exaggerated by the fact that the number of *good* tracks in the cone around the lepton (which provides the normalisation for the fake rate) is of the order of one, namely the lepton itself; whereas in the jet, most of the particles from the jet (about five) are successfully reconstructed. The ability to create fake tracks rises quadratically with the number of hits (real or noise) and so for situations where there are already many hits on a wafer, noise significantly exacerbates the problems.

It should also be noted that combinatorial track-finding algorithms take significantly longer to execute in the presence of noise - for the algorithm used here, the time taken rises as the cube of the number of hits.

10.3.4 Luminosity and occupancy

The effect of pile-up at a luminosity of 10^{34} cm^{-2}s^{-1} (corresponding to a mean of 20 additional minimum-bias events) on the track-finding for isolated high p_T leptons can be seen in Figures 112 and 113 and section 10.2.4. With one per mille noise, the differences are not huge. The (strip) occupancies at this luminosity are of the order of two per mille (0.2%) at inner radii. It is interesting to consider whether the effect of pile-up is equivalent to noise or not. Noise is modelled as being random, while pile-up leads to real particle tracks. In principle, the uninteresting tracks can be found and their hits removed from the search; however there is also the possibility to confuse these tracks with nearby tracks which come from the hard scatter.

In the figures, results (solid triangles) are shown for which the noise in the silicon (0.2, 0.4 and 0.6%) is arranged to match the particle flux through the 30 cm layer for luminosities of 1, 2 and $3 \cdot 10^{34}$ cm^{-2}s^{-1}. Correspondingly the noise in the Pixels is arranged to match the fluxes through the 11.5 cm layer. It is hard to tell whether the first triangles match the points on the curves with pile-up at low values of noise.

10.3.5 Stereo layout

Stereo superlayers are considered better than crossed-strip detectors because the number of ghost hits is reduced. The effective *cell size* for a stereo detector of length L is of the order of $L \times L\alpha$ where α is the stereo angle. For a crossed-strip detector of width W, the effective cell size is $L \times W$. The compromise made with stereo is in the z resolution. However, since only moderate precision in z is required, this is not a problem provided some vertexing layers exist at smaller radii (i.e., the pixel detectors). An alternative incorporating the small cell size of stereo and reducing the z precision can be envisaged. This is referred to as *striplet*. A striplet is modified z-strip which has a much wider pitch but a shorter length, so that several may be fitted across the width of a silicon wafer (the design is conceptual, and may not necessarily be realisable).

Table 34: Performance of different designs of the silicon strip detectors for b-jets from top.

Design	With pixels	
	Efficiency (%)	Fake rate (%)
(a) Stereo	97	3
(b) Crossed-strips	89	6
(c) Striplets	96	2
	Without pixels	
	Efficiency (%)	Fake rate (%)
(a) Stereo	86	19
(b) Crossed-strips	45	109
(c) Striplets	76	29

In table 34, comparison is made between (a) the proposed stereo design, (b) a crossed-strip design with the same number of channels (consequently $\sigma_z = 2\sigma_{r\phi}$) and (c) a striplet design, with the usual ϕ-strips and where the z-strips have a resolution of $\sigma_z = 500\mu$m and a length of 4.8 mm, again preserving the channel count. The statistical errors are of the order of a per cent. b-jets provide a very sensitive test of the ability of a design to correctly associate the $r\phi$ and z projections of tracks. The test is enhanced by the absence of pixels, which are very powerful in resolving ambiguities. Without pixels, there are only four superlayers of silicon, and there is a significant loss in track-finding efficiency due to losses at the hit level. Therefore, without pixels, the algorithm parameters described in section 10.2.4 were modified, consequently increasing the efficiencies, but also increasing the fake rates.

From the table, it can be seen that a stereo design offers the best performance for pattern-recognition. In the case of the crossed-strips in the absence of pixels, approximately half of the tracks are confused in projection, causing there to be as many fakes as there are correctly reconstructed tracks.

As already pointed out, the algorithm used for a particular design should be constructed to fully utilise the good design features and cope with the weaker ones. The algorithm used for these studies, while using a fairly general approach, was actually set up and tuned using crossed-strips, but with stereo designs anticipated. By modifying the algorithm parameters to accommodate possible inefficiencies, only one degree of freedom remains, thus providing poor rejection against fakes. If all sources of inefficiency are removed and the cuts are tightened, the stereo design remains very significantly more powerful than the crossed-strips. It is probable that the performance of the crossed-strips in the absence of pixels could be improved by a global algorithm (see footnote 7), and while the TRT cannot remove *irreducible* fakes (see next paragraph for definition), it should remove some random combinations.

The proposed design of the stereo layers has alternating u-ϕ and v-ϕ superlayers. One reason for this is the desire to avoid reconstructed ghost hits produced by nearby tracks lining up in all four superlayers and forming so called *irreducible* fake tracks. In Ref. [114] it was shown that the irreducible fake rate for a set of u-v superlayers was at the per cent level. However, it turns out that the rate of fake tracks from random coincidences exceeds this and so little difference in track-finding performance is seen between (u-ϕ, v-ϕ, u-ϕ, v-ϕ), (u-ϕ, u-ϕ, u-ϕ, u-ϕ) and (u-v, u-v, u-v, u-v).[8]

10.3.6 Stereo angle

The choice of the stereo angle α is a compromise between large angles to obtain good z and dip angle resolution and small angles to reduce the ghost rate and facilitate construction of modules with small inactive regions.

The behaviour of the space-point resolution as a function of α was given in section 10.2.1. In the absence of pixels, the z resolution is effectively $\sigma_{\|}$, which for $\alpha = 40$ mrad and with an effective strip resolution of $20 \oplus 10 = 22$ μm gives $\sigma_z = 780$ μm - better than the 1 mm precision advocated in Ref. [115] for K^0_s reconstruction. The resolutions of the dip angle λ and the parameter z_0 (defined in section 10.2.2) are shown in Fig. 114. More information can also be found in Ref. [116].

The track-finding has been studied using the b-jets. With the Pixels, the fake rate is fairly constant for α varying between 0 and 100 mrad, while the track-finding efficiency shows a small improvement of the order of one per cent. The Pixels clearly dominate the pattern-recognition and the small improvement in efficiency comes from better tracking in the r-z plane during the search phase of the track-finding. Once again, a more powerful test of the pattern-recognition capability of the silicon strips is provided by studying the reconstruction in the absence of pixels. Again, the efficiency shows a similar small improvement, but there is a significant rise in the fake rate as α increases. This is shown in Figure 115.

[8] For the u-v design, the angle between the u and v strips is the same as the angle between the u and ϕ or v and ϕ in the other designs, namely 40 mrad.

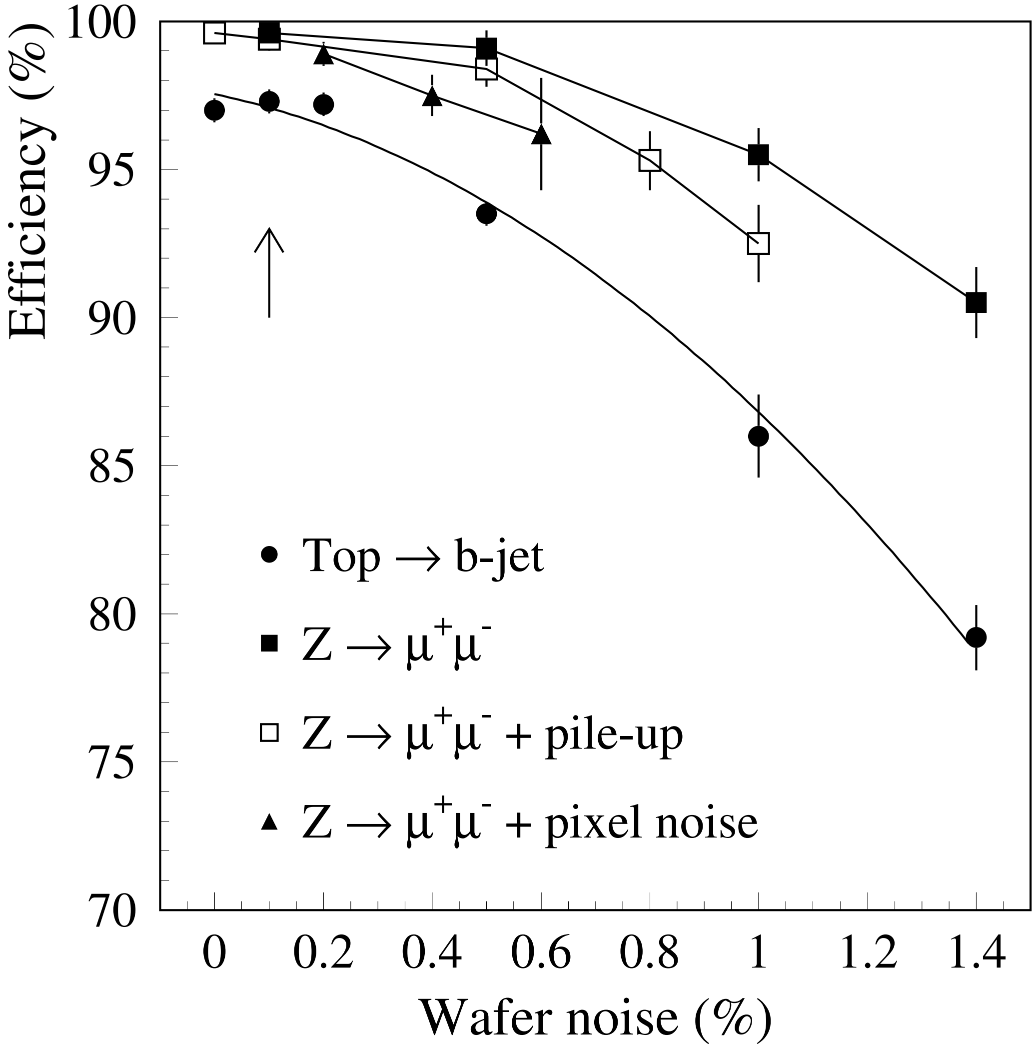


Figure 112: Efficiency as a function of noise on the wafer.

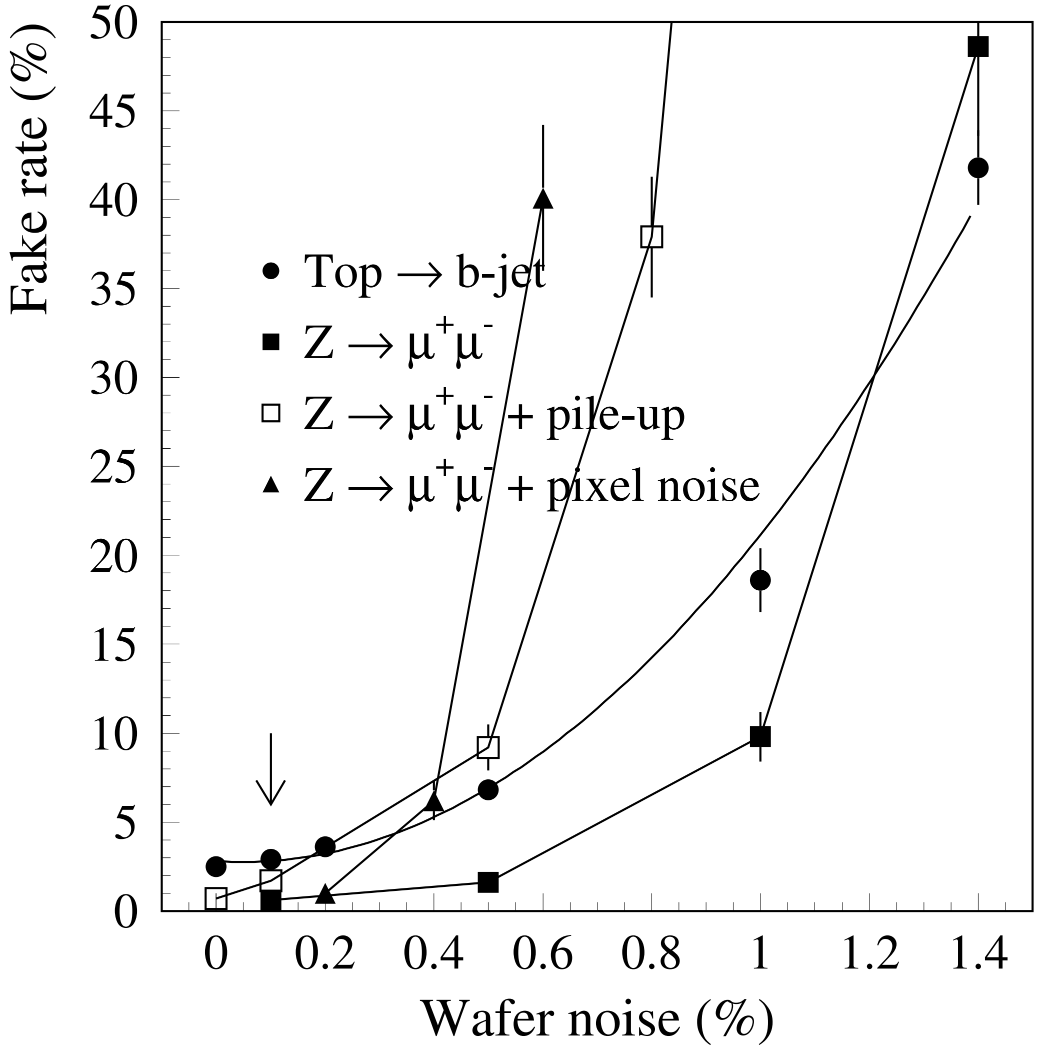


Figure 113: Fake rate as a function of noise on the wafer.

