Investigation of Opto-electronic Data Transmission Methods for Experiments at the CERN Large Hadron Collider

by

Steven James Oglesby.

A thesis submitted to the Faculty of Science

of The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Physics and Space Research,

Faculty of Science,

University of Birmingham.

September 1997

Synopsis

This thesis covers work carried out by the author on the investigation of data link technologies applicable for use in Large Hadron Collider detectors. The work has been pursued in particular for use in the readout of the ATLAS Semiconductor Central Tracker. Three technologies have been evaluated: optical analogue readout using passive modulators, optical binary readout using a custom LED and photodiode package, and electrical binary readout over shielded twisted pair cables.

In each case the performance of the links has been evaluated for use in ATLAS SCT environment. Particular requirements are placed on the links in terms of:

- data transmission quality
- reliability
- radiation tolerance
- material
- cost

Each link has advantages and disadvantages. The binary based links were chosen over the analogue link in March 1996. The choice was based on the combination of a large cost saving and more advanced technical solutions for the binary based readout. A choice for the baseline readout for the SCT was made in May 1997. The optical link was chosen over the electrical link because of its low material contribution to the Inner Detector and because of worries about electromagnetic noise radiation and pick-up in the electrical option.

The optical binary link, however, is still not a fully proven solution for the SCT readout and is an ongoing development project.

Author's Contribution

The author was supported throughout this work on a Co-operative Award in Science and Engineering (CASE) studentship, jointly funded by the Particle Physics and Astronomy Research Council (PPARC) and Marconi Defence Systems (MDS).

The work on the analogue optical link was performed as part of the RD23 collaboration. The basic tests on the link were carried out in Birmingham. The Phase 2 device characterisations were carried out by Dr. F.Vasey and myself at CERN. Irradiations tests of the MQW modulators and the testbeam were carried out by Dr. I.Dawson and myself. The workload was split approximately 50:50, with an emphasis on the data analysis for myself. The noise studies on the analogue optical link were carried out at GEC Marconi Defence Systems at Stanmore and in Birmingham.

The irradiation studies of the LEDs and photodiodes in the binary optical link package were carried out at Birmingham and CERN. The Birmingham optogroup members assisted in the measurements for the Birmingham based tests.

The author assisted in the preliminary tests of the STP cables for the binary optical link.

Acknowledgements

I would like to acknowledge the funding I have received from both the Particle Physics and Astronomy Research Council and GEC Marconi Defence Systems that has supported me throughout my Ph.D.

I would like to thank all the people who have assisted me in the work that I have carried out for the last three years: My supervisor Professor John Dowell and all the members of the Birmingham optogroup (past and present): Doctors Ian Kenyon, John Wilson, Jim Homer, Pedja Jovanovic and Ian Dawson, and Harry Shaylor have all helped in the generally smooth running of the project while I have been here; The members of the RD23 collaboration and the ATLAS link groups who have contributed to the development of the different links. In particular I would like to thank Doctor Francois Vasey for coordinating the RD23 project and Doctor Robert Kowalewski for carrying out the same role for the ATLAS link group and Doctor Tony Weidberg for his work on the binary optical link; Nick Try and Nick Green at GEC Marconi Defence Systems who coordinated the work on the analogue optical work and gave advice on the optical noise study; the participants of the ATLAS SCT testbeam and irradiation studies (especially the coordinators for all those night-shifts); the Birmingham University Dynamitron operators for their continual battle to get the most out of the machine; the Birmingham technicians, Roger Harris, Dennis Grant, and Ian McGill for sorting out all the last minute jobs I came up with; and Laurie Lowe and Frank MacDonald for computer things without which there would be many empty gaps in this thesis.

I would like to thank all the people I have shared an office with during my stay at Birmingham. My predecessor working on the project Ken Webster, now Doctor, for his much needed advice during my first few months here. Doctor Ian Dawson for his enthusiastic contribution to the project and for all the lively debates while working as a post-doc here.

Finally Adam Connors, who has shared my office for the last year and who will inherit it (and everything in it), thanks for not working too hard and showing me up.

I would also like to thank all the people who have helped me not do work during my time in Birmingham: I wish all the physics footballers the best of luck in the coming seasons and thank them for the good games of football (although in most games we were pretty bad); all the students at Birmingham, starting from the beer era and working down to the shandy era we have, Chris Dodenhoff, Stuart Robertson, Jim Clayton, Andrew Bayes, Steve Clewer, Paul Davies, Mark Pearce, Keith Norman, Tim Nicholls, Vicky Hudgson, Mark Venables, James Bloomer, Kirstee Hewitt, Andrew Bell, Brian Earl, Adam Connors, Scott Talbot, Shaun Ashby, Mike Thompson, Paul Norman; The honorary students Dave Rees, Steve Hillier, Dave Evans, Lee West, Paul Newman and Ian Brawn; the skiers and drinkers at CERN (too many too mention); the bikers and drinkers here, Simon Hemington, James Fothergill, Mark Holmes and the Cosenors Gym posse; my house-mates Craig Luff, Dan Hodson, Zaffer Iqbal and Pei Leung; and the Birmingham old boys, Paul Cooper and John Wingfield.

Finally, but not least, I would like to thank my family for their love and support.

Steven Oglesby. September, 1997.

Science is a two headed beast, Arthur. One head is good. It brings us advances like aspirin. But the other head - Oh, beware the other head of science, Arthur! It bites!

The Tick (Ben Edlund)

I hope that posterity will judge me kindly, not only as to the things which I have explained, but also to those which I have intentionally omitted so as to leave to others the pleasure of discovery.

René Descartes

Contents

List of Figures

List of Tables

Chapter 1

Physics Overview

1.1 Introduction

Particle physics involves the study of matter at its most fundamental level to further mankind's quest to answer the question "What is the world made of?". Many theoretical models have been proposed to describe the phenomena observed in experiments. However, advances in experimental techniques have often found further substructure to each set of 'elementary' particles found. It is only during the past few decades that the picture has stabilised to one involving a small set of seemingly indivisible constituents, whose interactions are described by just four forces. The Standard Model (SM) provides the framework of this description. Although all experimental data are consistent with the SM, it is not regarded as a complete theory. For instance, it lacks an explanation for the existence of only three generations of quarks and leptons and contains many free parameters such as the particle masses which must be measured empirically. Related to this is a whole section of the theory, known as the Higgs sector, which has never been observed or tested directly.

This chapter presents a brief survey of the SM and beyond.

1.2 The Standard Model

The SM [1] describes the electromagnetic, weak and strong interactions between the elementary particles of nature using a pair of renormalisable quantum field theories. The fourth force in nature, gravity, has only thus far been understood in terms of the classical field theory structure of General Relativity and is not included in the model. However, at the energies at which the particles are studied the gravitational interaction is negligible and can therefore be ignored.

1.2.1 The Elementary Particles

The elementary particles are believed to be structureless and hence point-like. They can be split into two subgroups. The fermions, having spin 1/2 and obeying Fermi-Dirac statistics, combine to form the macroscopic matter we observe in everyday life. The bosons, having integer spin and obeying Bose-Einstein statistics, mediate the interactions between the fermions.

The fermions can be subdivided into two groups of particles: the quarks, which are influenced by the strong force, and the leptons, which are not. There are six particles of each type:

- three quarks carrying fractional charge $-1/3e$ (d, s, b),
- three quarks carrying fractional charge $+2/3e$ (u, c, t),
- three charged leptons (e, μ, τ),
- and three neutral leptons $(\nu_e, \nu_\mu, \nu_\tau)$.

They can be arranged in three symmetrical generations:

$$
\begin{array}{cccc}
 & \multicolumn{3}{c}{\text{Generations}} \\
 & \text{I} & \text{II} & \text{III} \\
\hline
\text{Quarks} & \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} \\
\text{Fermions} & \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} & \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} & \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}
\end{array}
$$

The properties of the first generation are replicated by the subsequent generations. However, the masses of the particles increase across the generations with the exception of the neutrinos which are consistent with being massless.

The forces between the fermions are manifestations of the exchange of virtual field quanta (the bosons). Each field and therefore force has an associated exchange particle : the electromagnetic force is mediated by a massless photon, the weak force by massive $W^{\pm}$ and Z^{o} bosons, and the strong force by an octet of massless gluons which carry a colour charge. The range of the electromagnetic and weak forces is governed by the uncertainty principle. To emit a boson the fermion must borrow energy for a short time. For a massive particle a larger energy must be borrowed which is only possible for a shorter time. The weak force couples to particles with the same strength as the electromagnetic force, but only works over very short distances. The strong force couples much more strongly than the other forces. However, its range is limited to short distances by the self-interaction of the gluons.

Each elementary particle has an anti-particle partner with opposite charge and magnetic moment [2].

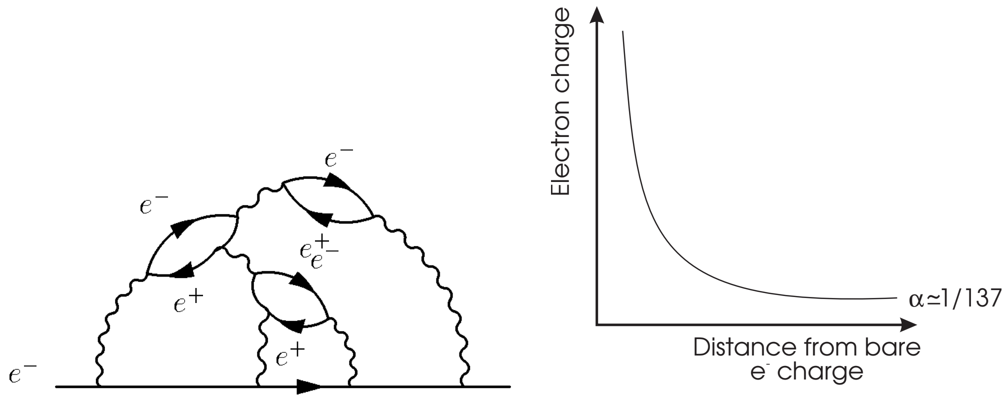
1.2.2 Gauge Field Theories

The interactions of the particles are described by gauge field theories. Two important properties of these theories are those of renormalisability, which allows predictive calculations to be made, and gauge invariance, which is the key to describing the interactions.

Renormalisation

Without renormalisation the calculations made with the field theories would be of no practical value because the physical predictions would be divergent. The process of renormalisation involves a reparameterisation of the basic equations of the theory in terms of unobservable bare quantities such as the bare mass and bare charge. All the infinities in the theory are then grouped together in the bare quantities leaving the physical observables calculable.