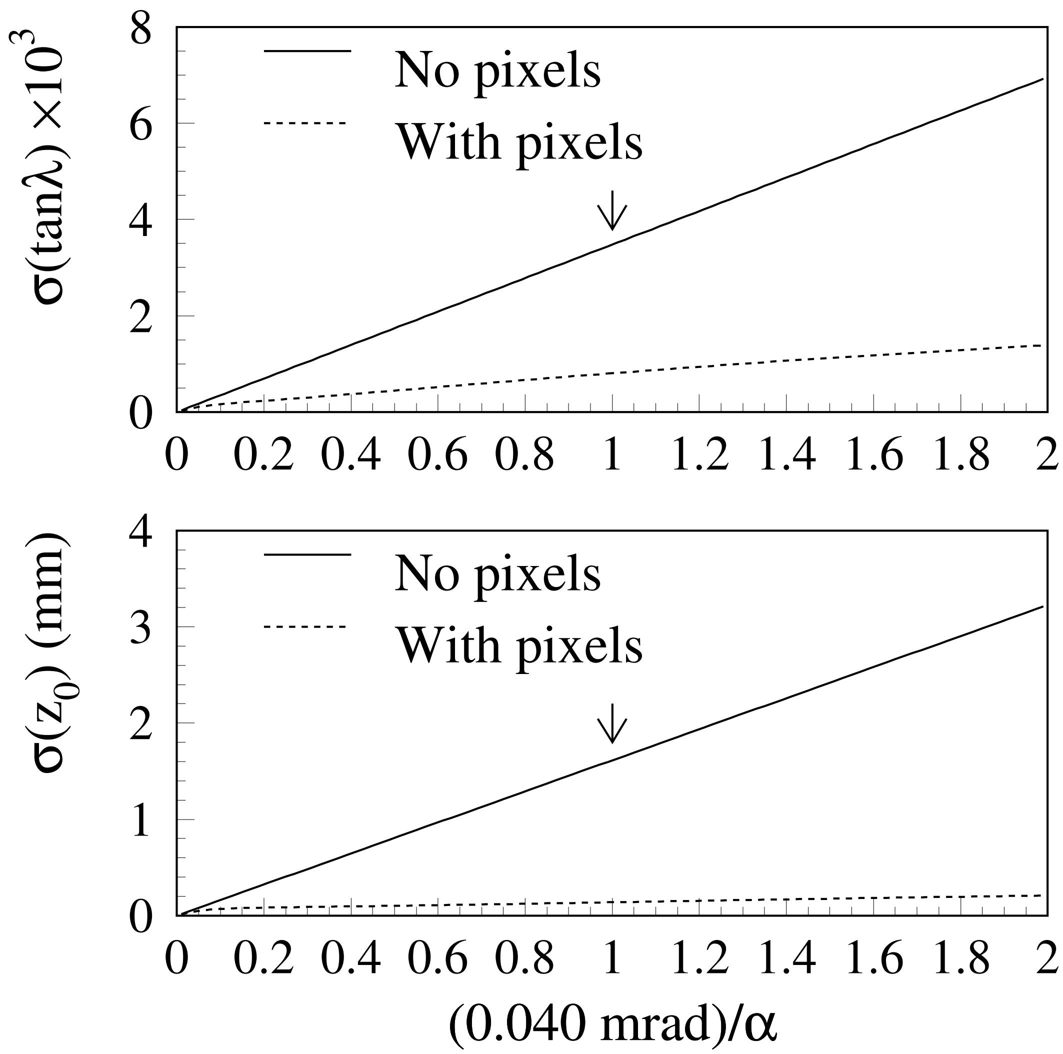


Figure 114: Resolution of fitted parameters $\tan\lambda$ and z_0 as a function of the inverse of the stereo angle α, normalized to the default of 40 mrad.

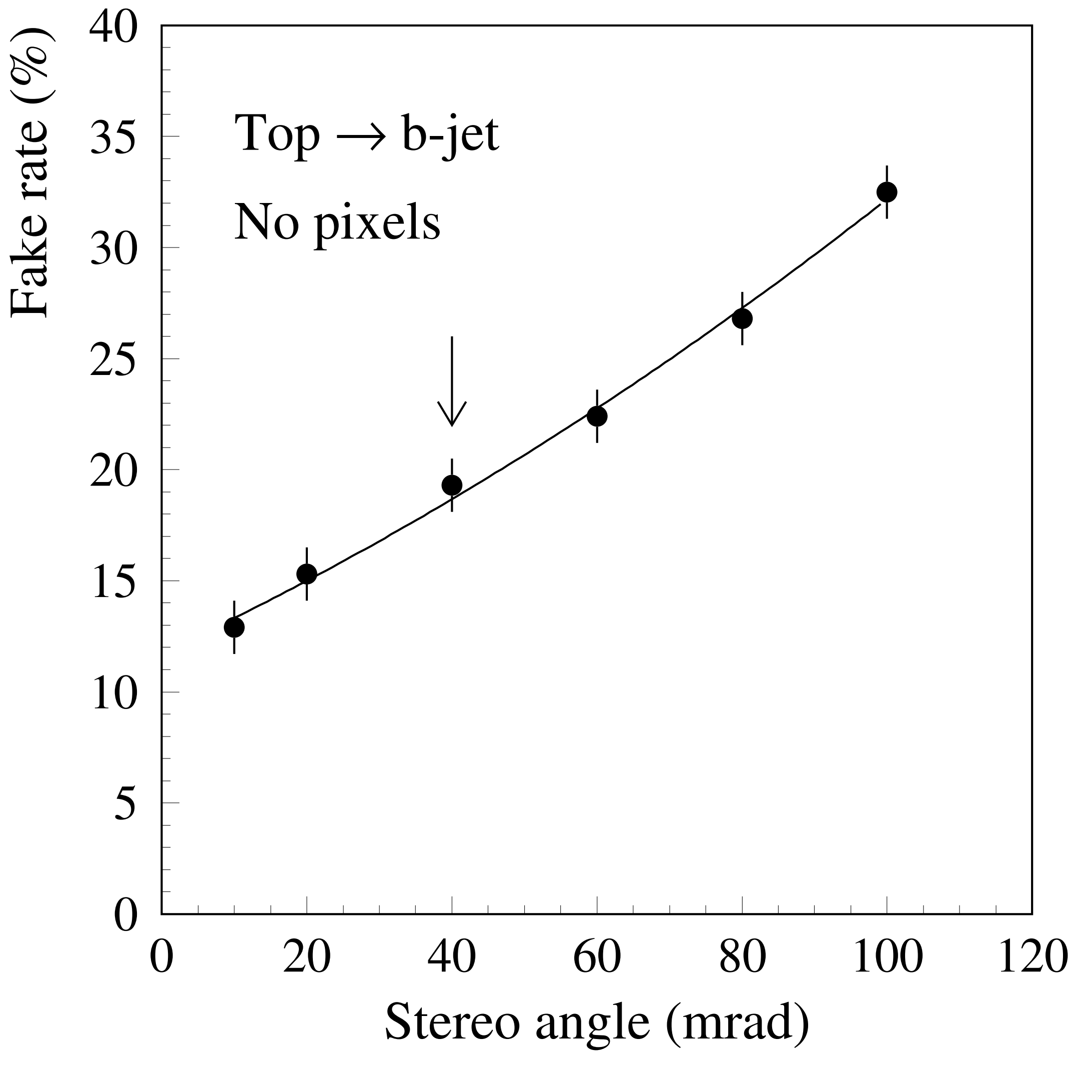


Figure 115: Fake rate as a function of the stereo angle α.

10.3.7 Silicon strip pitch

For pattern-recognition, there are two advantages of a finer strip pitch: Firstly it implies a better hit resolution. Therefore the extrapolation errors at planes where hits are to be searched for are smaller, providing greater discrimination between the correct and incorrect hits for a track candidate - this is compromised at low p_T by multiple scattering. Secondly, since the distance required to resolve two close hits is mainly a function of the number of strips separating them rather than the pitch itself, a finer strip pitch improves the two-hit resolution which can be achieved. The track-finding efficiency for b-jets as a function of the silicon strip pitch is shown in Figure 116 for tracks with $p_T > 6$ GeV/c - for lower p_T, the effect is less marked. The efficiency loss which can be seen arises from the inability to resolve individual hits in the silicon. The fake rate is fairly constant.

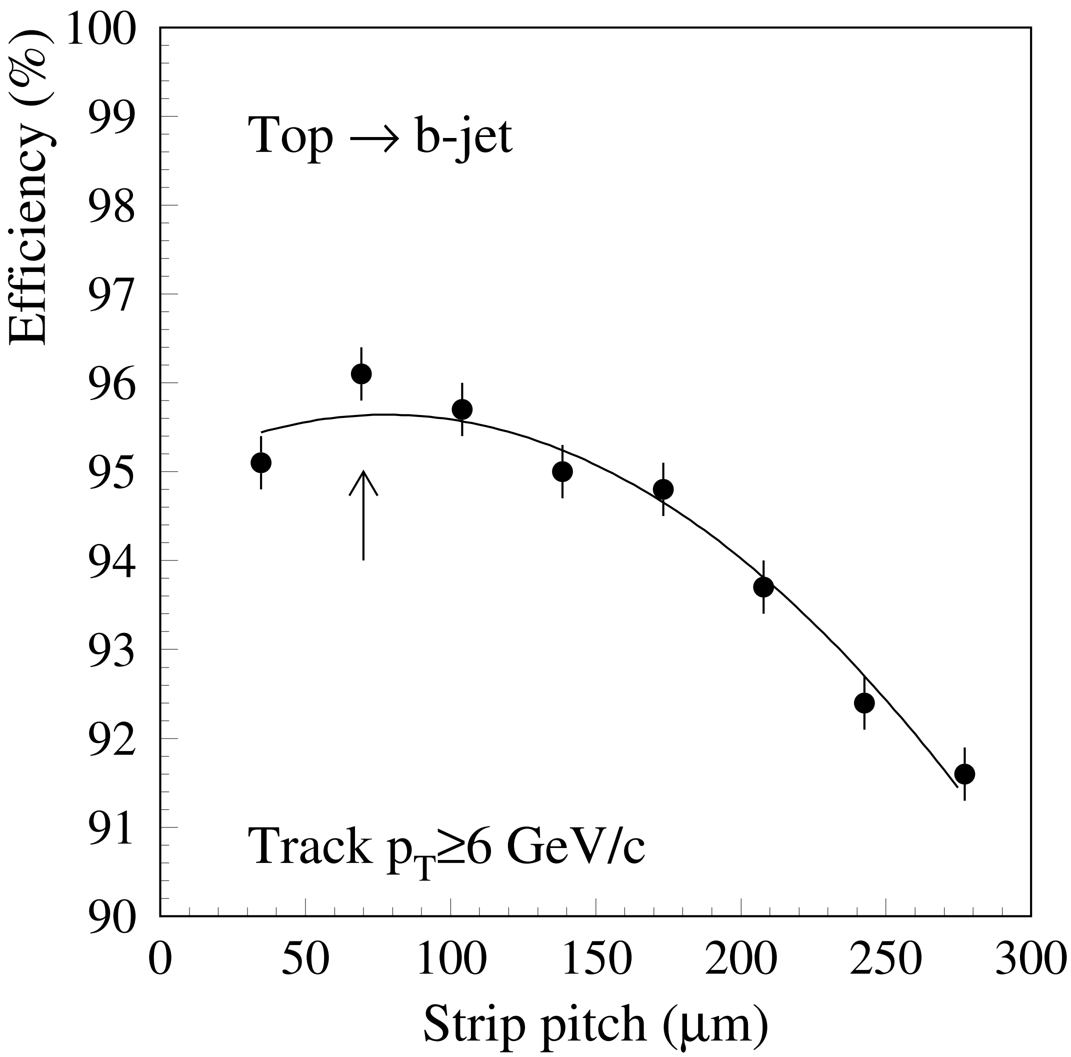


Figure 116: Efficiency as a function of silicon strip pitch.

The variation of transverse momentum resolution for large p_T (ignoring multiple scattering) is shown in Figure 117. More information can also be found in Ref. [116]. The effect of the 10 μm systematic causes the curves to flatten off as the intrinsic silicon resolution improves.

10.3.8 Alternative layouts

Macropixels Should it prove technically feasible, *macropixels* consisting of large pixels might provide an attractive replacement for the stereo superlayers. Macropixels providing resolutions of 29 μm by 580 μm in $r\phi$ and z respectively are considered. Two schemes could be envisaged: (a) macropixels (*Mpix*) at $r = 30$ and 50 cm with crossed-strips (*Xstrip*) (as in section 10.3.5) at $r = 40$ and 60 cm, (b) four planes of macropixels. Results comparable with Table 34 are shown in Table 35. The remarks about global algorithms made in section 10.3.5 are again relevant here.

Table 35: Performance of different designs incorporating macropixels for b-jets from top.

Design	With pixels	
	Efficiency (%)	Fake rate (%)
(a) Mpix + Xstrip	95	3
(b) All Mpix	99	1
	Without pixels	
	Efficiency (%)	Fake rate (%)
(a) Mpix + Xstrip	72	34
(b) All Mpix	98	2

Using scheme (a) and the same analysis used for tau's as described in section 10.2.4, in 98% of events, exactly three tracks were reconstructed.

Since both the strips and the macropixels have worse space-point resolution in $r\phi$ than a stereo superlayer, the momentum and impact parameter resolutions degrade by about 20 and 30% for schemes (a) and (b) respectively.

All stereo layers Should pixel technology prove to be unrealisable for the ATLAS Inner Detector, a detector consisting of only stereo layers in the Barrel could be constructed. A barrel consisting of five superlayers (v-ϕ, u-ϕ, v-ϕ, u-ϕ, v-ϕ) at 20, 30, 40, 50 and 60 cm has an efficiency for reconstructing tracks in b-jets of 92.5 ± 0.7%, falling by one per cent as the stereo angle is increased from 40 to 100 mrad. The corresponding fake rate rises from 3 to 5%.

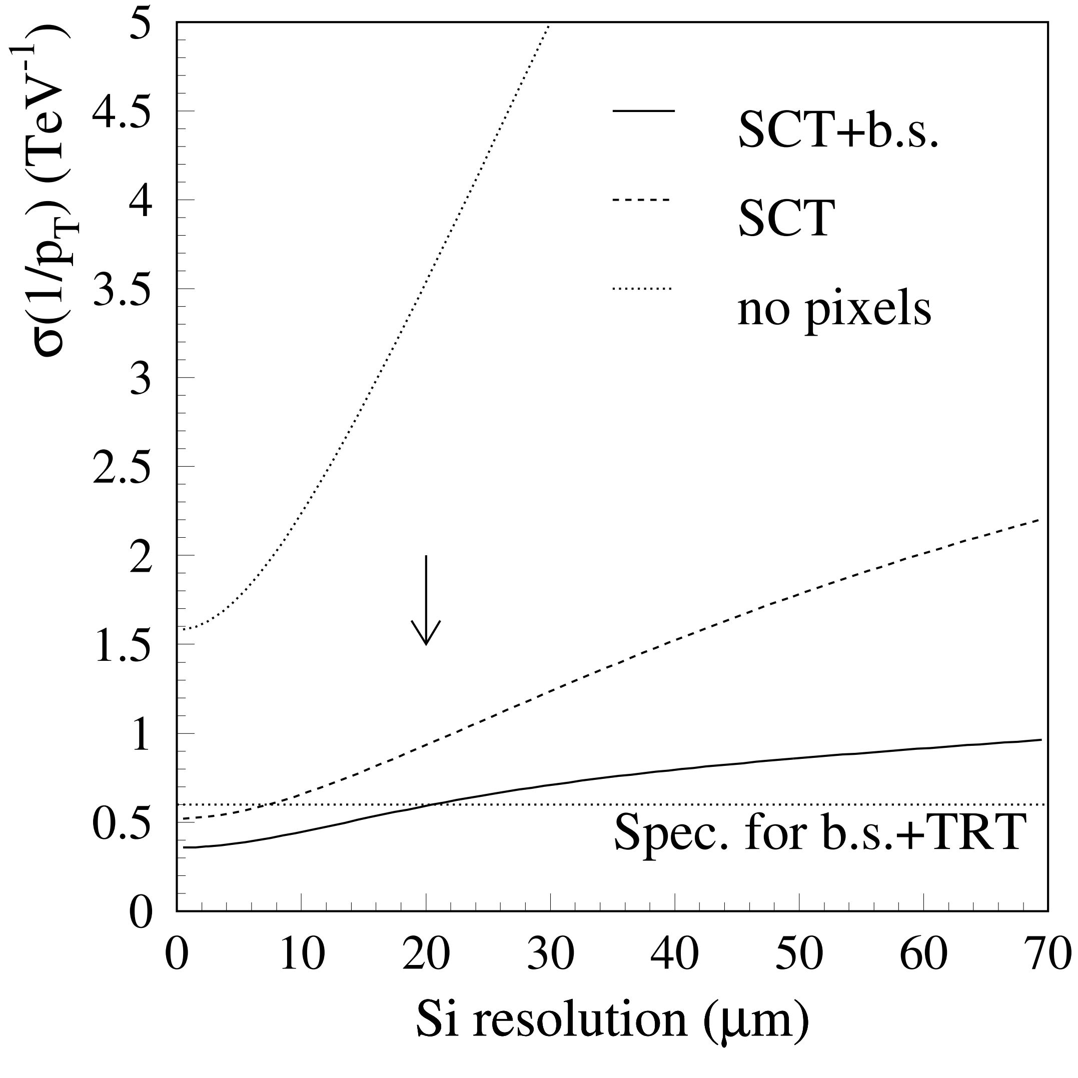


Figure 117: Transverse momentum resolution at large p_T as a function of intrinsic silicon strip resolution (in addition, there is a fixed systematic of 10 μm). Curves are shown for the SCT with a beam-spot constraint; SCT with no constraint; just silicon strip detectors with no pixels. The specification for the Inner Detector is intended to be met with TRT included and with a beam-spot constraint.

11 Prototyping for the ATLAS SCT.

A five or six year module construction, assembly, and installation phase is envisaged for the ATLAS Inner Detector. Comprehensive test and prototyping programs will be undertaken in the years 1995–1997 in preparation for this, to develop, test, and optimise the individual parts of the Inner Detector.

Two large prototype systems will be constructed to study system aspects of the Barrel and Forward Inner Detectors in test beams. This section describes the R&D activities needed to address the open technical issues for the ATLAS SCT. This includes the construction of a prototype sector of the Inner Detector Barrel system.

11.1 Summary of available prototyping and test beam results.

Test beam measurements obtained with LHC compatible electronics show that it should be possible to build and operate a silicon tracker system for the LHC with high efficiency, low noise, and the necessary resolution. The most critical parameter for the overall tracker performance is the noise of the front–end electronics. Below we summarise results supporting this conclusion.

11.1.1 KEK and HERA results.

Four layers of silicon detectors were placed in a 4 GeV negative pion beam at KEK. The detectors [117], were AC–coupled double-sided, 6 cm long crystals with 50 μm pitch and polysilicon biasing resistors of about 200k Ω. They were developed by the Japanese ATLAS collaborators with Hamamatsu for high luminosity operation. The signals were amplified, shaped, and discriminated by bipolar chips (UCSC-LBIC [118] or LBL-ATT [119]) and the digital results were sampled and buffered in digital pipeline CMOS chips (UCSC-CDP [120]) at 40 MHz asynchronously to the beam triggers.

Nominal readout pitch was 50 μm while other readout pitches (100, 150, 200 μm) were also implemented and measured. This was done by ganging together 2,3, or 4 50 μm detector strips. The ϕ-angular dependence up to 54^0 was measured. Efficiencies, resolution, and noise were measured while varying thresholds and detector bias voltages [121]. Excellent matching of noise and response was observed for all of approximately 750 channels.

Figure 118 shows the efficiency and noise background versus threshold for the p-side. The n-side looked very similar (see Tables 36-38), for 50 μm and 100 μm readout pitch. The noise is given for two 40 MHz time buckets, and has been determined in the signal events by cutting at $\pm$10 strips away from the projected hit. There is, in addition, an increased occupancy correlated with the position of the track which is almost independent of threshold. This is shown in Figure 119. The position resolution is measured to be slightly better than (readout pitch)/$\sqrt{12}$.

Table 36: Efficiency and occupancy for various pitches and two thresholds.

Pitch μm	V_{th} fC	ϵ(p) %	O(p) 10^{-4}	ϵ(n) %	O(n) 10^{-4}
50	1	99.98	0.2	99.7	0.25
50	2	98.0	0.13	96.8	0.13
100	1	99.92	0.7	-	-
100	2	99.5	0.3	-	-
100/150/200	1	-	-	99.6	3.0
100/150/200	2	-	-	98.8	0.4

The noise and signal distributions demonstrate a comfortable operation range with low noise and high efficiency. Table 36 shows efficiencies and occupancy for two thresholds, 1 fC and 2 fC, for different pitch values. The signal to noise ratios are shown in Table 37. The slightly smaller values for the n-side can be improved by increasing the bias voltage. For an occupancy of 10^{-3}, the ATLAS target, a threshold to noise ratio slightly below 4 is found, as shown in Table 38. Upcoming runs in early 1995 include tests of irradiated detectors.

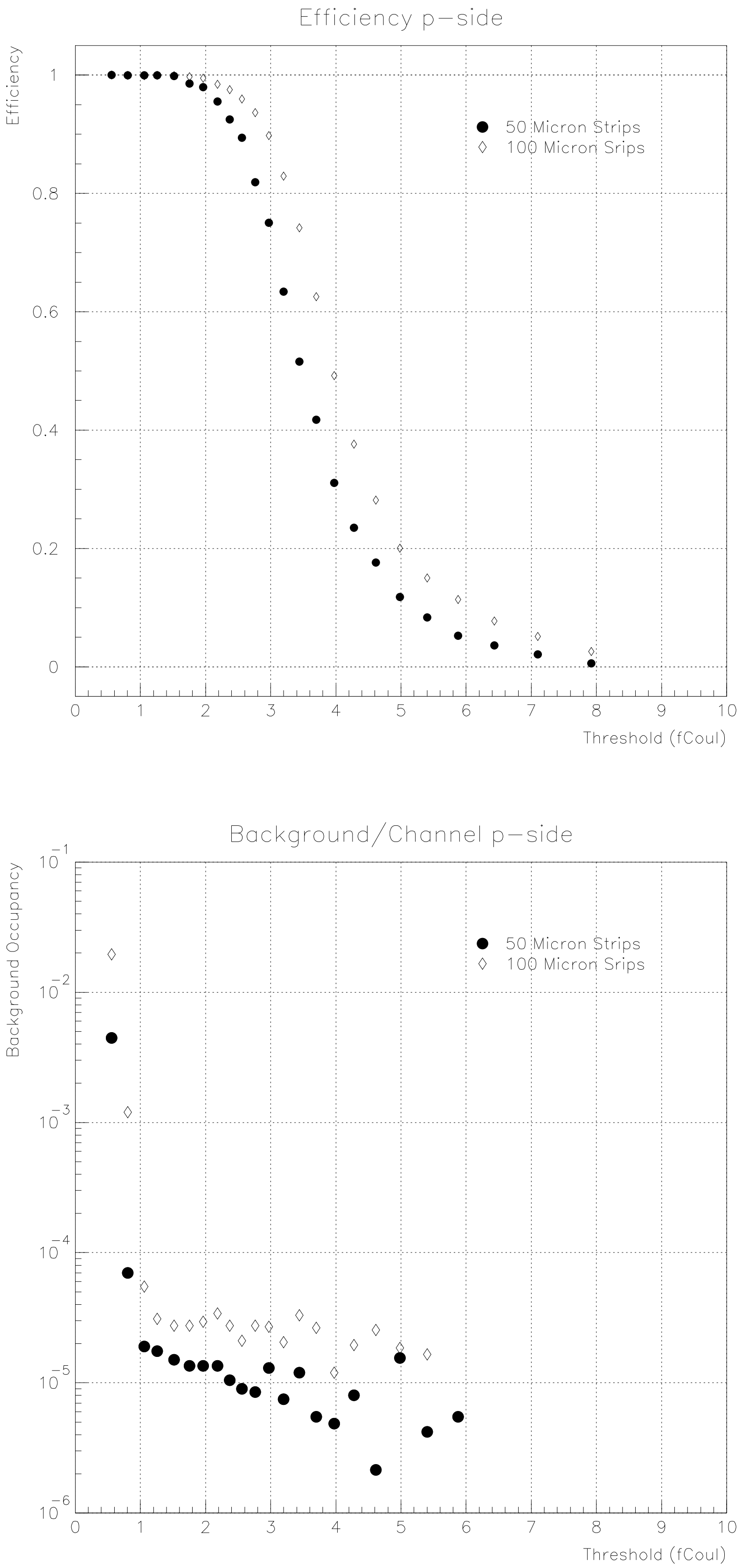


Figure 118: Efficiency and noise versus threshold for the p-side.

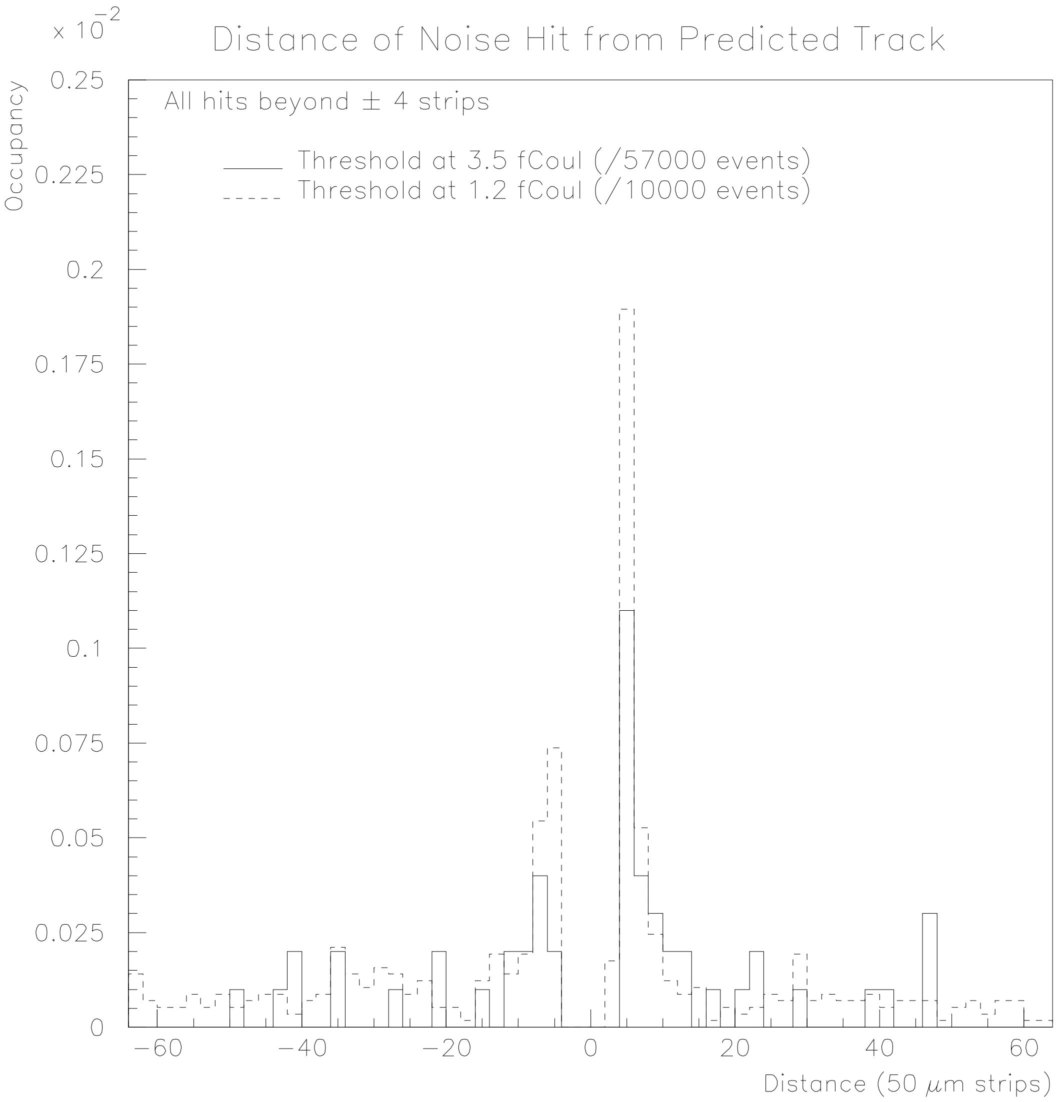


Figure 119: Spatial distribution of the noise hits.