This process can be visualised for the case of the electron charge as shown in figure 1.1. In the quantum field theory description the electron will be continually emitting and re-absorbing photons which can annihilate to form e^+e^- pairs. An example of a possible charge cloud formed by this process is shown in the diagram. The attraction of opposite charges leads to the e^+s being preferentially closer to the electron resulting in the charge cloud being polarised with the positive charges closer to the electron. The bare charge of the electron is therefore shielded. It is only if we fully penetrate the charge cloud that we will observe the bare charge of the electron. Experimentally the electron charge is measured using a test charge. It is observed that the charge of the electron increases from the physical charge, e, measured at infinity as the test charge is brought closer to the electron. To replicate this effect in the theory, the value of the charge is treated as a fixed parameter and the coupling constant α which represents the coupling strength of the photon to charged particles is allowed to 'run', increasing as the magnitude of the square of the four-momentum transfer of an interaction (Q^2) increases.



(a) Charge cloud. (b) Measured charge.

Figure 1.1: Shielding of the electron bare charge.

Gauge Invariance

Within gauge field theories the fermions are described by complex fields and therefore possess an unobservable phase. The Lagrangian of the system (a mathematical term describing the energy-momentum dynamic of the system) is constructed from these fields. Since the physical variables do not depend on the phase, it seems reasonable to to require a theory which is invariant under changes of this phase, historically called gauge rotations. Demanding such gauge invariance under transformations which are arbitrarily different at all space-time points, however, results in noticeable local changes in the field. The situation can be recovered by introducing a further set of fields into the theory called the gauge fields, which describe the (gauge) bosons. The local changes due to a gauge transformation applied to the fermionic fields are then compensated by the same transformation applied to the gauge fields. For example, a local change of phase of the electron field can be compensated for by the emission of a photon.

1.2.2.1 Quantum Electro-Dynamics

The first successful gauge field theory was that of Quantum Electro-Dynamics (QED) [3]. Demanding gauge invariance of a single charged fermion to the phase transformation of the group U(1) introduced the gauge field of the photon. The theory reproduces the electromagnetic interaction with incredible precision.

1.2.2.2 Quantum Chromo-Dynamics

Following the success of QED, Quantum Chromo-Dynamics (QCD) [1, 4, 5] was formulated to describe the strong interaction of the quarks. For perturbative calculations to be made the coupling constant of the interaction must be small. Thus higher order corrections with strength α^2 can be treated perturbatively. This is reasonable for QED where the coupling constant $\alpha \approx 1/137$, however, for the strong force the coupling constant $\alpha \approx 1$. It would therefore seem impossible to form a perturbative theory for the strong interaction. However, it has been possible to construct a field theory due to the nature of the gluons.

In QCD the quarks carry a 'colour' charge associated with the SU(3) group. The colours are designated red (R), green (G) and blue (B). This gives the quarks an extra three degrees of freedom. This has been measured experimentally in the cross-section for $e^+e^- \rightarrow q\overline{q}$ which is three times larger than would be expected if no colour charge existed. Demanding gauge invariance of the theory produces an octet of gauge fields, the gluons, which are the mediators of the strong force. The difference from QED comes from the fact that the gluons themselves carry a colour charge (cf. photons, which do not carry an electromagnetic charge) and can therefore couple to themselves at three and four gluon vertices. The theory is termed non-Abelian. This leads to a phenomenon known as asymptotic freedom, where the strength of the attractive force between two quarks increases with their separation. The quarks are therefore confined to bound states as observed in experiment.

To reproduce the behaviour of asymptotic freedom the coupling constant for the strong interaction is made to 'run'. However, the running is opposite to that of QED and, although for low Q^2 interactions the coupling constant is still close to one, it is found that for high Q^2 interactions the coupling constant is lower and sensible perturbative calculations can be made.

1.2.2.3 Electroweak Theory

There are two obstacles to producing a gauge theory to describe weak interactions. The first is the different properties of the neutral current Z, which couples to both left and right-handed particles, and the charged Ws, which couple to only left-handed particles. Such a difference in boson properties cannot be produced by demanding the gauge invariance of a simple theory. The second obstacle is the fact that the gauge bosons are massive, but any addition of explicit mass terms will destroy the gauge invariance of the theory. The Electroweak theory [6–8] avoids these problems by the introduction of some ingenious constructions. The first problem was solved by considering the weak and electromagnetic interactions jointly, and the second by introducing the Higgs mechanism to the combined theory.

Invariance to the gauge transformations of the combined groups $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, associated with the conserved quantities of weak isospin T and weak hypercharge Y respectively, is demanded by the theory. This results in the introduction of a triplet of gauge fields W^1, W^2 and W^3 coupled to the weak isospin current, and a single B° coupled to the weak hypercharge current. An important property of the theory is that the weak isospin current only couples to left-handed particles, this is denoted by the L subscript on the SU(2) group. The left and right-handed particles are therefore treated differently in the theory and are written as SU(2) doublets and SU(2) singlets respectively:

$$\text{Doublets (left-handed):} \quad \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \begin{pmatrix} u \\ d' \end{pmatrix}_L \begin{pmatrix} c \\ s' \end{pmatrix}_L \begin{pmatrix} t \\ b' \end{pmatrix}_L$$

$$\text{Singlets (right-handed):} \quad (e)_R \ (\mu)_R \ (\tau)_R \ (u)_R \ (d)_R \ (c)_R \ (s)_R \ (t)_R \ (b)_R$$

Note that right-handed neutrinos are not found in nature.

The quark eigenstates d', s' and b' which appear in the doublets do not correspond to the physical mass eigenstates d, s and b, but consist of linear combinations of these states, as expressed by the Kobayashi-Maskawa matrix [9]. The 3×3 mixing matrix contains three real mixing angles and a complex phase factor $e^{i\delta}$. The phase factor results in a theory which necessarily violates CP-invariance (the combination of charge (C) and parity (P) conjugation). The size of the violation is not however predicted in the SM.

Linear combinations can be made of the W^1 and W^2 fields that are associated with the weak charged currents, according to,

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp W^2). \tag{1.1}$$

However, since the Z^o is observed to couple to both left and right-handed particles it cannot be simply associated to the neutral W^3 field which couples only to left-handed particles. It is actually formed by a linear combination of the W^3 field with the neutral B^o field, which couples to both left and right-handed particles. By careful choice of the combination of these two fields bosons corresponding to the Z^o boson (Z) and the photon (A) can be constructed,

$$\begin{aligned} Z &= -B^o \sin\theta_W + W^3 \cos\theta_W \\ A &= B^o \cos\theta_W + W^3 \sin\theta_W, \end{aligned} \tag{1.2}$$

where θ_W is the weak mixing angle.

1.2.2.4 Spontaneous Symmetry Breaking and the Higgs Mechanism

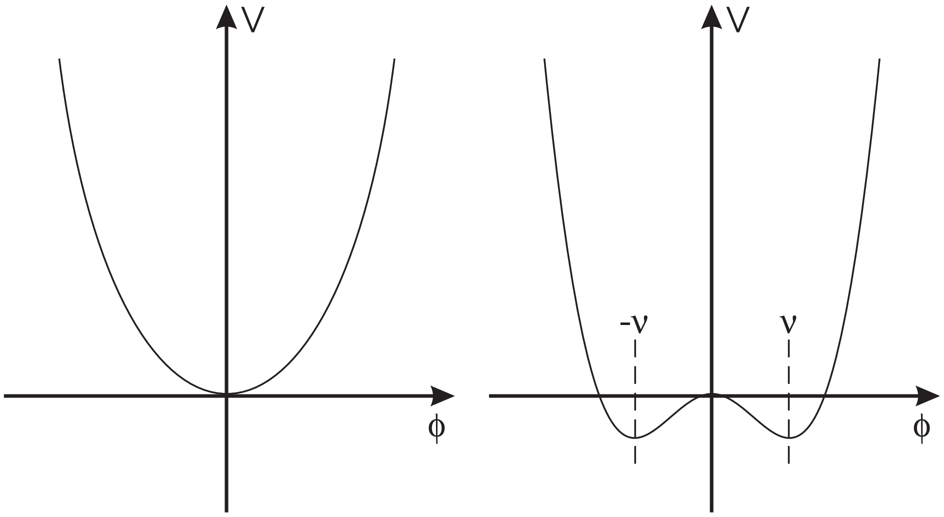
We now turn to the problem of giving the particles masses without destroying the gauge invariance of the theory. This is achieved using the procedure of spontaneous symmetry breaking [1]. A Lagrangian is constructed that possesses a symmetry which is not seen in the physical vacuum state. The symmetry of the Lagrangian is therefore said to be spontaneously broken by the choice of the vacuum state. The overall Lagrangian is gauge invariant, as required by the gauge theory, however the vacuum is not, allowing mass terms to be generated.

Consider the example of a Lagrangian L describing a scalar field ϕ with kinetic energy T and potential energy V,

$$L = T - V = \frac{1}{2}(\partial_\mu \phi)^2 - \Big(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda^2\phi^4\Big), \tag{1.3}$$

where μ and λ are coupling parameters measuring the strength of the interaction.

Clearly L is symmetrical under the operation that replaces ϕ with $-\phi$. However, by varying the coupling parameter μ the two types of potential shown in figure 1.2 can be produced. For the case $\mu^2 > 0$ the vacuum corresponds to the absence of the field ($\phi = 0$) but for the case $\mu^2 < 0$ the potential has minima at non-zero values of the field ($\phi = \pm\nu$). By choosing a vacuum value of either ν or $-\nu$ for the Higgs field, the symmetry is spontaneously broken.



(a) $\mu^2 > 0$, $\lambda > 0$.

(b) $\mu^2 < 0$, $\lambda > 0$.

Figure 1.2: Possible simple Higgs potentials.

In the SM the masses are generated by the process of spontaneous symmetry breaking using the Higgs mechanism [10]. Four new Higgs fields are introduced into the Lagrangian, arranged as a weak isospin doublet of complex scalar fields with hypercharge 1,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^o \end{pmatrix} \quad \text{with} \quad \begin{array}{l} \phi^+ \equiv (\phi_1 + \imath\phi_2)/\sqrt{2} \\ \phi^o \equiv (\phi_3 + \imath\phi_4)/\sqrt{2} \end{array}$$

As before, the choice of the couplings allows a Higgs field to be produced with a non-zero vacuum expectation value thereby endowing the vacuum with weak isospin and hypercharge. In fact, there exists a continuous set of degenerate ground states, $|\phi_o|$, which differ from one another in the phase of their weak isospin. The correct choice of a particular ground state

ϕ_o from this set allows the breaking of the $SU(2)_L \times U(1)_Y$ symmetry while leaving the U(1) symmetry of QED intact. This choice is known as the unitarity gauge. Mass terms are therefore generated for the W^+, W^- and Z^o bosons, whilst the photon remains massless. During this process three of the scalar fields that form the Higgs scalar doublet disappear. This leaves one scalar field which itself acquires a mass term and is found to describe a massive, spin-zero boson: the Higgs boson (H).