Table 37: Signal to Noise ratio for the various detectors.

Side	Pitch μm	N fC	S fC	S/N
p	50	0.16	3.5	22
p	100	0.21	4.0	19
n	50	0.21	3.4	16
n	100/150/200	0.21-0.26	3.7	15

Table 38: Threshold to noise ratio for an occupancy of 10^{-3} in two 40MHz bins.

Side	Pitch μm	N fC	V_{th} fC	V_{th}/N
p	50	0.16	0.65	4.0
p	100	0.21	0.80	3.8
n	50	0.21	0.75	3.6
n	100/150/200	0.21-0.26	0.90	3.8

A binary readout system has also been in operation in ZEUS for about 6 months [122]. Although the crossing frequency is lower than at LHC, key characteristics, for example threshold control and diagnostic capability, are directly applicable to ATLAS. At a threshold level of 0.78 fC, the occupancy does not exceed 10^{-3}. System operation is stable and predictable.

11.1.2 H8 results.

Test structures consisting of AC coupled 6 cm single sided silicon detectors readout by complete 32 channel CMOS chips were operated and tested in the H8 test beam at CERN. One test chip per detector was used. The detectors were placed in the middle of an 8 plane high precision silicon telescope with a single plane resolution of 3 μm. The telescope was arranged as four pairs of orthogonal detectors, giving two high resolution space points on each side of the test setup. The tracks were reconstructed in the telescope and used to find the impact point at the test detectors. Signals, either defined by a single strip or by adding the signals from two adjacent readout strips, were searched for in the three strips closest to the impact point. One of the two test detectors was installed on a rotation table so that the ϕ dependence of resolution and efficiency could be measured. Approximately 9 million events were recorded.

Table 39: A number of single sided detectors with 0, 1, or 2 floating strips were tested.

Diode Pitch	Readout Pitch	Biasing	Delivered by
50	50	Foxfet	CSEM
50	100	Foxfet	MICRON
50	100	Polysilicon	SINTEF
67	200	Foxfet	MICRON

The signals from the detectors were amplified, shaped, stored, and processed in two CMOS chips: FELIX, bonded to a 32 channel analogue multiplexer [123] and APV3, for which a discrete multiplexer was used. Details about these chips can be found in Section 5 of this document. All 32 analogue channels of the chip were then transferred to the counting room and digitised. All results, unless otherwise stated, refer to the final deconvoluted output of the chips (as described in the same section).

Fig 120 shows results from a typical run with a 50–100 μm detector connected to a FELIX chip. A signal to noise ratio (Figure 120a) around 10.5 is found, defined as the sum of two adjacent strips divided by the noise in the strip with the highest signal. Figure 120b shows the efficiency versus noise of the system with

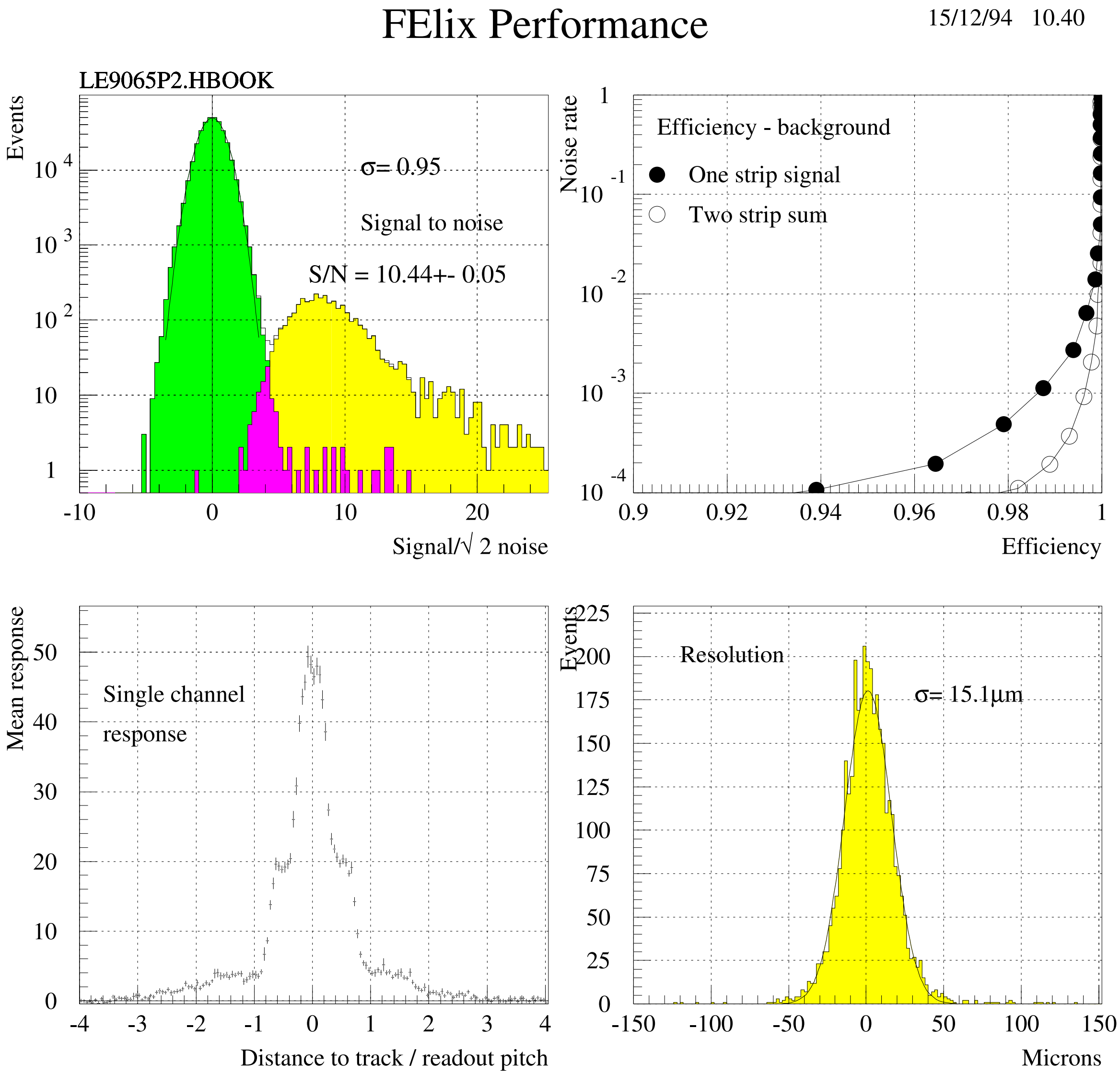


Figure 120: FELIX results. S/N, efficiency versus noise, detector response for 100μm readout pitch and one floating strip in between, resolution.

a variation in the threshold which defines a signal. The signal was searched for in in the window mentioned above and the noise was measured in channels more than 6 readout strips away from the track. Figure 120c shows the response of a single strip as a function of the distance to the particle impact point. The response defined here by the ADC response of the signal in the strip in the middle of the plot.

Figure 120d shows the residual between hits and tracks, essentially the hit precision, using

$$\eta = \frac{PH_l}{PH_l + PH_r}$$

Data was taken with the four detectors described at the beginning of this section to examine the effect of varying pitch. Figure 121 shows the resolutions achieved. The various points at any given readout pitch indicate the resolutions obtained for varying signal to noise values between 7 (APV3, deconvoluted) and 20 (FELIX, peak).

The conclusion is efficiencies well above 99.5% with noise below 10^{-3} are obtained. With a proposed ATLAS readout pitch of 112.5 μm the specified resolution of 20 μm can easily be obtained. The crucial parameter is the efficiency for a given noise, which is of course controlled by the signal to noise value of the front–end electronics.

The shape of the signal indicates that there will be no problem with hits from the previous bunch crossing confusing the triggered event. Further improvements of the two chips have been implemented and the improved FELIX will be tested in the KEK test beam early 1995, in a 1 T magnetic field, while the APV5 chip will be tested later in 1995.

Several other measurements were performed and the analysis is still in progress. In particular, resolution and efficiency as a function of angle, and measurements of irradiated detectors are expected to provide interesting information [124].

11.2 Future developments and prototyping activities.

11.2.1 FE electronics prototyping activities.

The readout electronics R&D and prototyping is related to the development of binary, analogue and digital systems for the ATLAS SCT. The final system will consist of a preamplifier/shaper, a pipeline structure (digital or analogue) and readout and control systems for the individual chips and modules. The goal is to select the optimal system by the end of 1995. On the same timescale, convincing progress related to radiation hard implementations of the various parts of these systems will be demonstrated. This will occur through the design and fabrication of vital parts of the IC's constituting the various systems in radiation hard technologies. A full radiation hard implementation of an optimised system is planned for 1996.

Some of the most crucial topics to be addressed and studied during 1995 are listed below.

- The noise performance of the preamplifier and pipeline systems for the various options will be measured.
- Can common mode baseline shifts be controlled well enough in a system that sparse readout of digitized analogue information or of binary information is possible?
- A study of a full optical analogue readout system will be made.

These items will be confronted in a number of prototype developments, lab tests and test beam measurements the coming year.

A complete binary system will be developed and tested in the H8 beamline at CERN in 1995. The front-end electronics is subdivided into to 128 channel IC's, an analogue chip using bipolar transistor technology and a digital IC fabricated in CMOS. Preliminary versions of these circuits are being used and have been successfully tested at HERA and in the 1993–1994 test beam at KEK as described above. Further tests in the 1995 test beam will include new optimised analogue front–end chips and the goal is to study system aspects of the binary system. This includes measuring the efficiency and resolution of the silicon system as function of angle of incidence, threshold setting, and detector bias voltage. General studies of the signal to noise ratio of the system will be carried out. As a part of these tests a realistic hybrid will be developed which includes all necessary functions to operate the detectors and front–end circuits including power distribution.

In addition to a bipolar front–end preamplifier two optimised CMOS preamplifier systems integrated with analogue pipelines will be tested in 1995.

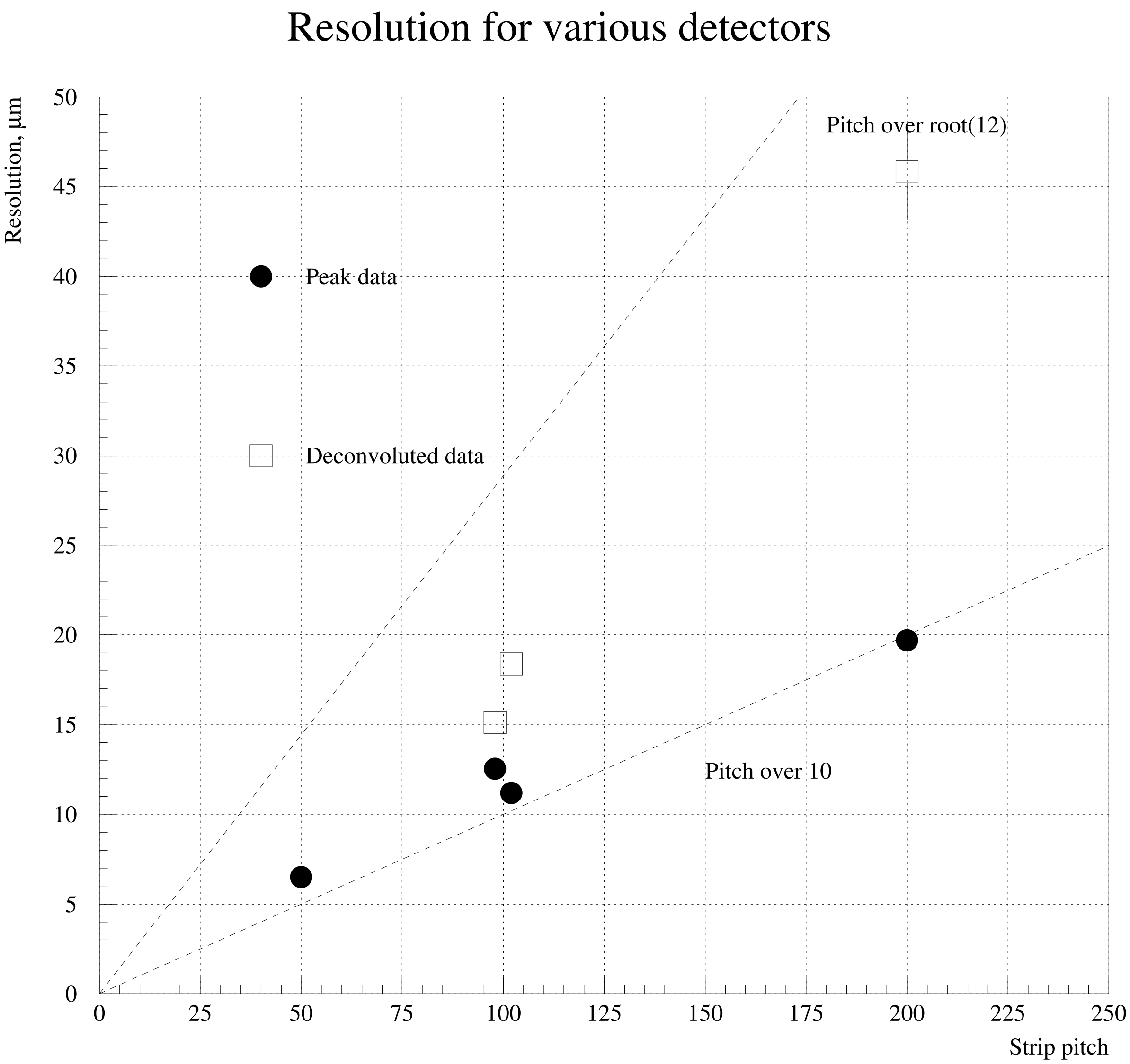


Figure 121: Resolution versus readout pitch for peak and deconvoluted signals. The S/N varies between 7 and 20 for the two modes of the two chips, and the resolution varies accordingly. With S/N $\approx$ 15-20 a resolution of readout pitch/10 is obtained.