The same Higgs doublet that generates the boson masses is also sufficient to give masses to the leptons and quarks. Terms are added to the electroweak Lagrangian that describe the interaction of the Higgs doublet to the fermions. When the symmetry of ϕ is broken these terms resolve into two components. The first component corresponds to the mass term of the relevant fermion, the other describes the coupling of the Higgs boson to that fermion. The fermions couple to the Higgs boson with a strength proportional to their mass.

Although the Higgs mechanism describes the process of mass generation, the only directly predicted mass ratio is that between the W and Z masses (M_W and M_Z respectively),

$$M_W = M_Z \cos\theta_W, \tag{1.4}$$

which is found to agree well with experiment. The masses of the fermions and the Higgs boson appear in the theory only as free parameters. We can however place some constraints on the mass of the Higgs boson (m_H). An upper limit arises from the self-interaction strength of the Higgs doublet. For meaningful calculations to be made using perturbation theory m_H must be smaller than $\sim 1TeV$. A lower limit can also be derived. For the electroweak model to be stable the Higgs field must produce a minimum above the electroweak ground state. This requires $m_H > 95GeV$ (1σ) in the SM [11].

1.3 Beyond the Standard Model

1.3.1 Grand Unified Theories

The electroweak theory describes the weak and electromagnetic interactions within a single theoretical framework which shows impressive agreement with experiment. However, the degree to which it unifies the interactions is open to question. The $SU(2) \times U(1)$ group is a product of two separate gauge transformations: the SU(2) group with coupling strength

g and the U(1) group with coupling strength g'. The ratio of the coupling strengths is not predicted and must be measured experimentally.

If we wish to form a theory which relates the coupling strengths, the SU(2) and U(1) symmetry groups must be embedded into a larger group, G. It would seem natural also to include the SU(3) group of the strong interaction in this unification. The relationship between the groups can be written symbolically as,

$$G \supset SU(3) \times SU(2) \times U(1). \tag{1.5}$$

A grand unified theory (GUT) can be constructed by demanding invariance under gauge transformations of the group G. All the coupling constants (g and g' of the electroweak theory, and α_s, the colour coupling constant) would be related to a universal coupling constant, g_{GUT}.

Figure 1.3 shows the running of the three coupling constants with the four-momentum transfer (Q) of an interaction. It can be seen that at some large energy, $M_X \sim 10^{15} GeV$, all three interactions could possess the same coupling constant. Interactions above this energy are described by the unified symmetry group, G. Below this energy however, G is spontaneously broken into the groups SU(3) $\times$ SU(2)$_L$ $\times$ U(1)$_Y$. Present day experiments observe interactions with $Q \simeq 30 GeV$.

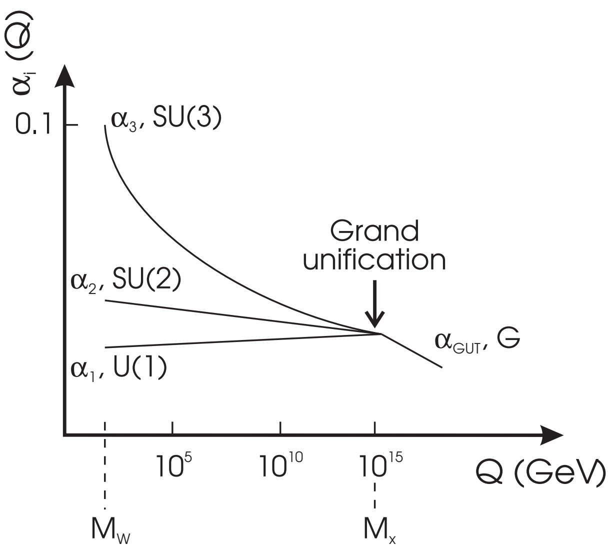


Figure 1.3: The variation of the coupling constants with Q. The speculative unification of the strong and electroweak interactions occurs at $Q \sim M_x$. ($\alpha_i \equiv g_i/4\pi$).

The simplest choice for the unified group, G, was shown to be the SU(5) symmetry group, by Georgi and Glashow [12] in 1974. There are a total of 24 gauge bosons generated by this model. After spontaneous symmetry breaking these are resolved into the 12 gauge bosons of the SM plus an additional 12 'leptoquarks' which couple the leptons to quarks. Interactions violating lepton and baryon number conservation are therefore possible, such as the decay of the proton.

However, GUTs in general suffer from the so-called hierarchy problem. The problem arises from the different energy scales of the different interactions: from the QCD scale of order $\simeq 1GeV$, through the electroweak scale of $M_W \simeq 100GeV$, to the GUT scale of $M_X \simeq 10^{15}GeV$. Higgs scalars in the electroweak theory were introduced to generate mass by the spontaneous breaking of the gauge symmetry. Similarly, Higgs scalars will be required to break the SU(5) symmetry, but with mass $\simeq M_X$ rather than $\simeq M_W$. The light Higgs will receive radiative corrections to their mass from the GUT Higgs, resulting in the H, W, and Z masses of the electroweak sector becoming unstable. This can only be avoided by arranging clever cancellations of the radiative corrections working to a precision of 1 part in 10^{13} $(M_W : M_X)$. Such precise cancellations are not possible in the minimal Georgi-Glashow GUT, or many of the more complicated theories.

1.3.2 Supersymmetry

Supersymmetry (SUSY) [1, 4] theories appear to provide a solution to the hierarchy problem. All SUSY theories propose a symmetry between the fermions and the bosons. New supersymmetric partners, named sparticles, are postulated for all the fundamental particles as indicated in table 1.1.

In SUSY theories the radiative corrections applied to the light Higgs masses must include contributions from the the coupling of the light Higgs particles to both the heavy Higgs bosons and heavy sHiggs fermions. Since the radiative corrections to the bare mass terms from virtual boson loops and virtual fermion loops are of opposite sign, the contributions will cancel exactly. The light Higgs and SM boson masses therefore remain stable.

The symmetry between the particles and sparticles must be broken at some stage, to explain why no sparticles have yet been observed. However, for SUSY theories to solve the hierarchy problem this must occur at an energy scale similar to the electroweak Higgs sector. The new SUSY particles must therefore have masses $\lesssim 1TeV$.

Table 1.1: Supersymmetric particles.

Particle	Spin	Sparticle	Spin
Quark q	$\frac{1}{2}$	sQuark $\tilde{q}$	0
Lepton l	$\frac{1}{2}$	sLepton $\tilde{l}$	0
Photon γ	1	Photino $\tilde{\gamma}$	$\frac{1}{2}$
Gluon G	1	Gluino $\tilde{G}$	$\frac{1}{2}$
W	1	Wino $\tilde{W}$	$\frac{1}{2}$
Higgs H	0	sHiggs $\tilde{H}$	$\frac{1}{2}$

Most SUSY theories require the associative production of the new particles. This leads to the idea of a conserved quantum number, R (+1 for particles and −1 for sparticles). This implies that at least one SUSY particle (the lightest) must be stable. Sparticles would pass through a detector undetected, however, their presence should be evident. For instance, the production of a squark could be manifest in the decay $\tilde{q} \rightarrow q\tilde{\gamma}$. The missing photino in the decay products would result in a large momentum imbalance in the event. To date, no evidence has been observed for the existence of SUSY particles up to masses of $80GeV$.

The Large Hadron Collider is designed to allow the investigation of physics at the TeV energy scale. It will help to further our understanding of the SM or show us that it is flawed.

Chapter 2

The Detector System

2.1 Introduction

CERN, the European Laboratory for Particle Physics, houses the leading particle accelerator facilities in Europe. It is the home of the Large Electron Positron Collider (LEP) which has led the way in testing our understanding of the Standard Model (SM). However, by the end of the decade the measurements possible at LEP energies will be largely completed. The precision of the LEP measurements points the way towards exciting discoveries that may be made at higher energies as required to complete our understanding of the SM and its possible extensions.

The Large Hadron Collider (LHC) [13, 14] is the next generation of particle accelerator. It will be built in the existing 26.7km circumference LEP tunnel, and is scheduled to start operating in the year 2005.

2.1.1 The Large Hadron Collider

The LHC will consist of two storage rings that will be filled with protons delivered from the Super Proton Synchrotron (SPS) and its pre-accelerators at $0.45TeV$. The layout of the LHC and associated accelerator complex is shown in figure 2.1. The two counter-rotating proton beams are accelerated to 7-on-7 TeV by Radio Frequency (RF) cavities. The magnetic field required to bend the protons around the LHC ring at these energies is $8.36Tesla$. Niobium-titanium superconducting magnets operating at $1.9K$ are used in a unique configuration that produces the magnetic fields required for both beams. The two beams will be maintained for several hours, colliding at the experimental areas with a centre of mass energy of $14TeV$, until they become so degraded that the machine must be emptied and refilled. An alternative mode of operation collides beams of heavy ions in place of the protons.

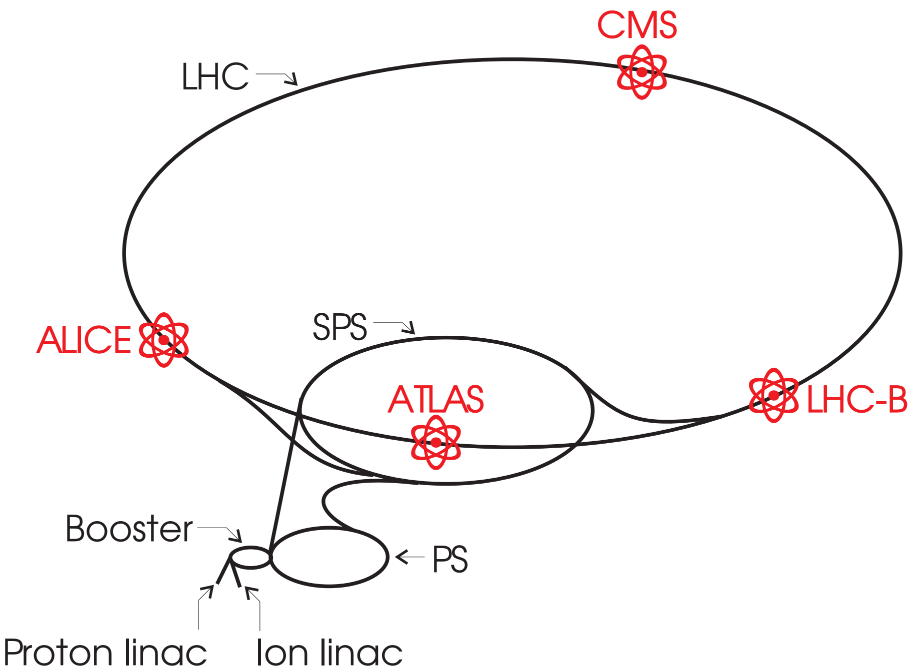


Figure 2.1: The layout of the LHC accelerator complex.

2.1.2 Physics Studies

Physics studies of interest at the LHC include the search for leptoquarks, the search for the lightest stable sparticle, and the search for evidence of quark compositeness. However, the primary interest is in the investigation of spontaneous symmetry breaking at the electroweak scale.

Possible manifestations of the spontaneous symmetry breaking mechanism include the existence of the Standard Model Higgs boson (H) or the quintet of Higgs particles ($H^{\pm}$, h, H, and A) predicted by the Minimal Supersymmetric extension of the Standard Model (MSSM). In order to cover the full possible mass range for the SM Higgs, the search strategy requires detectors at LHC to be sensitive to a variety of final states:

$$
\begin{array}{lr}
H \rightarrow b\bar{b} & \text{mass range } 80 < m_H < 100GeV; \\
H \rightarrow \gamma\gamma & 90 < m_H < 150GeV; \\
H \rightarrow ZZ^* \rightarrow 4l^{\pm} & 130 < m_H < 180GeV; \\
H \rightarrow ZZ \rightarrow 4l^{\pm}, 2l^{\pm}\ 2\nu & 180 < m_H < 800GeV; \\
H \rightarrow WW, ZZ \rightarrow l^{\pm}\nu\ 2jets, 2l^{\pm}\ 2jets & m_H <\sim 1TeV.
\end{array}
$$

In addition, the search for the MSSM Higgs will require sensitivity to processes such as:

$$
\begin{array}{ll}
A \rightarrow \tau^+\tau^- & \rightarrow e\mu + \nu s \\
 & \rightarrow l^{\pm} + hadrons + \nu s; \\
H^{\pm} & \rightarrow \tau^{\pm}\nu \\
 & \rightarrow 2jets.
\end{array}
$$

The channels available for the search for the Higgs bosons are limited to those visible above the dominant hadronic background present at LHC. The expected cross-sections for the Higgs processes are small across the whole mass range, requiring the LHC to run at a high luminosity ($10^{34}cm^{-2}s^{-1}$).

Further models for symmetry breaking in the electroweak sector exist, all of which predict new particles with masses up to $\sim 1TeV$. The LHCs ability to reach these energy scales will inevitably further our understanding of the mechanism for symmetry breaking, either through the discovery of new particles or the revelation of their absence.

During the first three years of operation the LHC will run at a lower than maximum luminosity of $10^{33}cm^{-2}s^{-1}$ and will allow the study of top and bottom-quark physics. The SM predicts CP violation due to the three quark generations (see chapter 1). The violation of a combined charge (C) and parity (P) conjugation was observed for the weak interaction in the decay of neutral kaons [15]. Measurements of the B-meson decays:

$$
\begin{array}{ll}
B_d^o & \rightarrow J/\psi K_s^o, \\
B_d^o & \rightarrow \pi^+\pi^-, \\
B_s^o & \rightarrow J/\psi\phi,
\end{array}
$$

will allow a better determination of the quark flavour mixing parameters which are not predicted in the SM.

Studies of rare decays of the b and t quarks will also be possible over the three year low luminosity operational period.

The LHC has an additional mode of operation which is to accelerate and collide beams of heavy ions, such as Lead, with a centre of mass energy of $1250 TeV$. At high temperatures and densities a phase change from 'normal' hadronic matter to a deconfined state, known as a guark-gluon plasma (QGP), is theorised. Discovery of a QGP and measurement of its thermodynamic properties allows the study of strong interactions under different conditions to normal experiments. It has also been proposed that the early Universe existed as a QGP, so measurements in the laboratory could further our understanding of cosmology.

There will be four detectors at the LHC: two general-purpose proton-proton experiments named **A** **T**oroidal **L**HC **A**pparatu**S** (ATLAS) [16] and **C**ompact **M**uon **S**olenoid (CMS) [17], a third experiment named **A** **L**arge **I**on **C**ollider **E**xperiment (ALICE) [18] is designed for the study of heavy ion physics, and the final experiment LHC-B [19] is proposed to study beauty-physics. This thesis concerns the design of the ATLAS experiment. However, the requirements of all the detectors are largely set by the LHC operating conditions and are therefore similar.

2.2 The ATLAS Detector

The ATLAS detector is a general-purpose proton-proton experiment designed to exploit the full physics potential of the LHC. The broad scope of the optimisation required for the physics tasks mentioned in section 2.1.2 is expected also to maximise the detector's potential for the discovery of new, unexpected physics.

The proposed ATLAS detector is shown in figure 2.2. A particle travelling outwards from the interaction point first meets the inner tracking detector. This detector system consists of an inner section of high-resolution pixel and strip detectors, surrounded by an outer section of continuous straw-tube tracking detectors. Encompassing the inner detector is a super-conducting solenoid which produces a 2T magnetic field over the inner detector region allowing the particle momentum to be determined.

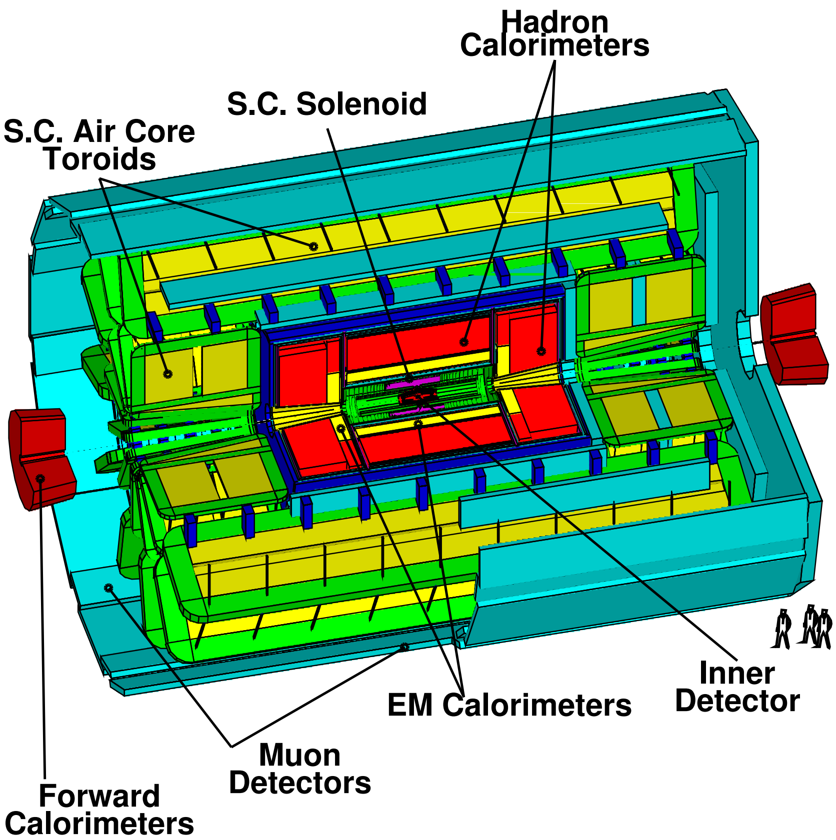


Figure 2.2: Schematic of the proposed ATLAS detector.

The inner tracking detector is followed by the ATLAS calorimetry. Liquid Argon (LAr) technology is used for the electromagnetic calorimetry in the form of a barrel and two endcap calorimeters. The LAr endcaps also form part of the hadronic calorimeter along with a scintillator tile calorimeter, which forms a barrel around the LAr calorimeters for the full length of the detector.

The overall dimensions of the ATLAS detector, 11m in radius and 42m long, are defined by the muon spectrometer which forms the final section of the detector. The muon spectrometer consists of a second magnet system formed of three super-conducting air-core toroids: a barrel section and two end-caps. Three layers of tracking chambers are placed throughout the magnetic volume.

The detector subsystems are described in more detail in sections 2.2.2 to 2.2.5. Preceding this is a brief review of the coordinate systems used in these descriptions.

2.2.1 ATLAS Coordinate Systems

The three coordinate systems used in describing the detector are:

- Cartesian coordinates (x, y, z), where z lies along the beam axis, and x and y are perpendicular to this in the horizontal and vertical directions respectively;
- cylindrical coordinates (R, ϕ, z), where z lies along the beam axis, R is the perpendicular distance from this axis, and ϕ is the azimuthal angle about the z-axis;
- pseudo-rapidity coordinates (r, η, ϕ), where r is the distance from the origin (defined as the interaction point), ϕ is the azimuthal angle about the beam axis, and η is the pseudo-rapidity, defined as,

  $$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.1}$$

  θ is the angle made to the beam axis. Pseudo-rapidity is the high-energy approximation of the Lorenz-invariant variable rapidity.

2.2.2 The Inner Tracking Detector

The inner tracking detector [20] must perform several basic functions in order for ATLAS to meet the physics requirements set out previously in section 2.1.2. During high luminosity running these functions are:

- Tracking coverage over the pseudo-rapidity range $|\eta| \leqslant 2.5$;
- Efficient momentum measurement of leptons over the range $|\eta| \leqslant 2.5$;
- Tagging of b-jets;
- Measurement of the primary vertex position.

At low luminosity, reconstruction of secondary vertices from b-quark and τ decays is also required.

A cross-sectional view of the inner tracking detector is shown in figure 2.3. The tracker is split into three regions, a barrel and two identical forward regions. It is further split into three subsystems: the Pixel Detector, the Semiconductor Central Tracker (SCT), and the Transition Radiation Tracker (TRT). The design parameters of these systems are listed in table 2.1.