One option is a charge sensitive amplifier followed by a 75 ns CR–RC shaping filter (FELix). To recover the initial charge and time resolution, a deconvolution algorithm will be executed on the events passing the level 1 trigger. This is done by an analogue processor at the output of the analogue delay buffer built specially for this application. A 32 channel version exists and an improved 128 channel version optimised for 15-20 pF input will be available for tests in 1995. A 128 channel 20 MHz analogue multiplexer will be bonded to this chip in order to test it. A radiation hard version of this chip (APV5) which includes a fast analogue multiplexer will also be available next year. Both these chip will be installed on fairly complete hybrids and read out in an analogue manner in the H8 test beam. The purpose of these tests is to demonstrate high resolution and efficiency for a system of several modules with analogue readout of full CMOS chips. As an important part of this program the goal is to have complete 4–optical channel links available for testing in Spring 1995. These will be used partly for tests of the full scale analogue readout chain from a single silicon module using the APV5 chip. These opto-electronic links will have discrete components (lasers, photodiodes and optical couplers) at the readout board.

Another CMOS solution is a charge-sensitive, high-gain, single-stage, fast amplifier (NICON) with rise-time around 15 ns. This amplifier can accept both AC and DC input connections and has a performance which is very robust against parameter variations and leakage current variations. This amplifier can be used for detector capacitances above 20 pF. The amplifier is integrated with an analogue memory where both the storage and shaping functions are simultaneously performed. This memory stores two consecutive samples for each channel. It is followed by a differential readout amplifier and an ADC on chip serving all 32 channels (ADAM). The chip is available and will be used to equip 3 relatively complete modules in the spring 1995, in the H8 beamline. The system will be read out digitally, without sparsification, and provides important tests of system threshold stability, baseline stability, and protection against parasitic effects in fairly complete chips, in addition to the tests of the individual ADAM chips.

Both analogue pipelines for the CMOS chips mentioned will be tested also in connection with a bipolar analogue front–end amplifier. These systems will be tested in 1995 (but not as full modules) in order to compare their relative merits in terms of signal to noise, efficiencies and stability. The preliminary noise performance of these systems look very promising as an alternative to a full CMOS solution read out in an analogue or digital manner.

It is foreseen to have an analogue pipeline followed by sparse readout and an ADC as one integrated chip available in 1995. Measurements of this chip will provide important information about the feasibility of this readout scheme. The key parameter to be investigated in such a system is the baseline stability [36].

Another important point which is addressed by the prototyping program is the control and monitoring system (CM) for the silicon detector modules. One essential part of the control and monitoring is the distribution of the BC–clock and the Level 1 trigger–accept. In addition, each subdetector will need a number of test, slow–control, monitoring and reset functions. Typical examples are temperature measurements, local clock adjustments, calibration signals, reset functions, discriminator level setting, and biasing. In connection with the developments mentioned above, several of these functions will be implemented and tested. In the H8 test beam a central system will provide the overall synchronization such that a number of modules can be tested in parallel. Also, the services for the modules will be implemented using realistic service cables (biasing, control, monitoring) to simulate ATLAS conditions.

11.2.2 Module mechanical and thermal design and prototyping.

Two module concepts will be prototyped for the 1995 test beam. A decision on the final design is foreseen for the end of 1995.

The prototyping has the following aims:

- Analysis of the module performance and projected cost to determine the most appropriate techniques for the final module.
- Development of realistic models of detector and electronics services, together with the examination of the interaction of these services with the mechanical precision and stability of the modules.
- Study of the temperature distribution across the module with irradiated detectors and a cooling channel operating at a fixed temperature.
- Development of manufacture and assembly techniques suitable for large scale production.

- Investigation of thermally conducting adhesives, especially with regard to their long term performance in a radiation environment.

The module concepts are the Z–module and the Rϕ-module described in detail in Section 6.

- Relatively complete versions of the Z-module will be studied in the 1995 test beam. Two variations of this concept will be tried out. One is based upon narrow aluminium cooling pipes suitable for a carbon fibre support structure and the other with the coolant contained within the beryllium staves which also acting as support (due to the high cost of extruded beryllium, aluminium staves will be used in the 1995 test beam).
- Several version of the Rϕ modules will be developed for the 1995 test beam.

11.2.3 Mechanical support structures.

Two mechanical support structure concepts are being studied.

A full length cylinder will be manufactured, from composite materials for mechanical studies, to ensure that the required specification can be achieved and to fully understand construction techniques. It will provide a cross-check for finite element analysis calculations. Similarly it is planned to construct a full length cylinder from barrel staves and to study its stability under different loadings. Module fixation methods and heat transfer to the cooling system have also to be prototyped very accurately.

A mock-up to study service routing is also foreseen, especially inside the barrel, to study interconnections with the modules, and at the barrel end face, to optimise connectivity, routing and feed–through to pass the pixel and silicon enclosure. The goal is to choose the best solution before the end of 1995.

For 1996 a 30^0 sector will be constructed to answer the following specific questions :

- How to assemble the detector modules, mount them on the mechanical supports and monitor a stable internal alignment?
- How to maintain and monitor the temperature of the system at 0^0C or lower?
- How to supply services to the detectors without introducing excessive material and without deforming the structure and causing uncontrolled stresses?

In addition the sector will act as a space–frame model so that the integration of detectors, installation, access, and maintenance procedures can be determined.

The sector is a half–length, $30°$ segment of the barrel region. The support for the segment will be as realistic as possible. Aluminium will be used to obtain some affordable compromise to the stability and material thickness provided by a beryllium structure in case this solution is chosen.

11.2.4 Cooling systems.

It is intended to develop a common cooling strategy for all barrel detectors.

Both binary ice and liquid coolant are being considered for the cooling. In the case of the stave support structure the cooling is integrated with the support. In the case of the cylindrical carbon fibre support structure the cooling is provided by aluminium pipes attached to the cylinder.

In connection with the module testing in test beam in 1995 the thermal stability of these modules will be measured and modelled both during data taking and in more extensive tests with binary ice afterwards. Monitoring of the temperatures in the system is another important aspect which will be included in the module prototypes.

11.2.5 Detectors.

R&D goals for the silicon strip detectors :

- High voltage operation. Particularly for ATLAS-B we need clear evidence of satisfactory operation of a device with single-sided lithography after inversion if n-type substrate is adopted. For all devices, the operation at 300V needs to be established and failure modes both before and after irradiation need to be understood and correlated, if possible, to allow better device screening pre-irradiation. (For

ATLAS-A with n-type substrate, the movement of the high field region after inversion to the strip side may reveal undetected weaknesses. Does this design also require lithography to define the p-doped non-readout face?) Statistics on strip device failures and rates of catastrophic failures (unable to hold high voltage without drawing more than 5 mA total) need to be evaluated.

- Strip bias. Only the direct coupled approach and the use of polysilicon resistors have been studied extensively by the RD2 and RD20 collaborations. However, punch through and FoxFET biasing have proved successful in less exacting radiation environments. For such 'reach-through' techniques to be adopted we need first to see the dynamic resistance and voltage drop in n-type silicon after surface and bulk damage at LHC levels. Biasing of p-type detectors using such techniques (for each given option for n-strip isolation) has yet to be demonstrated in detail and the performance after bulk and surface damage to LHC levels will need to be established before such designs can be considered.
- Interstrip resistance. Work exists on p-strips in n-type which suggest adequate isolation can be achieved after inversion (required bias voltages?) but studies of signal collection for partial depletion requires further study. The interstrip isolation for n-strips in n-bulk has been studied in RD20 but corresponding work for spray compensation and any p-bulk designs remain to be evaluated after bulk and surface damage. (This parameter should ideally be evaluated at LHC frequencies.)
- Interstrip capacitance. Once again the n-bulk options have been studied in some detail for both n and p strips but the corresponding work for p-type substrates remains to be pursued to the same depth. There is concern that the solutions requiring n-strips with p-isolation are difficult to design with the required initial interstrip capacitance. The consequences of incomplete depletion needs also to be studied. (This parameter should ideally be evaluated at LHC frequencies.)
- Coupling capacitors. The actual voltage tolerances need to be established and processing techniques which then show the required breakdown characteristics over a large number of strips need development. Of particular concern is double-sided designs requiring read-out near zero volts on both faces.
- Further studies of annealing of detectors and the expected properties after the 10 year scenario need further investigation, particularly those relating to surface damage. The effects with p-bulk should be shown to conform to expectations.
- Designs with large gaps between implants allow lower capacitance albeit at the cost of lower charge collection efficiency if sparse read-out is adopted. Simulation results suggesting dangers of incomplete depletion and greater susceptibility to breakdown have yet to be fully corroborated.
- Charge collection. The issues of collection times in partially depleted devices, effects of bulk and surface damage, any dependence on biassing techniques, holes vs electrons, efficiencies, trapping and the effects of the magnetic field have largely only been studied in simulation. All these require further empirical investigation.

11.2.6 Summary

The R&D programme is summarized in Table 40.

11.3 The module prototyping program in 1995.

Most of the test beam studies for the SCT system will take place in the H8 beamline at CERN. In the beamline a high precision silicon telescope has been installed. With 8 planes of silicon detectors (32×32 mm^2), 25 μm pitch, 50 μm readout pitch a single plane resolution of around 2 μm is obtained.

In the middle of the telescope, various ATLAS silicon systems are being operated and tested. In front of the silicon telescope another silicon telescope is used to test pixel systems; behind that, a TRT system is being studied followed by calorimeters.

In April/May 1995 several individual front–end readout chips will be tried out. Three modules with front–end chips with integrated ADCs (ADAM), realistic mechanical and thermal construction will be installed and tested.

In August 1995 two near complete module systems, with several front–end readout chips, binary or full analogue optical readout will be installed and operated in the H8 test beam.

Table 40: Summary of major R&D items and plans for completion.

Item	Technical Issue	How Addressed	Time scale
Binary readout	Stability of threshold	Operate several modules in beam	1 year
Analogue Readout	Analogue Optical data-transfer	Build and operate a system for a module	1 year
Digital Readout	Stability of threshold, sparse readout, data flow	Build and operate a full system	15 months
Amplifier systems	Improve signal/noise	Further design optimization	1 year
All FE electronics	Radiation hardness	Design in rad. hard technologies	2 years
Module mechanics	Thermal stability	Develop and operate the two module versions in test beam, lab measurements, thermal modelling	15 months
	Production methods	Build several modules, define production schedule and set up production lines	2 years
Mechanical Support	Mechanical stability	Build two full cylinders	6 months
	System Assembly	Construction of a Barrel Sector	18 months
Cooling system	Thermal stability, reliability	Operate and monitor cool modules	1 year
	Integration with support structure	Construction of a Barrel Sector	18 months
Detectors	Radiation hardness	Measurements of irradiated detectors, design optimization	2 years

The 1995 milestones :

- Build and operate realistic modules. Measure efficiency at a noise level of 10^{-3} or lower, measure space and time resolution of the system.
- Perform some of these measurements with irradiated detectors.
- Operate a full optical readout system.
- Verify the thermal and mechanical properties of the modules.

The systems will be operated at conditions similar to the final LHC conditions (i.e simultaneous read and write) and with realistic service cables.

At the end of the test beam period a pixel, silicon strip, TRT and calorimeter system will be operated in common mode as a preparation for the BSP system test the year after. Various system tests of operation, track reconstruction, two track resolution, alignment and readout can be performed in this period.

11.4 The Barrel Sector Prototype Program, system tests in 1996–97.

The performance a 30° sector of the ATLAS Barrel Inner Detectors will be evaluated in a magnetic field. The sector will be installed inside the Morpurgo magnet, with a magnetic field of 1.8 T, in the H8 beam at CERN. The aperture of the magnet limits the efficient length of the module to 1.0-1.2 m. The SCT part of the sector will be equipped with modules with close to final layout and performance.

The BSP is an important milestone in demonstrating that a large ensemble of the ID subdetectors can operate together and provide the performance required by the physics goals of ATLAS. The specific goals of the BSP are :

- Demonstrate the required momentum resolution for electrons.

  This requires that the sector be placed in a magnetic field (in front of the electromagnetic calorimeter module), that all layers of a sector are equipped, and that the material distribution is realistic.

- Prove pattern recognition and vertex reconstruction for hadronic jets.

  A hadron beam with an interaction target consisting of thin foils will allow the sector to be used as study tool for developing track and vertex reconstruction algorithms.

- Study system issues :
  - Assembly
  - Operation at temperatures close to 0°.
  - Alignment
  - Evaluation of electronic system performance
  - Readout and trigger studies
  - Synchronization, control, and monitoring of the detector system
  - Offline studies

  The complete understanding of structural behaviour will require additional engineering prototypes. In particular it will be necessary to fabricate at least one full size cylinder.

- Study the transition, including services, between the Barrel and Forward trackers.

  Because of the limited aperture of the available magnet, this may require a dedicated test beam run outside the magnet.

With these goals in mind the priority is to equip the BSP with fully functional detectors, mounted on a realistic structure, and read out. The use of radiation tolerant components (first unirradiated and later irradiated) is seen as a second stage.

Overview of the BSP time schedule:

1995
- Design of Sector structure (mechanical, cooling, cabling...)
- Definition and first tests of alignment procedure
- Magnet plus cryogenics available in H8
- Further integration into and development of the ATLAS test beam DAQ
- Development of common clock, trigger, control distribution system for the BSP
- Synchronized operation and testing of fully engineered prototype subdetector modules

1996
- Commissioning of magnet in H8
- Construction of sector with integrated cooling system
- Sector partially equipped with readout
- Test beam evaluation with magnet
- Data for alignment, resolution, and pattern recognition studies
- Construction of radiation tolerant modules

1997
- Evaluation of radiation tolerant modules
- Combined BSP and calorimeter data
- Continuing studies on alignment, resolution, pattern recognition

12 Project Planning

This section will be completed this year (1995).