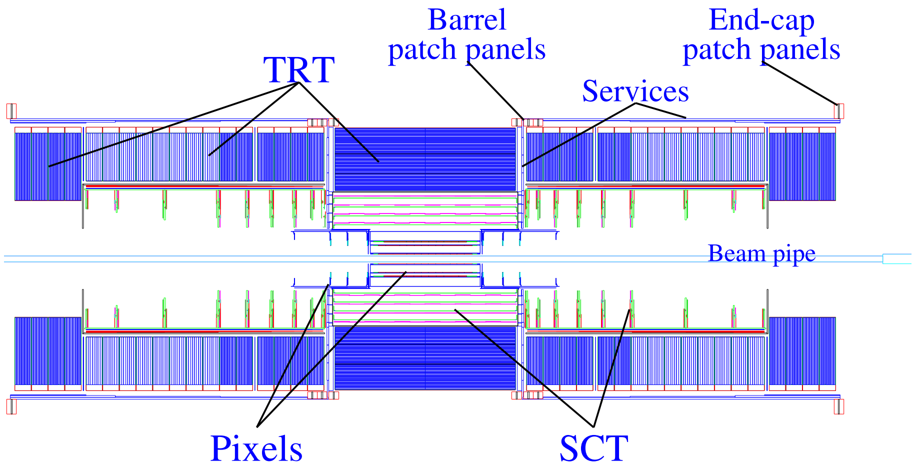


Figure 2.3: A cross-section of the ATLAS inner tracking detector through the beam axis [20].

2.2.2.1 The Pixel Detector

The very high track density in the Inner Detector necessitates the use of high granularity detectors, additionally the spatial and momentum resolution targets demand high precision in the measurements. The innermost layers of the detector are therefore made up of pixel

Table 2.1: Parameters of the Inner Tracking Detector.

System	**Position**	**Area** (m^2)	**Resolution** σ (μm)	**Channels** (10^6)	η **coverage**
Pixels	1 removable/replaceable barrel layer	0.2	$R\phi = 12, z = 60$	16	± 2.5
	2 barrel layers	1.4	$R\phi = 12, z = 60$	81	± 1.7
	4 end-cap disks on each side	0.7	$R\phi = 12, z = 60$	43	$1.7 - 2.5$
Silicon strips	4 barrel layers	34.4	$R\phi = 16, z = 580$	3.2	± 1.4
	9 end-cap wheels on each side	26.7	$R\phi = 16, R = 580$	3.0	$1.4 - 2.5$
TRT	Axial barrel straws		170 (per straw)	0.1	± 0.7
	Radial end-cap straws		170 (per straw)	0.32	$0.7 - 2.5$
	36 straws per track				

detectors to provide the highest possible granularity. Three barrel pixel layers and 4 disks in each forward region will be installed for the initial low luminosity running. The pixel arrays have elements of size $50\mu m$ in $R\phi$ by $300\mu m$ in z. The readout chips contain individual electronics for each pixel and are bump bonded to the detector substrates to achieve the required density of connections.

Radiation damage will eventually render the first barrel pixel layer inoperable. This layer was originally designed for improved secondary vertex finding during the low luminosity running. However, new physics simulations have demonstrated the value of the improved b-tagging offered by this layer for Higgs and super-symmetric searches. It is therefore possible that this layer will be replaced after several years running rather than removed.

2.2.2.2 The Semiconductor Central Tracker (SCT)

At larger R, where lower granularity is acceptable, silicon strip detectors are used. The barrel uses four layers of silicon microstrip detector modules. The modules are made of back-to-back detectors which have their strips running at $40mrad$ angle relative to each other. The modules are mounted with the one set of strips running along the z-axis (the other at the small stereo angle) so that they measure the ϕ coordinate of a particle directly from the hit strip and the z coordinate by using the back to back pair; the R coordinate is known from the location of the detector. The forward detectors are of similar construction with the strips running radially and therefore measure ϕ directly, R using the stereo angle, and z from the detector location. A greater precision in the ϕ coordinate than in z and R is required, as ϕ lies in the bending plane of the magnet. The SCT environment and modules are explained in context of the data readout in chapter 3.

The required combined momentum resolution of the precision tracking elements is $\Delta p_T/p_T < 0.3$ at $p_T = 500GeV$ for $|\eta| < 2$. It is relaxed to $\Delta p_T/p_T < 0.5$ at $p_T = 500GeV$ for $|\eta| = 2.5$. The Inner Detector cavity size and magnetic field strength then result in a required detector resolution of $\sim 25\mu m$ in $R\phi$ (the bending plane). A lower resolution in the non-bending plane of $\sim 600\mu m$ can be tolerated. The requirement comes from the need to distinguish the charge of particles (eg. to distinguish electrons and positrons).

2.2.2.3 The Transition Radiation Tracker (TRT)

The number of precision layers in the inner tracker is limited due to their need for local electronics to achieve the required readout speed. The electronics increase the cost, power dissipation and radiation thickness of the detector. This has led to the choice of a continuous straw tube tracker at the outer radius of the inner tracker. The straw tubes composing the TRT are basically drift tubes which can operate at high rates by virtue of their small $4mm$ diameter. The tubes are filled with xenon gas adding an electron identification capability by allowing detection of transition radiation photons (hence TRT) created in radiators which are interleaved with the straw tubes. In the barrel, the tubes run parallel to the z-axis and therefore measure the $R\phi$ coordinate of a hit. In the forward regions they are oriented radially to measure ϕ and z directly, with R determined indirectly from the first and last straws hit giving the trajectory of the track. The spatial resolution and granularity of this detector is less than the precision layers, but this is acceptable at larger radii where the track density is lower. Each track in the range $|\eta| \leqslant 2.5$ must cross at least 36 straws, allowing tracks to be followed continuously, an obvious advantage for pattern recognition.

2.2.3 Calorimetry

The function of the calorimetry system is to measure the energy of photons, electrons and hadrons, as well as determining the missing transverse energy (transverse energy is defined by $E_T = Esin\theta$). It is achieved by the 'total absorption' of the particles in suitably instrumented blocks of matter. The incident particle cascades in the calorimeter material producing secondary particles of decreasing energy. Sensors distributed in the absorber measure the multiplicity of the showers which is proportional to the shower energy.

Electrons and photons interact electromagnetically, whereas hadrons are degraded by undergoing strong interactions. The differing characteristic interaction lengths and products of the two processes require different optimisations. The calorimetry is therefore split into hadronic and electromagnetic calorimetry systems.

Figure 2.4 shows a schematic of the proposed ATLAS calorimetry. The inner barrel and endcap sections of the calorimeter use liquid argon (LAr) technology. An outer barrel scintillator tile calorimeter surrounds these detectors and determines the size of the calorimetry system, 4.25 m outer radius and 12.2 m length. Electromagnetic calorimetry is performed within the LAr sections. Hadronic calorimetry is performed mainly within the tile calorimeter, with the LAr endcaps extending the measurement to higher η. Calorimetry is extended to very high η using a dense integrated forward calorimeter.

The barrel cryostat contains both the barrel electromagnetic calorimeter and the solenoidal coil which provides the inner tracking detector 2T field. The coil is placed in front of the calorimeter and is integrated with the inner wall of the cryostat to minimise the material before the calorimeter. Two endcap cryostats contain the electromagnetic and hadronic endcap calorimeters as well as the integrated forward calorimeter.

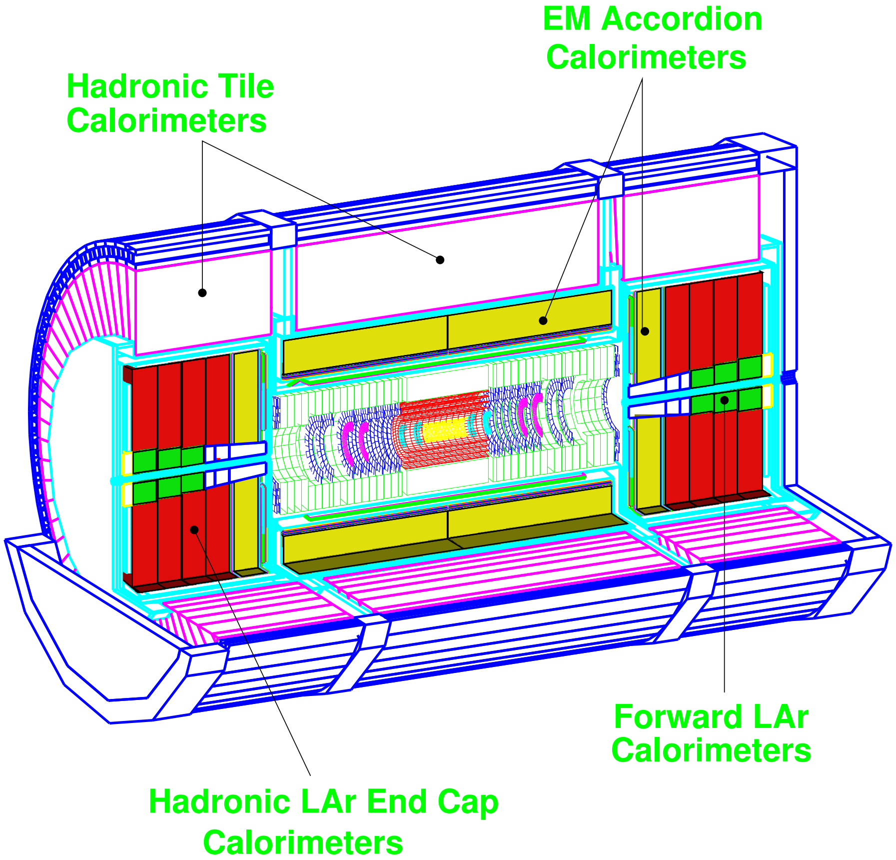


Figure 2.4: 3-D view of the ATLAS calorimetry [16].

2.2.3.1 The Electromagnetic Calorimeter

The electromagnetic calorimeters use lead absorber plates in a liquid Argon ionisation medium. There are 1024 accordion shaped absorber plates, which are oriented radially in ϕ

with the waves running outwards from the centre. Readout electrodes are located midway between the absorber plates.

The barrel electromagnetic calorimeter covers the pseudo-rapidity range $|\eta| \leqslant 1.4$. This coverage is extended to $|\eta| \leqslant 3.2$ by the endcap electromagnetic calorimeters. Accordion shaped lead absorbers are again used, this time arranged radially with the waves running parallel to the beam axis.