12.1 Schedules

For the Technical proposal, a schedule was prepared assuming first physics runs in 2003. Working back, this led to the requirement to make all major choices relating to the baseline by the end of 1995, and to start production in 1997.

12.2 Costs

A cost estimate was made making some assumptions on technical choices. The dominant cost is the cost of the detectors which is based on an extensive market survey. Details of the cost model can be found in the ATLAS Cost planning document v4.0 submitted to CORE. The total costs of the Pixel system, the silicon strip system and the GaAs strip system have been estimated as 12.8 MSFr, 30.8 MSFr and 4.8 MSFr respectively.

12.3 Responsibilities and Resources

The division of responsibilities will only be finalised once technologies have been fixed and a baseline agreed. The resources of the 31 (13, 6) groups interested in constructing the Silicon strips (Pixels, GaAs strips) seem to be well matched to the cost estimates given above. The interested groups are shown in Table 41.

Table 41: Institutions interested in the construction of the SCT (by country)

Si pixel detectors	Prague; Marseille; Bonn, Dortmund, MPI Munich, Wuppertal; Florence, Genoa, Milan; LBL, Hawaii, UC Irvine, Wisconsin.
Si strip detectors	Melbourne, Sydney; Montreal; Prague; Helsinki; Bonn, Dortmund, Hamburg, MPI Munich; Hiroshima, KEK, Kyoto UE; Oslo; Cracow INP and FPNT; Moscow State U, Novosibirsk; Uppsala; Bern, CERN, Geneva; Birmingham, Cambridge, RAL, Liverpool, QMW London, Oxford; LBL, UC Irvine, UC Santa Cruz, Wisconsin.
GaAs strip detectors	Sydney; Freiburg; IHEP Protvino; Glasgow, Lancaster, Sheffield.

Appendices

A Vendor Specifications for Detectors

To make a final choice of technology and to identify companies interested in undertaking production on this scale, we are forced to seek accurate costings from companies for three detailed sets of device specifications. We also list some general 'performance requirements' which relate to the desired performance after 10 years exposure to the hostile radiation environment of the LHC. For an understanding of the cost differences, essential for the overall design, we request quotations based on these lists of specifications. However it is very important for us to also understand **what are the 'cost drivers' for different companies** and information volunteered on **reduced costing for stated relaxations of the specification** would be most welcome.

A further important aspect of this exercise is information for the preparation of the Technical Proposal of the entire experiment for the Autumn of this year. Costings, as realistic as possible, need to be presented in this document to ensure that the overall costs of the experiment can remain within budget. This exercise is therefore also aimed at providing the first systematic survey of companies interested in this technology to find what overall price we should anticipate. Great care will be taken to treat quotations provided in the context of this inquiry with the strictest confidentiality and no company names will be associated with the final costings disclosed in the Technical Proposal of the ATLAS experiment or in any other disclosures of those involved in this exercise. The CERN management has agreed to supervise this activity to ensure the strictest fairness is administered.

Delivery times for prototypes and for the final production run are to be negotiated with the individual manufacturers, but should clearly be compatible with $2-3$ years delivery, starting 1996/97, of the order. (Clearly if a company only has capacity for a fraction of the total order this should be specified with the quotation.) It is anticipated that two mask sets will be required to allow 12000 detectors of 2 detector designs, one set having strips of $40mrad$ to the $z-$direction to provide stereo information. In the case of double-sided detectors only 6000 wafers total are required. Companies should quote for orders of 1000, 2000, 3000, 6000 (single-sided) and for the full order. Quotes for prototypes, capable of delivery for evaluation in 1995 are also invited for 25, 50, and 100 evaluated detectors. Clearly a company's ability to produce detectors to the specification needs to be established before a final order could be placed.

For the final order we reserve the right to buy from more than one manufacturer if several companies have produced successful prototypes and can demonstrate the required production capacity at the agreed specification.

ATLAS-A Single-sided Capacitively Coupled $r\phi$

- **Overall area** 60×60 mm (Scribe to scribe)
- **Tolerance on lateral dimensions** $\pm 20\mu m$. Edge to $< \pm 10\mu m$ w.r.t. fiducials
- **Thickness** $300\mu m$
- **Tolerance on thickness** $-20\mu m + 0\mu m$
- **Flatness** $\pm 10\mu m$
- **Number of strips** 1536
- **Number read out** 512
- **Substrate** n-type or p-type high-ρ silicon (both n-strip) (See performance requirements)
- **Strips** High doped n-implant. Pitch: $37.5\mu m$
- **Strip dimensions**: Width $27.5\mu m$ (depends on isolation), length $\sim 60\ mm$
- **Read-out pitch** $112.5\mu m$ (both ends)
- **Guard to scribe distance** $500\mu m$
- **Guard rings total width** $700\mu m$
- $C_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- $R_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- $C_{\text{coupling}} > 20 \times C_{\text{TOT}}$ $(ie >\sim 30pF/cm)$
- $R_{\text{implant}} < 100k\Omega/cm$
- $R_{\text{aluminium}} < 20\Omega/cm$
- R_{bias} **or** $\frac{dV}{dI} > 400k\Omega$ (& $R_{poly} < 2M\Omega$) (For bias, see performance requirements)
- **Uniformity** Strips at same potential to $\pm 1V$ at $2\mu A$
- **Reach through voltage** Voltage drop to strip $< 20V$ at $300V$ and $2\mu A$
- **Depletion voltage** $< 100V$ p-type, $< 150V$ n-type
- **Operating voltage** $300V$ (See performance requirements)
- **Total leakage** *incl* **guard** $< 50\mu A$ at $300V$
- **Maximum leakage per strip** $100nA$ at $300V$
- **Coupling dielectric** < 2 shorts
- **Capacitor breakdown** $> 100V$
- **Defective strips** < 2, metal/diode breaks or shorts to neighbours
- **Strip linearity** Within $2\mu m$ of nominal
- **Test structures** Large diode, large capacitor, MOS capacitor, resistors (To be negotiated)
- **Evaluation** Contact holes to implant (Prototype detectors only)
- **Read-out** $200 \times 50\mu m$ bond pads, $\geq$ two rows, daisy-chainable.
- **Passivation** Detectors to be passivated (Process negotiable)

ATLAS-B Single-sided Direct Coupled $r - \phi$

- **Overall area** $60mm \times 60mm$ (Scribe to scribe)
- **Tolerance on lateral dimensions** $\pm 20\mu m$. Edge to $< \pm 10\mu m$ w.r.t. fiducials
- **Thickness** $300\mu m$
- **Tolerance on thickness** $-20\mu m + 0\mu m$
- **Flatness** $\pm 10\mu m$
- **Number of strips** 768
- **Number read out** 768
- **Substrate** n-type (p-strip) and p-type (n-strip) prices sought (see performance requirements) (Single-sided lithography assumed)
- **Strips** High doped p/n-implant. Pitch: $75.0\mu m$
- **Strip dimensions**: Width $15\mu m$, length $\sim 60mm$
- **Read-out pitch** $75.0\mu m$ (both ends)
- **Guard to scribe distance** $500\mu m$
- **Guard rings total width** $700\mu m$
- $R_{\text{aluminium}} < 25\Omega/cm$
- $C_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- $R_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- **Depletion voltage** $< 100V$ p-type, $< 150V$ n-type
- **Operating voltage** $300V$ (See performance requirements)
- **Total leakage** *incl* **guard** $< 50\mu A$ at $300V$
- **Maximum leakage per strip** $100nA$ at $300V$
- **Defective strips** < 2, metal/diode breaks or shorts to neighbours
- **Strip linearity** Within $2\mu m$ of nominal
- **Test structures** Large diode, MOS capacitor (To be negotiated)
- **Read-out** $200 \times 50\mu m$ bond pads, $\geq$ two rows, daisy-chainable.
- **Passivation** Detectors to be passivated (Process negotiable)

ATLAS-C Double-sided Capacitively Coupled $r\phi$

- **Overall area** $60mm \times 60mm$ (Scribe to scribe)
- **Tolerance on lateral dimensions** $\pm 20\mu m$. Edge to $< \pm 10\mu m$ w.r.t. fiducials
- **Thickness** $300\mu m$
- **Tolerance on thickness** $-20\mu m + 0\mu m$
- **Flatness** $\pm 10\mu m$
- **Number of strips** 1536
- **Number read out** 512
- **Substrate** n-type or p-type high-ρ silicon (See performance requirements)
- **Strips** High doped p&n-implants. Pitch: $37.5\mu m$
- **Strip dimensions**: Width $27.5\mu m$ (depends on isolation), length $\sim 60\ mm$
- **Read-out pitch** $112.5\mu m$ (both ends)
- **Guard to scribe distance** $500\mu m$
- **Guard rings total width** $700\mu m$
- $C_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- $R_{\text{inter-strip}}$ Read-out – read-out (See performance requirements)
- $C_{\text{coupling}} > 20 \times C_{\text{TOT}}$ ($ie >\sim 30pF/cm$)
- $R_{\text{implant}} < 100k\Omega/cm$
- $R_{\text{aluminium}} < 20\Omega/cm$
- R_{bias} **or** $\frac{dV}{dI} > 400k\Omega$ (& $R_{poly} < 2M\Omega$) (For bias, see performance requirements)
- **Uniformity** Strips at same potential to $\pm 1V$ at $2\mu A$
- **Reach through voltage** Voltage drop to strip $< 20V$ at $300V$ and $2\mu A$
- **Depletion voltage** $< 100V$ p-type, $< 150V$ n-type
- **Operating voltage** $300V$ (See performance requirements)
- **Total leakage** *incl* **guard** $< 50\mu A$ at $300V$
- **Maximum leakage per strip** $100nA$ at $300V$
- **Coupling dielectric** < 2 shorts through oxide at $150V$ (Both faces)
- **Capacitor breakdown** $> 150V$
- **Defective strips** < 2, metal/diode breaks or shorts to neighbours
- **Strip linearity** Within $2\mu m$ of nominal
- **Test structures** Large diode, large capacitor, MOS capacitor, resistors on both sides (To be negotiated)
- **Evaluation** Contact holes to implant (Prototype detectors only)
- **Read-out** $200 \times 50\mu m$ bond pads, $\geq$ two rows, daisy-chainable
- **Passivation** Detectors to be passivated (Process negotiable)

Performance Requirements Exposure to radiation of silicon detectors induces changes in effective doping, driving the detectors more $p-$type, increases the leakage currents and (with charged particles) leads to a build up of positive charge at the silicon−dielectric interface.

To achieve adequate depletion depths after 5 to 10 years of LHC operation, the detectors should be able to operate reliably at a bias voltage of $300V$. In addition, even at the operating temperature foreseen of 0^oC, leakage currents of up to $2\mu A$ must be anticipated. (However, The initial specifications assume measurement at room temperature.) Furthermore, oxide charges can build up to levels of $3 \times 10^{12} cm^{-2}$ before saturating even after modest ($100kRad$) doses. It is essential that detectors be as tolerant as possible to all these changes in their operating conditions. In particular, the electronics will show degraded noise performance if the inter-strip resistance were to drop below $\sim 10M\Omega/cm$ or the inter-strip capacitance were to rise above $0.5pF/cm$. (**Inter-strip meaning between two adjacent read-out strips**.) It is hoped that the total capacitance seen by the electronics, inter-strip + further neighbours + back-plane + bond wires, can be kept to $< 18pF$ for a module of $2\times6cm$ length for the duration of the experiment. An initial total capacitive loading of the electronics of $15pF$ is required for sufficient signal/noise to ensure full efficiency.

The choice of technology for the isolation of n−strips in n−type or p−type silicon must allow the required inter-strip resistance to be achieved after 10^{14} protons/cm^2 and 10^{14} neutrons/cm^2 whilst still not giving an inter-strip capacitance much above $0.5pF/cm$.

In the case of p−strips in n−type material, these requirements have been shown to be met, but the operation after the material reverses type is expected to require that the depletion region extends to the read out surface. For such inverted detectors, the operation at high voltage is therefore particularly critical with further constraints coming from the need to see the signal within $25ns$.

The interest in p−type material comes from these concerns coupled with the advantages of the higher mobility of electrons. Depending on the cost differential, a rapid R&D programme with this technology has to be foreseen. The preferred choice for the final detector with p-type silicon will be that with the lowest possible boron content. However, since this material will probably not be readily available in small quantities, $> 8k\Omega cm$ p-type will be acceptable for the prototypes.

A further variant that has also been proposed is to work with $250\mu m$ thick detectors, but the consequent reduction in signal in this case places even more severe constraints on the electronics performance.

It is therefore important that after irradiation, operation at $300V$ should still be possible and with less than 1% of the channels showing $\geq 5\mu A$ current. The large guard region specified is anticipated to be needed to keep guard currents below $1mA$ at $300V$ after irradiation. Vendors are invited to give quotes for lower operating voltages, in addition, should this requirement be a significant 'cost driver'.

Although no biassing scheme is specified, the requirements listed are most easily met using polysilicon resistors although this technology involves more masks and implants than for punch-through or FoxFET biassing. Companies are invited to quote for whichever technology(s) they have available, but the requirements of bias (dynamic) resistance $> 400k\Omega$ at the final leakage current, substrate doping and oxide charge needs to be proven for techniques other than polysilicon.

The defect specification is fairly strict and if this looks to be a 'cost driving' item, costs at looser specs are invited in addition. A similar comment also applies to the capacitor breakdown, the device dimensions and the leakage currents/operating voltage, so long as a quotation is always provided for the specifications given above.

Should a company elect to provide indications of prices with relaxed specifications, the differential cost with respect to the full specification should be given. It is anticipated that full costings only be given for the three detector types listed, with the differential costs of the different options (a requirement in the case of the substrate type for direct coupled devices) provided as appropriate.

As a final comment, delivery on a best effort basis cannot be accepted so companies must specify what quality assurance can be guaranteed prior to delivery. This, and other performance specification related issues, may best be discussed with the ATLAS SCT representatives in institutes local to the manufacturer prior to submission of a final quotation.

Summary Sheet

- **Company Name:**
- **Preferred Technical Solutions** Choice of substrate, biassing, capacitor dielectric, guard structure etc.
- **Capacity** Examples of previous orders with delivery timescales.
- **Prototyping** Time and resources required.
- **Prices** (In confidence if preferred) for full specifications.
  Number **ATLAS-A** **ATLAS-B** **ATLAS-C**
- Prototype Batch
  (25/50/100)
- 1000
- 2000
- 3000
- 6000
- 12000

- **Relaxed Specifications** Price differences with chosen relaxations of key cost driving items.
- **Quality Assurance** Facilities at company and list of routine inspection procedures.

B Prototype ATLAS Data Transmission Protocol

Data Transmission Protocol. Data from physics events as well as monitoring and status information will be sent from the on-board electronics in serial form using NRZ–L format. This means that a “1” is represented as a high level (light on for optical) and a “0” is represented as a low level (light off for optical) during the bit time interval. No return to baseline (no return to zero) will occur between consecutive bit levels. Each serial link will transmit complete data from some subset of a detector subsystem (i.e. data is not logically multiplexed across fibers). The actual number of channels or readouts serviced by one link will be determined by the specific readout architecture. The protocol is defined in three levels. The first level defines a Data Packet which is in common with all architectures and all types of data. The second is the Data Header which is also in common with all architectures but defines the type of data present in the packet. The third level is the actual Data Format which is specific to the readout architecture. One Data Packet will be sent for each accepted Level 1 Trigger or each request for information (via control command).

1. **Data Packet:** Each Data Packet will start with a Preamble consisting of the following bit pattern (data is shifted out to left with leftmost bit out first): 11101 (5 bits total). Each Data Packet will end with a Trailer consisting of the bit pattern: 100000000 (9 bits total). Between the Preamble and Trailer any data length field is allowed (including zero).

2. **Transmitting Algorithm for the Data Packet:** For each data packet sent the transmitter will initiate the transmission with a Preamble. Note that it is assumed that a Trailer has ended a preceding Data Packet and that idle cycles between Data Packets are effectively in “0” state. If an initial Data Packet is sent immediately upon power-up, the transmitter must assure that a minimum number of “0” cycles, equal to the length of a Trailer, have appeared on the link before starting the Preamble.

   As the data is serially transmitted, the transmitter will monitor outgoing data. After any sequence of 7 consecutive “0”’s (one fewer than the number of “0”’s in the Trailer) , the transmitter will insert a “1” into the outgoing data sequence. This of course does not apply when the transmitter is sending the Trailer.

3. **Receiver Algorithm for Data Packet:** When Receiver is in idle state (i.e. not currently accepting a Data Packet), it is constantly looking for a match to the Preamble pattern. With each bit read, the Receiver will compare the most recently received 5 bits against each of the patterns of Table 3 looking for a match (leftmost bit read first):

Uncorrupted Preamble	11101
Single bit flip	11100 11111 11001 10101 01101

   This action is immune to any single bit error within the Preamble or within the 5 bits preceding the Preamble. When a match is found, the Receiver goes into Receiving State.

   When Receiver is in Receiving State (i.e. actively processing a Data Packet), it must monitor all incoming data looking for a match to the Trailer pattern. When Trailer pattern is found, Receiver ends receipt of that Data Packet and goes to Idle State.

   When Receiver is in Receiving State, if seven “0”’s (one less than in the Trailer) are found followed by a “1”, the Receiver will remove the “1” from the data sequence.

4. **Data Header:**

   The first 13 bits of data in a Data Packet after the Preamble will be a header with the form Table 4:

   - **DT (Data Type):** 0: Level 1 trigger data; 1: Information data
   - **Level 1:** Current count of Level 1 Trigger modulo 256 since the last system reset. This field can be used for event building by the DAQ and also to monitor for lost data. The length of 256 was chosen to be safely longer than the Level 2 latency or buffer length.

DT (1 bit) (0=LVL1) (1=data)	8 bits LVL1 num IDPT	4 bits BC num IDP	left read 1st

- **BC:** Current count of Beam Crossing modulo 16 since the last system reset. This field is not really to tag Beam Crossing numbers since it wraps around so quickly. It is intended to monitor for clock pulses lost by the on-detector electronics.
- **IDPT:** The Information Data Packet Type is an 8- bit field to designate what kind of information data is being sent. Type codes are not defined yet, but they could be the 8-bit Sub-command code used in the Control Function Protocol to request the information. (See description of Control Function Protocol in following section.)
- **IDPT number:** Current count of Information Data Packets (IDP's) modulo 16 which have been sent since the last system reset. Note that there is only one rolling count of the number of IDP's on each link, not an individual count for each type of IDP. This is to minimize the number of counters in on-detector electronics. The purpose of this count is to facilitate detection of lost IDP's by the off-detector electronics as the Level 1 number is used to detect lost trigger data.

5. **Level 1 Trigger Data:**

   If the DT value in the Data Header is "0", the data being sent is in response to a level 1 trigger with the on-detector readout system in normal data taking mode. The format of the rest of the Data Packet until the Trailer is dependent upon the readout architecture.

6. **Information Data:**

   If the DT value in the Data Header is "1", the data being sent is in response to a request for information via the control system, an error status being reported, or data resulting from the on-detector readout system being in some type of diagnostic mode.

7. **Data Format for Silicon Micro-strip Analogue Readout Architecture:**

   The data format for this readout architecture has not been worked out in detail. Presumably, after the digital information contained in the Data Header, a sequence of analogue pulses will follow. The first being a calibration pulse followed by a pulse for each strip readout by the serial link. Any further definition of data format for this architecture awaits input from the groups working on it.

8. **Data Format for Silicon Micro-strip Digital Readout Architecture:**

   The data format will consist of a sequence of a strip address field followed by one or more pulse height fields. This pattern will be continued for each cluster of hits. A cluster is defined as any group of consecutive strips whose pulse height exceeds the threshold of the sparsification circuit. In the case that Neighbor Logic is active, the nearest neighbor strips of any exceeding threshold will be included in the cluster. The address of the cluster transmitted before the sequence of pulse heights will be the address of the first (or lowest address value) strip of the cluster.

   Each address field will start with a "0" in the most significant (first read) bit position. Each pulse height field will start with a "1" in the most significant (first read) bit position. Since the number of bits in an address field and a pulse height field are fixed, the receiving circuitry has a well determined, if not, trivial task to separate the data.

   Table 8 is an example of a data sequence of two clusters, the first cluster having 3 strips reported, the second cluster has one:

   IC/Ch Adrs: 3 bits of IC address followed by 7 bits of channel address each read most significant bit first. PH: 7 bits of pulse height.

   The 10 bits of address assumes the current ATLAS baseline of 128 channels per IC and 8 ICs per double sided module. The assumed design will use 2 serial fibers per module initially with the possibility that

0 IC/Ch addr	1 PH	1 PH	1 PH	0 IC/Ch addr	1 PH
10 bits	7 bits	7 bits	7 bits	10 bits	7 bits

one fiber can service the entire module if one fiber dies. Also, the 7 bit pulse height is tied to the digitization specification which presently assumes a 7 bit ADC.

9. **Data Format for Silicon Micro-strip Binary Readout Architecture:**

   The data format will consist sequentially of data from each readout IC in the module. Data from each IC will start with a Chip Header of three bits. If there are no hits in the IC, that is all the data transmitted from that IC. If the Chip Header indicates at least one hit is present, it will be followed by a 16-bit Map. Each bit in the Map represents the OR of 8 channels in the IC. A "1" in the Map indicates at least one hit in those 8 channels. Following the Map will be a sequence of 8-bit Zoom Patterns one for each "1" in the Map. Each bit in the Zoom Pattern represents one channel. A "1" in that bit position indicates that channel was hit.

   Data from an IC with no hit channels: Chip Header 100

   Data from an IC with a hits in channels 36, 97 and 102 (counting channel numbers 0-127 per IC):

Chip Header	Map	Zoom Pattern	Zoom Pattern
111	0000100000001000	00001000	01000010

Control Function Protocol. Level 1 Triggers and Control Function requests will be sent to the on-detector electronics via the same serial link. This reduces the cost, power consumption and material in the detector. Information will be transmitted on this serial link in Unipolar RZ format running at the 40 MHz clock rate of the machine and the on-detector electronics. A "1" will be represented by a high level (light on for optical) and a "0" will be represented by a low level (light off for optical) during the bit time interval which will be half a clock period. Between bit time intervals the signal will return to a low level (light off for optical). A valid Level 1 Trigger will be denoted by three bits "110" (the "0" received last) in time with three consecutive clock pulses. The selected beam crossing of such a trigger will be timed to the last of the three pulses minus a fixed level 1 latency. This can be repeated in succession. For example, a pattern of "110110" would denote two Level 1 Triggers on beam crossings 3 and 6 relative to the time of the first received pulse.

Other Control Functions will be initiated by a sequence of three consecutive pulses reading "101". In this case, these three bits must be followed by another 4-bit field which could define any one of 16 commands. Only 5 possible patterns of these 4 bits are chosen for use since other patterns are less immune to misinterpretation if transmission errors occur during a rapid sequence of Level 1 Triggers. Four of these 5 commands request immediate, fast response by the on-detector electronics (e.g. reset). These commands require no further fields to be transmitted. The last of these commands demands a following sub-command field of 8 bits to designate any one of 256 commands which may require parameters and may take considerable time to execute. Some of these "slow" commands may even put the electronics into a non-data taking mode. The number of fields and their length following each sub-command are determined by the specific command and a fixed (not variable) for each command. We have not defined any of these commands yet (some examples are given below). Some command codes should be in common across readout architectures but many will be specific to one type. Most of these command codes await a definition of the control architecture, but this protocol should be generic enough to service all that is planned.

One of the possible parameters required for a "slow" command might be a chip ID so that the command will only apply to a single front-end chip. For example, if the command is to load a mask register, only one chip is specified and only one mask is shipped as part of the parameter list. A restriction is placed upon any command which requests information to be returned by the on- detector electronics. Any such command must not specify a chip ID but rather must be global to all chips serviced by the link carrying the command. This simplifies the on- detector electronics in that Data Packet are formed in the standard way for all types of data or information.

The following table summarizes some possible Trigger/Control messages (specific Commands and Sub-commands are only examples):

Code 3 bits	Command 4 bits	Sub-command 8 bits	Parameters depends upon subcommand	Description
110	none	none	none	Level 1 Trigger
101	0100	none	none	Global Reset
101	0010	none	none	not yet defined
101	1110	none	none	not yet defined
101	0101	none	none	not yet defined
101	0111	0000 0000	none	Turn off Clock-thru-mode
101	0111	0000 0001	none	Turn on Clock-thru-mode
101	0111	0100 0001	0011, 128-bit mask	Set Mask Register Register for chip 3
101	0111	1000 0001	none	Read Mask Register for all chips
101	0111	0100 0100	0001 00 1000	Set Threshold DAC on chip 1 to 8
101	0111	1000 0100	none	Read Threshold DAC for all chips

List of Figures

List of Tables

References

[1] Pixel detector backup document to support the ATLAS technical proposal. *ATLAS Internal Note, INDET-NO-086*, 1994.

[2] D. Munday K. Jakobs and B. VanEijk. On the requirements for the central tracking system in ATLAS. *ATLAS internal document GEN-TR-150*, Feb. 1993.

[3] The ATLAS inner detector. *CERN/LHCC/93-24/, Mar.*, 1993.

[4] D. Froidveaux and M. Parker. The performance specifications of the ATLAS inner detector. *ATLAS internal document INDET-NO-046*, May 1994.

[5] A. Fasso et al. *Proc. IV Int. Conf. on Calorimeters and their Applications, World Scientific, Singapore, 493*, 1994.

[6] Aurenche et al. DTUJET-93 sampling inelastic proton–proton and anti-proton–proton collisions according to the two component dual parton model. *Comp. Phys. Commun 83:107-123*, 1994.

[7] Gorfine and Taylor. Particle fluxes and damage to silicon in the ATLAS inner detector. *ATLAS internal document INDET-NO-30*, Nov 1993.

[8] Hamburg group. Standard luminosity profile.

[9] A. Van Ginneken. Non-ionizing energy deposition in silicon for radiation damage studies. *FERMILAB-FN-522*, 1989.

[10] P. Aarnio, M. Huhtinen. Estimation of pion induced displacement damage in silicon. *HU-SEFT-R-1993-02*, 1993.

[11] A Chilingarov et al. Radiation studies and operational projections for silicon in the ATLAS inner detector. *Nucl. Instr. Methods.*

[12] A. Reichold and F. Fares. *ATLAS Internal Note, INDET-NO-070*, Nov. 1994.

[13] Status report of the RD8 collaboration. *CERN/DRDC 94-32*, Oct. 18, 1994.

[14] G. Lutz at al. *ATLAS Internal Note, INDET-NO-056*, 1994.

[15] A. Bischoff et al. Breakdown protection and long-term stabilization of Si-detectors. *Nucl. Instr. Methods A326 27-37*, 1993.

[16] G. Cesura et al. *ATLAS Internal Note, INDET-NO-053*, 1994.

[17] R. H. Richter. Strip detector optimization using two dimensional device simulation. *Perugia Conference*, 1994.

[18] Specification List for SCT.

[19] R. Apsimon, private communication. Power constraints.

[20] E. Barberis et al. Operational experience with the front-end electronics for the LPS silicon microstrip detectors. *SCIPP 94-29*, 1994.

[21] E. Barberis et al. A fast shaping amplifier-comparator integrated circuit for silicon strip detectors. *IEEE Trans. Nucl. Sci. NS-40/4 740-743*, 1993.

[22] Analog front-end electronics for the SDC silicon tracker. *Nucl. Instr. Methods A342 205-213*, 1994.

[23] I. Kipnis et al. An analog front-end bipolar-transistor integrated circuit for the SDC silicon tracker. *IEEE Trans. Nucl. Sci. NS-41/4 1095-1103*, 1994.