Stringent requirements are placed on the electromagnetic calorimetry by the physics processes at LHC in terms of acceptance, dynamic range, particle identification, energy resolution, and direction measurement. The search for rare processes, such as $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4e$, requires excellent acceptance. A large rapidity coverage is therefore required. Simulation studies have shown that the coverage provided by the electromagnetic calorimeter allows high efficiency reconstruction of essentially all of the interesting high transverse momentum (p_T) physics channels.

The electromagnetic calorimeter must be capable of identifying and reconstructing electrons and photons over a wide energy range. At the low end of this range the calorimeter is required to reconstruct electrons with $E_T \gtrsim 2GeV$ produced in semi-leptonic b-decays. The upper limit comes from the signature of $Z' \rightarrow ee$ and $W' \rightarrow e\nu$ decays with energy up to $5TeV$. The signal per readout cell at this upper mass limit combined with the noise per cell results in a dynamic range requirement of 16 bits in the electronics.

The most demanding channels for the energy resolution are again $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4e$. A resolution $\Delta E/E \leqslant 10\%/\sqrt{E} \oplus 1\%$ is considered to provide adequate sensitivity to these channels after a few years integrated luminosity at LHC. The accordion geometry allows the use of fast signal shaping reducing the noise term due to pile-up from minimum bias events. The total noise resolution for an electromagnetic shower has an r.m.s of $400MeV$ at design luminosity and $\eta = 0$.

The optimum segmentation of the electromagnetic calorimeter is a balance between the performance issues such as electron and photon identification, position resolution, and pile up and electronic noise contributions against cost and technical constraints such as the routing of signals out of the cryostats. Hadronic jets form the dominant background to electrons and photons. However, the showers produced are longer and broader than those of isolated electrons and photons. The longitudinal and transverse segmentation of the calorimeter measure the shape of the shower and therefore form an essential tool for the rejection of jet backgrounds. The most stringent requirement comes from the search for the $H \rightarrow \gamma\gamma$ decay which requires a rejection factor of $\sim 10^4$ against single jets. The segmentation of the barrel electromagnetic calorimeter is $\Delta\eta \times \Delta\phi \approx 0.025 \times 0.025$ in the second sampling region and $\Delta\eta \times \Delta\phi \approx 0.025 \times 0.05$ in the third sampling region. At larger η, in the endcap calorimeters, a coarser segmentation of $\Delta\eta \times \Delta\phi \approx 0.05 \times 0.05$ is used. The combination of this segmentation with data from the hadronic calorimeter gives a jet rejection factor of ~ 2000. The remaining background is due mainly to the single high p_T π^o's from jet fragmentation. A fine-grained, position sensitive device is required to distinguish the two photons from the decay of the π^o and so reduce this background. This is achieved by segmenting the first sampling region into narrow strips of size $\Delta\eta \times \Delta\phi \approx 0.003 \times 0.1$. This first sampling region effectively works as an integrated pre-shower detector.

A pre-sampler, located directly behind the cryostat inner wall, precedes the electromagnetic calorimeter in the barrel region. It performs two functions: to correct for the energy lost by a particle in the material in front of the electromagnetic calorimeter, and to help in measuring the direction of electromagnetic showers.

2.2.3.2 The Hadronic Calorimeter

The barrel hadronic calorimeter is based on a sampling technique using steel absorber material and scintillating tiles. The tiles are read out by fibres placed at both open edges. The tile calorimeter surrounds the cryostats that contain the LAr calorimetry and is contained in an outer support cylinder which also acts as the main flux return yoke for the 2T inner solenoid. The scintillating tiles are oriented in planes perpendicular to the beam axis and staggered in R to provide a good sampling homogeneity. Readout cells are defined by the grouping of the readout fibres going to photomultiplier tubes.

The hadronic calorimetry is extended to higher rapidity, $1.5 < |\eta| < 3.2$, by the endcap hadronic calorimeters. LAr technology is again used to handle the higher radiation levels in this region. Copper absorber plates are placed in a liquid argon ionisation medium which is read out by an electrode structure which forms an electrostatic transformer. The hadronic endcaps sit behind the electromagnetic endcaps in the endcap cryostats.

The major tasks of the hadronic calorimetry are to identify jets and to measure their direction and energy, to measure the total missing E_T, and to improve the particle identification capability of the electromagnetic calorimetry by measuring such quantities as leakage and isolation. Intrinsic effects limit the reconstruction of both jet energy and missing E_T. Jet energy measurement is limited by effects such as jet fragmentation, magnetic sweeping of charged particles, gluon radiation and non-interacting particles (neutrinos and muons). The pile-up noise from minimum bias events contributes a significant uncertainty to the measurement at the LHC design luminosity. A lower energy resolution than in the electromagnetic calorimeter is therefore sufficient. An energy resolution of $\Delta E/E = 50\%/\sqrt{E} \oplus 3\%$ in the barrel region and $\Delta E/E = 100\%/\sqrt{E} \oplus 10\%$ in the endcap region provides an adequate sensitivity to study all the physics processes of interest.

The segmentation of the hadronic calorimetry is governed by such processes as the decay $H \rightarrow WW \rightarrow l\nu + 2jets$ which requires the reconstruction of a high-p_T W to jet pairs. A well segmented calorimeter is required to give the necessary efficiency and background rejection. A segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ is sufficient for reconstructing jet pairs and a segmentation of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ in the endcap calorimeter is sufficient to reconstruct jets at high η needed for identifying WW/ZZ fusion events.

The required η coverage of the hadronic calorimetry is determined by the high acceptance required for the measurement of missing E_T. The loss due to prompt neutrinos and muons limits the measurement in the endcaps at $|\eta| \geqslant 3.2$. The coverage is extended up to the desired level, $|\eta| = 4.5$ by the integrated forward calorimeter.

2.2.3.3 The Integrated Forward Calorimeter

Excellent hermicity is provided by the forward calorimeter which completes the coverage for jets. The extreme particle and energy flux at high pseudo-rapidity values necessitates a high speed of response and radiation hardness. The calorimeter is based on LAr technology utilising a rod and tube structure. Problems caused by positive ion buildup causing distortion of the electric field and degradation of the signals are possible at high particle and energy fluxes. However, the very narrow LAr gap, $250\mu m$ compared with the more normal gap of $1.94mm$ used in the LAr barrel calorimeter, avoids this problem.

The forward calorimeter is made up of three modules, each having 12120 tubes. The first module performs the electromagnetic calorimetry. It is constructed of copper electrodes embedded in a copper matrix which acts as an absorber. The remaining outer two modules are the hadronic modules which are constructed of stainless steel electrodes embedded in a tungsten-alloy absorber matrix. The electromagnetic module and first hadronic module contain the bulk of the hadronic shower, the second hadronic module acts as a tail-catcher. The interaction length of tungsten is approximately half that of copper or iron and thus limits the transverse spreading of hadronic showers. This enables the forward calorimeter to have angular resolution, $\Delta\eta \times \Delta\phi \approx 0.15 \times 0.15$, comparable to that of a forward calorimeter located much further from the interaction point.

2.2.4 The Muon Spectrometer

The muon spectrometer [21] consists of a large volume magnetic field instrumented with precision tracking chambers and a dedicated trigger system. The performance requirements follow from the physics goals and background environment in LHC. Good momentum resolution, over the range $|\eta| \leqslant 3$, is required for the detection of decays such as $H \rightarrow ZZ^* \rightarrow 4l$ and $Z' \rightarrow \mu\mu$ above large backgrounds, the main components of which are decay muons, electrons and hadrons from punch-through, and neutrons and photons. The stand-alone performance provided by the dedicated trigger system safeguards against unanticipated backgrounds and event topologies.

The magnet system used by the muon spectrometer consists of three super-conducting air-core toroids: a long barrel toroid with two inserted endcaps. Each toroid is constructed of eight flat coils assembled radially and symmetrically around the beam axis. Each endcap toroid is housed in a single cryostat, while the coils forming the barrel toroid have individual cryostats. The light, open structure of the air-core toroid system minimises the effects of multiple scattering on the muon momentum measurement. A magnetic field of $0.5 - 2T$ covering $|\eta| < 3$ is produced by the magnet system.

The layout of the muon chambers in the spectrometer results from consideration of requirements such as the efficient use of the bending power of the magnets and the coverage of the spectrometer balanced against practical issues such as ease of access after installation and the size of chambers that can be easily produced and transported. In the barrel, the muon chambers are arranged in three concentric cylinders around the beam axis. Each endcap region is made up of four disks of chambers, concentric around the beam pipe. The positioning of the chambers is such that a particle which originates at the interaction point will traverse three chamber stations.

Figure 2.5 shows the position of the precision and trigger chambers. There are two types of precision chambers used on the spectrometer: monitored drift tube chambers (MDTs) and cathode strip chambers (CSCs), and two types of trigger chambers: resistive plate chambers (RPCs) and thin gap chambers (TGCs). The precision chambers provide precise tracking and therefore momentum measurements, but have readout times that are in general much greater than the LHC bunch crossing time. The trigger chambers provide less precise position and momentum measurements for the purpose of triggering, and also allow the measurements made by the precision chambers to be assigned to a specific bunch crossing. Multiple scattering in the chamber materials and magnet structures limit the spectrometer momentum resolution to $2-3\%$.

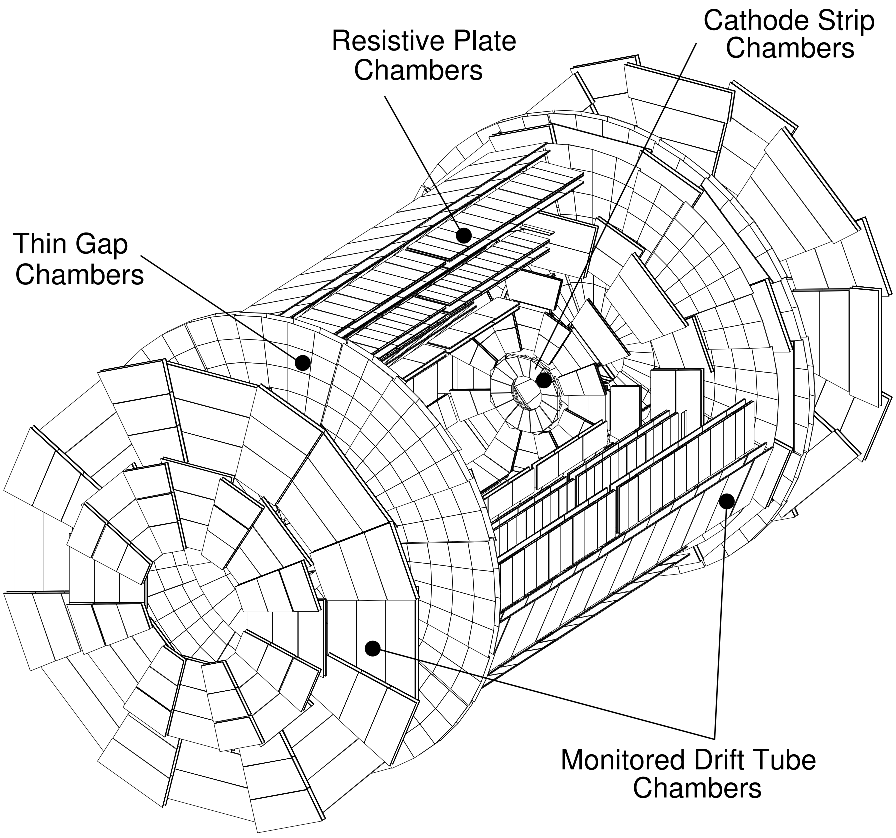


Figure 2.5: 3-D view of the muon chambers, indicating the regions in which the different technologies are used [21].

The MDTs and CSCs both consist of 2 sets of multi-layers, each made up of three or four planes of individual detectors. The multi-layers are separated by $150 - 350mm$, depending on the chamber's position in the detector, to allow for a local vector measurement. In the case of the MDTs the layers are made up from closely packed drift tubes, which have a drift time of $\sim 600ns$. The CSCs are constructed from layers of multiwire proportional chambers, giving readout times $< 25ns$. The higher granularity and radiation hardness of the CSCs allows them to be used in the high rapidity regions of the spectrometer where the MDTs cannot operate. However, cost prohibits their use throughout the whole detector. The CSCs are therefore employed only in a very small forward area ($2 < |\eta| < 2.7$). MDTs are employed throughout the remaining 99.5% of the area of the spectrometer.

The trigger system utilises RPCs in the barrel region, positioned at the middle and outer stations, and TGCs in the transition and endcap regions, positioned at the inner and middle stations, the choice being led by the same arguments as for the precision chambers. The RPCs are gaseous detectors with parallel electrode plates which are segmented into strips to provide positional information. The TGCs are multiwire chambers operated in saturated mode. They can operate at much higher rates than the RPCs. The trigger system also provides the only measurement of the track coordinate in the non-bending plane and is used in the pattern recognition algorithm.

2.2.5 The Trigger

A three level trigger system is employed by ATLAS to handle the very large event rate present at LHC. Each successive level refines the trigger decision and reduces the data rate to one that the next trigger level can process. The final trigger level performs full event building prior to making the decision of which events are written to permanent storage. A block diagram of the ATLAS trigger is shown in figure 2.6.

2.2.5.1 The Level-1 Trigger

The level-1 trigger accepts data at the full LHC bunch crossing rate of $40MHz$. The maximum allowable output rate is limited by the detector readout systems and the level-2 maximum trigger rate to $100kHz$. This requires that only one bunch crossing in 400 is selected. During the level-1 trigger processing the data from all parts of the detector are stored in

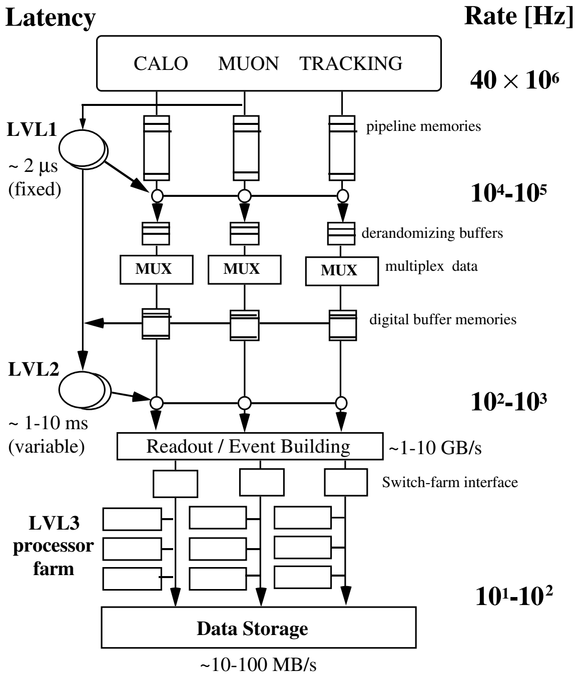


Figure 2.6: Block diagram of the ATLAS three level trigger system [16].

pipeline memories. This limits the acceptable latency (time to form and distribute the trigger decision) of the level-1 trigger to $\sim 2\mu s$, since the number of detector channels results in a rapid rise in cost as the pipeline length is increased.

Inner tracking information is not used in the level-1 trigger decision, due to the complexity of events at high luminosity and because the required reduction in the rate can be achieved without it. Reduced granularity data from the calorimetry systems and digital hit information from the muon trigger chambers are evaluated in separate processors before being combined in the level-1 central processor. The proposed signatures by which the level-1 trigger recognises an event of interest are made up from combinations of four basic building blocks: electron/photon triggers, jet triggers and missing-E_T triggers which are evaluated in the calorimetry processor, and muon triggers which are evaluated in the muon processor. The algorithms used to evaluate these signatures are hard wired, however some flexibility is provided by the ability to set the momentum and energy thresholds used. This is required to allow the level-1 trigger to react to unexpected physics and unforeseen situations. Additional lower thresholds are used to define Regions of Interest (RoIs) which are passed onto the level-2 trigger.

2.2.5.2 The Level-2 Trigger

If the level-1 trigger accepts an event, the data for that event are shifted from the pipeline memories in the front-end electronics into derandomising buffers. The data are then compressed, multiplexed and then read out of the detector to the barracks where they are stored in the level-2 buffer memories for the duration of the level-2 trigger latency. If either of the buffers becomes full an inhibit signal is sent to the level-1 central processor which stops the production of level-1 triggers. This results in the loss of events and is therefore highly undesirable.

The level-2 trigger aims to reduce the data rate by a factor of ~ 100. This is made possible by the use of full granularity data from the calorimetry, data from the precision muon chambers, and inner tracking data. Full calorimetry granularity sharpens the electron/photon, jet, and missing-E_T cuts. The electron trigger can be improved by using the inner tracking data to match high-p_T tracks with calorimeter hits. Finally the muon cuts can be refined by the combination of the precision muon chamber data and inner tracking data. In addition events may be required to satisfy extra signatures such as extra leptons at low-p_T or mass thresholds for secondary vertices.

The processing in the level-2 trigger can be split into three steps:

- feature extraction is used to reduce the data from each sub-detector from basic cell information to more physical parameters, for example cluster parameters in the calorimeter. Only data in the RoIs identified by the level-1 trigger are examined in order to reduce the quantity of data to be processed;

- features for one RoI from several sub-detectors are combined to form an object containing information on particle type, energy and direction;

- objects from all the RoIs are combined to form a global event decision.

2.2.5.3 The Level-3 Trigger

If the level-2 trigger accepts an event the full detector data is passed onto the level-3 trigger system. The level-3 trigger aims to reduce the event rate by a factor of ~ 10 to $10 - 100Hz$, corresponding to a data rate of $\sim 10 - 100 Mbyte/s$ for storage. Full event building is performed followed by event selection on the basis of physics signatures.

Full event data will be recorded for certain triggers, for example Higgs boson candidates, however only partial event storage will be made available for other triggers such as calibration triggers and high mass jet-pair triggers which may indicate new particles decaying to jets.

A complete solution has not yet been defined, since the rapid evolution of computing and communications software would make it obsolete by the time of its commissioning.

Chapter 3

Semiconductor Central Tracker

3.1 Introduction

The SCT [20] consists of a total of 4088 detector modules, 2112 make up the 4 layers of the barrel and 988 the nine wheels of each of the forward regions. A description of the modules is given in section 3.2. The detector and associated electronics design have followed several options and the ATLAS community's decisions on the baseline choice for the electronics has led the choice of readout. There are three readout technologies covered in this thesis and this requires an explanation of the two different detector and electronics options that they are designed to operate with. The two options: analogue and binary are discussed in section 3.3. However, it should be noted that the binary option was chosen as the baseline for the ATLAS SCT in March 1996. The physical and physics based requirements that the SCT demands of the readout are then covered in section 3.4. There then follows a brief discussion of the additional requirements of the pixel detector, in section 3.5, since a common readout choice for this detector is preferred.

3.2 Modules

Each module is made up from four silicon microstrip detectors, each with an active area $61.6 \times 62.0mm$. Pairs of detectors are then daisy-chained together to form two $12cm$-long detectors which are mounted back-to-back. In order to achieve the required noise performance the electronics must be situated as close to the detectors as possible. The front-end electronics sit on a low-mass hybrid which forms the connections to the strips.

3.3 Electronics

3.3.1 Binary Readout

In the binary scheme only the locations of hit strips are read out. A threshold cut is employed to decide if there has been a hit in a specific strip. The desired detector resolution is achieved by employing 768 strips spaced at $80\mu m$ across each $\sim 6cm$ detector. The resolution for a top-hat distribution is given by $width/\sqrt{12}$, resulting in a resolution of $23\mu m$ across the strips.

The processing blocks for the signals from the detector are shown in figure 3.1.The signals are first amplified before undergoing a threshold cut. The data are then stored in the level-1 pipeline buffer. Upon receipt of a level-1 trigger signal three bunch-crossings worth of data around the event are shifted into the readout buffer. The data then undergo a compression stage. Depending on the running mode one of four hit patterns is looked for in each strips data as shown in table 3.1. The strip addresses of those passing the selected hit pattern are multiplexed and sent to the binary data link for transmission.

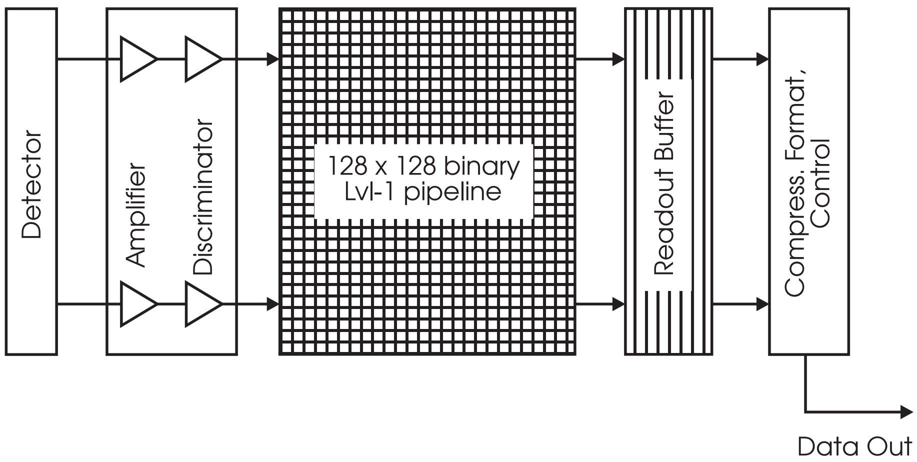


Figure 3.1: Block diagram of the binary readout processing steps.