[24] E. Spencer et al. To be published in *IEEE Trans. Nucl. Sci.*, 1994.

[25] W. Dabrowski et al. Fast bipolar front-end for binary readout of silicon-strip detectors. *CERN PPE/94-55*, 1994.

[26] I. Kipnis, private communication. 1994.

[27] E. Barberis et al. A low power bipolar amplifier integrated circuit for the ZEUS silicon-strip system. *3rd International Conference on Advanced Technology and Particle Physics, Como, Italy, 22-26 June, SCIPP 92/35*, 1992.

[28] M. French et al. The SDC silicon tracker DAQ. *Nucl. Instr. Methods A342 214-217*, 1994.

[29] Y. Unno et al. Beam test of the SDC double-sided silicon-strip detector, in preparation. *Conference Record of the IEEE Nuclear Science Symposium, San Francisco, Calif.,ISBN 0-7803-1487-5, pp. 38-42*, 1993.

[30] The RD2 collaboration. Collaboration report. *CERN/DRDC, 94-34, Aug*, 1994.

[31] RD2. *IEEE on DHARP2.*

[32] N. Bingefors et al. *Nucl. Instr. Methods A326 112.*, 19943.

[33] S. Gadomski, G. Hall, T. Hogh, P. Jalocha, E. Nygard and P. Weilhammer. The deconvolution method of fast pulse shaping at hadron colliders. *Nucl. Instr. Methods A320 217.*, 1992.

[34] R. Brenner et al. *Nucl. Instr. Methods A339 564.*, 1994.

[35] R. Brenner et al. *Nucl. Instr. Methods A339 477.*, 1994.

[36] CDF Collaboration. Design of the readout IC for the CDF SVX-II silicon-strip detector. *Proc of 3rd Int. Workshop on Vertex Detectors,Lake Monroe, IN.*, May 1994.

[37] Requirements of electronic component technologies for space applications. *ESA/SCC Basic Specification No 24300.*

[38] Requirements of monolithic microcircuits technologies. *ESA/SCC specification No 24390.*

[39] F.M. Newcomer et al. A fast, low power amplifier-shaper-discriminator for high-rate straw tracking systems. *IEEE Trans. Nucl. Sci. 40/4 630-636*, 1993.

[40] W. Dabrowski. private communication of measurement results at UCSC. 1992.

[41] N. Cartiglia et al. Radiation hardness measurements on bipolar test structures and an amplifier-comparator circuit. *Conference Record of the IEEE Nuclear Science Symposium, Oct. 25-31, 1992, Orlando, Florida (ISBN 0-7803-0883-2), vol. 2, pp. 819-821*, 1992.

[42] M. Dentan et al. Study of a CMOS-JFET-Bipolar radiation hard analog-digital technology suitable for high energy physics experiments. *IEEE Trans. Nucl. Sci. 40/6 1555-1560*, 1993.

[43] E.J. Kennedy et al. Radiation effects on JFETs, MOSFETs, and bipolar transistors as related to SSC circuit design. *Nucl. Instr. Methods A303 452-457*, 1991.

[44] LBL, I. Kipnis, private communication.

[45] Evaluation test programme for monolithic integrated circuits. *ESA/SCC Basic specification No 2269000.*

[46] C.Barns. Post radiation effects in integrated circuits. *RADECs proceedings Montpellier*, 1991.

[47] S. Tedja et al. Noise spectral density measurements of a radiation hardened CMOS process in the weak and moderate inversion. *IEEE Trans. Nucl. Sci. NS-39/4 804-808*, 1992.

[48] C. Haber, private communication.

[49] M. Acciarri et al. (The L3 SMD Collaboration). The L3 silicon microvertex detector. *CERN-PPE/94-122*, 1994.

[50] H.J. Ziock et al. Measurement of proton induced radiation damage to CMOS transistors and PIN diodes. *IEEE Trans. Nucl. Sci. NS-37/3 1238-1241*, 1990.

[51] H.J. Ziock et al. Tests of the radiation hardness of VLSI integrated circuits and silicon-strip detectors for the SSC under neutron, proton, and gamma irradiation. *IEEE Trans. Nucl. Sci. NS-38/2 269-276*, 1991.

[52] W. Dabrowski et al. Noise measurements on radiation-hardened CMOS transistors. *Conference Record of the IEEE Nuclear Science Symposium, Nov. 2-9, 1991, Santa Fe, New Mexico, IEEE Catalog no. 91CH3100-5, Vol. 3, pp. 1536-1540*, 1991.

[53] J. DeWitt. A 128-channel digital pipeline chip for silicon-strip detector readout. *UCSC SCIPP unpublished note.*, 1994.

[54] E.H.M. Heijne F.Faccio and P. Jarron. Study of device parameters for analog IC design in a 1.2 μm CMOS-SOI technology after 10 Mrad. *1992 IEEE NSREC proceedings*, 1992.

[55] F. Faccio et al. Noise characterization of transistors in a 1.2μm CMOS-SOI technology up to a total-dose of 12 Mrad (Si). *1994 IEEE NSREC proceeding.*, 1994.

[56] J.L. Leray et al. CMOS-SOI hardening above 100 megarad (SiO_2); limits and liabilities. *IEEE Trans Nucl. Science,NS-37*, 1990.

[57] G. Borel and C Terrier. A multi-Mrad hardened 8 bit 20 MHz flash ADC. *RADECS Montpellier proceedings, page 271.*, 1991.

[58] E.Dupont-Nivet et al. PJFET on Simox for rad-hard analog devices. *IEEE International conference on SOS/SOI Technologies, Key West, Florida, Oct.*, 1990.

[59] RD9. *CERN DRDC status report.*, 1994.

[60] J.D. Dowell (Birmingham) et al. *CERN/DRDC/91-41; DRDC/P31 Oct. 9*, 1991.

[61] M. S. Haben. Applications of opto-electronics in high energy physics. *PhD Thesis, University of Birmingham.*, 1993.

[62] A. Mosely et al. C. Kirkby, M. Goodwin. Low voltage InGaAs/InP multiple quantum well reflective Fabry-Perot modulators for optical communications at microwave frequencies. *IEEE Proceedings 139*, 1992.

[63] C. da Via et al. Lightwave analogue links for LHC detector front-ends. *Nucl. Instr. Methods A344 199-212.*, 1994.

[64] Europtics Ltd.

[65] Birmingham, Cambridge, Glasgow, Lancaster, Liverpool, Oxford, Queen Mary and Westfield College of London and Sheffield University and Rutherford Appleton Laboratory. A proposal to develop a prototype of a sector of the barrel tracking for ATLAS. *UK PPESP proposal PPESP/93/33 Aug*, 1993.

[66] R.B Nickerson and A.R. Weidberg. Clock & control. *ATLAS internal document INDET-NO-048*, May 1994.

[67] B.G. Taylor. Timing trigger and control distribution for LHC detectors. *Working review document, CERN/ECP.*

[68] R.B Nickerson and A.R. Weidberg. Clock & control for the test beam. *ATLAS internal note in preparation.*

[69] J. Gorbald. Specification for Doric.

[70] T. Ohsugi Y. Unno. *ATLAS internal document INDET-NO-062*, Sep. 1994.

[71] Y. Unno et al. *IEEE Nucl. Sci. Symp, San Francisco, Vol1. p38*, 1993.

[72] Annecy ID workshop. *ATLAS internal document TR-0142*, June 1994.

[73] A. Seiden.

[74] H.F. Sadrozinski and A. Seiden. Limits to the Use of Irradiated Silicon Detectors. *Santa Cruz Institute for Particle Physics, SCIPP 94/06*, March 1994

[75] H. J. Ziock et al. Temperature dependence of the radiation induced change of depletion voltage in silicon pin detectors. *Nucl. Instr. Methods A342 96*, 1994.

[76] A. Chilingarov et al. *Contribution paper (#0943) to the 27th Int. Conf. High Energy. Phys., Glasgow*, 1994.

[77] H. Sadrozinski. Marseilles ID workshop. *ATLAS internal document TR-0153*, Sep. 1994.

[78] Y. Unno. *Nucl. Instr. Methods A342 193; and references there-in*, 1994.

[79] D. Amidei et al., *NIM A350 73-130*, 1994.

[80] MRSI 505 Assembly Workcell, manufactured by MRSI, 25 Industrial Avenue, Chelmsford, MA, 01824, USA.

[81] H. Sadrozinski. Risk assessment of the silicon tracker operations.

[82] W.O. Miller et al. Thermal Runaway in silicon-strip Detectors *To be released as an ATLAS INDET*

[83] T. Kondo. *ATLAS internal document TR-NO-153.*, Sep. 1994.

[84] Superconducting Super Collider silicon tracker system R&D. *Los Alamos National Lab Report LA-12029, Dec.*, 1990.

[85] Dornier, G. Helwig. Highly dimensional stable composite structures. *Advanced Materials Workshop for High Precision Detectors, CERN 94-07*, 1994.

[86] E. Perrin. Note on stability of carbon fibre structures. *20-10-94*, 1994.

[87] S. Robitaille. *Advanced Materials Workshop for High Precision Detectors, CERN 94-07*, 1994.

[88] Nuclear Metals Inc., N. F. Leroy. High specific stiffness beryllium and beryllium alloys. *Advanced Materials Workshop for High Precision Detectors, CERN 94-07*, 1994.

[89] A. Holmes. Design of a carbon fibre based composite sandwich cylinder for possible use as a silicon detector support.

[90] Y. Unno et al. The $r\phi$ module. *Nov. 16*, 1994.

[91] E. Perrin. Carbon fibre sandwich cylinder for the barrel silicon detector. *Nov.*, 1994.

[92] RD2 Status Report. *CERN/DRDC/94-34/, Dec. 8*, 1990.

[93] C. Goessling. Module design selection. *Dortmund Workshop, Aug.*, 1994.

[94] Dow-Corning. Thermal grease dc-360.

[95] K. Daum. Alignment of the ATLAS Inner Detector. *ATLAS Internal Note, INDET-NO-060*, 1994.

[96] N. Bobroff. Recent advances in displacement measuring interferometry. *Meas. Sci. Technol. 4 907-926*, 1993.

[97] G.P. Barwood, P. Gill, W.R.C. Rowley. Laser diodes for length determination using swept-frequency interferometry. *Meas. Sci. Technol. 4 988-994*, 1993.

[98] R. Ohba, I. Uehira, S. Kakuma. Interferometric determination of a static optical path difference using a frequency swept laser diode. *Meas. Sci. Technol. 1 500-504*, 1990.

[99] E.R. Peck. A New Principle in Interferometer Design. *Journal of the Optical Society of America, Vol.38 No.1, 66*, 1948.

[100] E.R. Peck. Theory of the Corner-Cube Interferometer. *Journal of the Optical Society of America, Vol.38 No.12, 1015-1024*, 1948.

[101] M. Johnson. Fiber-end interferometric sensor using cooperative retro-reflectors. *Optics Letters, Vol.8 No.11, 593-595*, 1983.

[102] M. Johnson. Stabilized fiber-end retro-reflecting interferometer. *Applied Optics, Vol.23 No.15, 2629-2632*, 1994.

[103] J. Bocki et al. Silicon-strip vertex detector for ATLAS. *ATLAS Internal Note, INDET-NO-075*, Jan. 1995.

[104] S.Roe S.Gadomski and P.Weilhammer. A possible new layout for barrel part of ATLAS ID. *ATLAS Internal Note, INDET-NO-035*, Nov. 1993.

[105] S. Gadomski, G. Hall, T. Hogh, P. Jalocha, E. Nygard and P. Weilhammer. The deconvolution method of fast pulse shaping at hadron colliders. *Nuclear Instruments and Methods A 320 217-227*, 1992.

[106] W.Dabrowski, private communication

[107] T. Bowcock et al. Report on a silicon backup to the MSGC detectors of ATLAS. *ATLAS Internal Note, INDET-NO-069*, 1994.

[108] R. Clifft and A. Poppleton. *ATLAS Internal note, SOFT-NO-9.*, 1994.

[109] S. Haywood. B tagging in ATLAS. *ATLAS Internal Note, in preparation*, 1995.

[110] J. Baines et al. *ATLAS Internal Note, INDET-NO-066*, 1994.

[111] R. Hawkings and A. Weidberg. *ATLAS Internal Note, INDET-NO-052*, 1994.

[112] Particle Data Group. *Phys. Rev. D45*, 1992.

[113] N. Bingefors et al. The Delphi microvertex detector. *Nucl. Instr. Methods A328 447.*, 1993.

[114] S. Haywood. *ATLAS Internal Note, INDET-NO-044*, 1994.

[115] V. Buscher and T. Weidberg. *ATLAS Internal Note, INDET-NO-054*, 1994.

[116] M. Gilchriese. *ATLAS Internal Note, INDET-NO-058*, 1994.

[117] T. Ohsugi et al. Double-sided microstrip sensors for the barrel of the SDC silicon tracker. *Nucl. Instr. Methods A342 16.*, 1994.

[118] E. Spencer et al. A fast-shaping low-power amplifier-comparator IC for silicon-strip detectors. *IEEE N.S. Symp., Norfolk VA,* 1994.

[119] I. Kipnis et al. An analog front-end bipolar-transistor IC for the SDC silicon tracker. *IEEE Trans on N.S. 41, No 4, 1095*, 1994.

[120] J. DeWitt. A pipeline and bus interface chip for silicon-strip detector readout. *SCIPP 93-37 IEEE N.S. Symp., San Francisco, CA, Nov*, 1993.

[121] J. DeWitt et al. Signal-to-noise in silicon microstrip detectors with binary readout. *SCIPP 94-34 IEEE N.S. Symp., Norfolk VA,*, 1994.

[122] E. Barberis et al. Operational experience with the FEE for the LPS silicon microstrip detectors. *SCIPP 92-29, Submitted to Nucl. Instr. Methods*, 1994.

[123] RD20. Analogue MUX for RD20 FE chip readout. *RD20/TN-31.*

[124] Stapnes et al. H8 testbeam results on RD20 readout chips. *ATLAS Internal Note, in preparation*, 1995.