Table 3.1: Hit patterns used for data compression in binary readout (X signifies either a 1 or 0 is acceptable).

Mode	**Hit Pattern**	**Use**
Any hit	1XX,X1X,XX1	Detector alignment
Level sensing	X1X	Data taking
Edge sensing	01X	Data taking
Test	XXX	Testing

3.3.2 Analogue Readout

Analogue readout aims to transmit all the possible data from the SCT to the counting room. The signals coming from the 512 strips on one side of a silicon detector module undergo the processing steps shown in figure 3.2. Each signal undergoes a deconvolution [22] process which consists of a slow pulse shaping stage with amplification, sampling at three points of the resulting pulse shape and then a weighted summing of these samples to reconstruct the pulse height and timing. This process enables the signal to be associated to a specific beam crossing, but results in a small reduction in the accuracy of the pulse height measurement. The data are then stored in an analogue pipeline and shifted to the readout buffer if a level 1 trigger is received. The data are then multiplexed and passed to the analogue data link for transmission. Only 512 strips are required on each module due to increased resolution offered by the pulse height information of each strip hit. This enables interpolation of the hit position between two strips. A strip spacing of $112.5\mu m$ therefore offers the same resolution as the $80\mu m$ binary read out strips.

The binary readout of data has the advantages of robustness and a cost saving over the analogue system due to the smaller size of the front-end chips. The transmission of binary data has the additional advantages of lower performance requirements for both the linearity and dynamic range of the link. However, the analogue readout has a safety factor coming from making tracking decisions in the counting room. Any failure in the binary readout system due to, for instance, a high noise occupancy would lead to the loss of data. In the

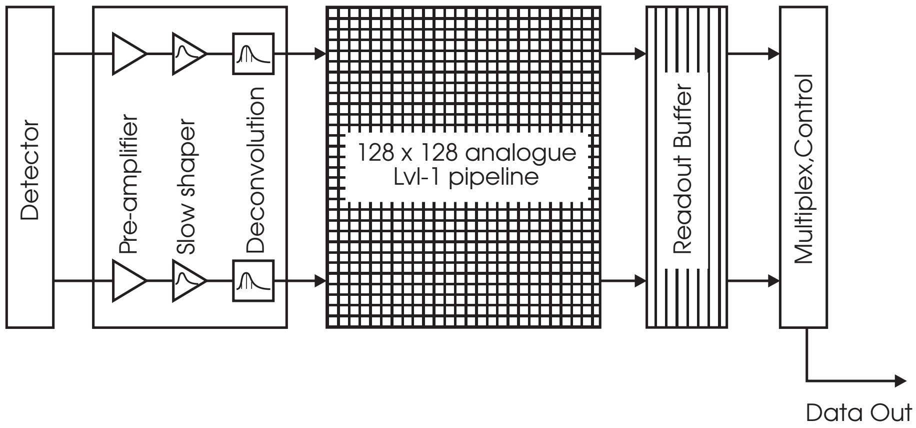


Figure 3.2: Block diagram of the analogue readout processing steps.

analogue readout the extra information from all the strips on a module may allow unforeseen problems to be solved.

3.4 Requirements

3.4.1 Data Transmission

In terms of the data that must be transmitted the system requirements are:

- the transmission of a $40MHz$ clock to the SCT module;
- the transmission of trigger and slow control signals to the SCT module;
- the transmission of $\sim 105M$ analogue signals $/s$[(a)] from the SCT module in the analogue readout scheme or $\sim 35Mbit/s$[(b)] of data from the SCT module in the binary readout scheme;
- a data loss $< 0.1\%$.

[(a)] 2×512 strips with a $100kHz$ maximum level-1 trigger rate

[(b)] 2×768 strips, maximum occupancy of 1%, and a $100kHz$ maximum level-1 trigger rate. 17 bit addresses are used for each hit, 4 extra bits are transmitted for additional hits forming a cluster (assume 50% of hits have neighbours), and 31 header and trailer bits are transmitted per link.

3.4.1.1 Analogue Transmission Quality

Analogue data can be distorted during transmission either by the addition of noise or by non-linear scaling of the signals. The signal to noise ratio, measured at the output of the front-end electronics, for a minimum ionising particle incident on a detector is expected to be in the range $\sim 30:1$. Detector output for signals up to 4 MIPs should be linear at the $5-10\%$ level, resulting in a dynamic range $\sim 120:1$. For a transparent data link the dynamic range must exceed that of the data to be transmitted while having small non-linearity. The data link requirements are therefore set at: dynamic range $> 250:1$[c], and non-linearity $\lesssim 2\%$. A larger non-linearity can be tolerated for signals greater than 4 MIPs.

3.4.1.2 Binary Transmission Quality

Errors introduced in the transmission of binary hit data result in the wrong address data being transmitted and must therefore be avoided. For a link to operate at a bit error rate (BER) $< 10^{-9}$[d] the required signal to noise ratio is $12.4:1$ assuming Gaussian distributed noise. The inclusion of a general engineering safety margin results in a signal to noise target of $20:1$.

A similar signal to noise requirement is imposed for the transmission of the digital clock, trigger and control signals to the module.

3.4.2 Magnetic Properties

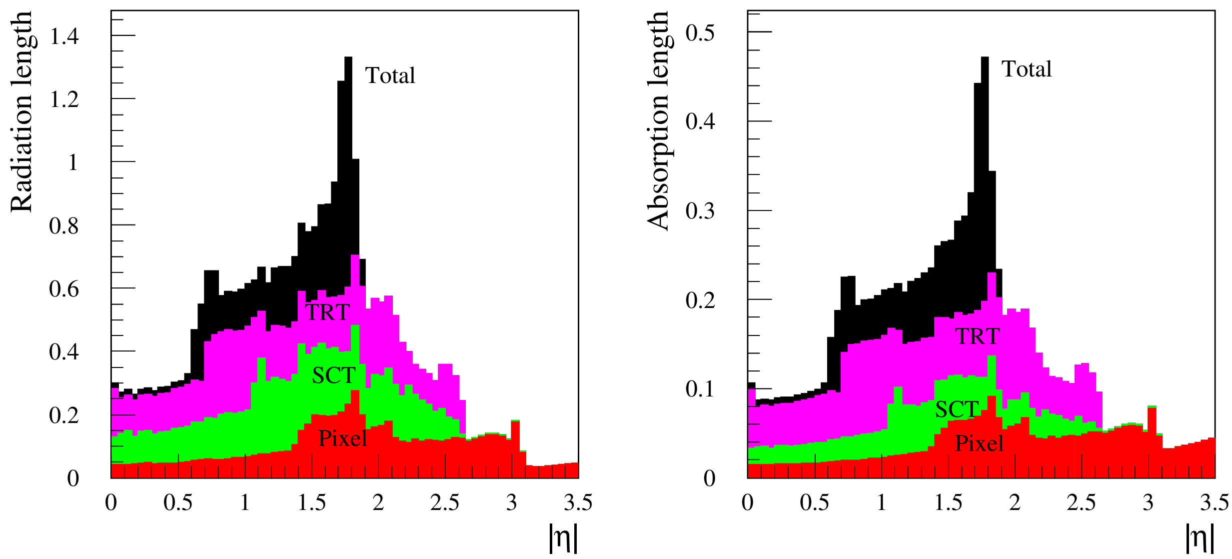
The front-end link components will be required to function in the $2T$ field of the Inner Detector. In addition to the transmission and reception properties being unaffected the packages used for the link components will be required to be non-magnetic.

3.4.3 Material

Minimising the amount of material in the Inner Detector is important for the physics performance of the whole ATLAS detector. This is because the function of the Inner Detector is to

[c]The noise contribution of the link is added in quadrature to the noise of the data being transmitted. A link with double the dynamic range of the data will therefore have $< 12\%$ contribution to the total noise.
[d]BER $< 10^{-9}$ is an industry standard for data transmission. In fact a BER $\sim 10^{-5}$ would fulfil the ATLAS requirement for data loss $< 0.1\%$

track the particles passing through it, but otherwise to interact with them minimally. The material can be viewed in terms of two important parameters: the radiation length, which characterises the effect the material will have on electromagnetically interacting particles; and the absorption length, which characterises the effect on strongly interacting particles. The material in the Inner Detector in terms of each of these two quantities is shown in figure 3.3. The plots show the cumulative amount of material after the pixel detectors, SCT, TRT, and the whole Inner Detector. The final contribution is from the services and patch panels which are outside the active tracker volume just in front of the calorimetry cryostat. In general material closer to the interaction point has a greater effect on the calorimetry.



(a) Radiation lengths. (b) Absorption lengths.

Figure 3.3: The cumulative material in the Inner Detector in terms of radiation lengths and absorption lengths [20].

A higher number of radiation lengths of material leads to tracks undergoing more multiple scattering, an increase in the multiplicity from secondary particles (eg from the conversion of γs coming from the decay of π^{o}s), a greater bremsstrahlung probability for electrons, and a greater conversion probability for photons.

A large number of absorption lengths of material leads to the absorption of hadrons and therefore lost tracks, large scatters of hadrons, and again to an increase in multiplicity from secondary particles.

The data link should contribute a minimal amount of material to the total of the Inner Detector. A reasonable limit is $\lesssim 2\%$ of a radiation length.

3.4.4 Power Budget

The power budget of the SCT is constrained by the increase of material required for additional cooling. The maximum power dissipation per channel of the SCT is set at $< 3.8mW$.

3.4.5 Reliability

The Inner Detector will not be accessible after the ATLAS detector has been commissioned. Reliability of the components is therefore a very important issue and the data links are required to work for the full lifetime of ATLAS. However, in the event of a component failure it is important that the overall operation of the detector is affected minimally. It is required that only the small amount of data processed by the failed component is lost or if possible the data are re-routed around the component so that no data is lost.

3.4.6 Radiation Environment

The radiation levels at the LHC experiments will greatly exceed those in present particle physics experiments.

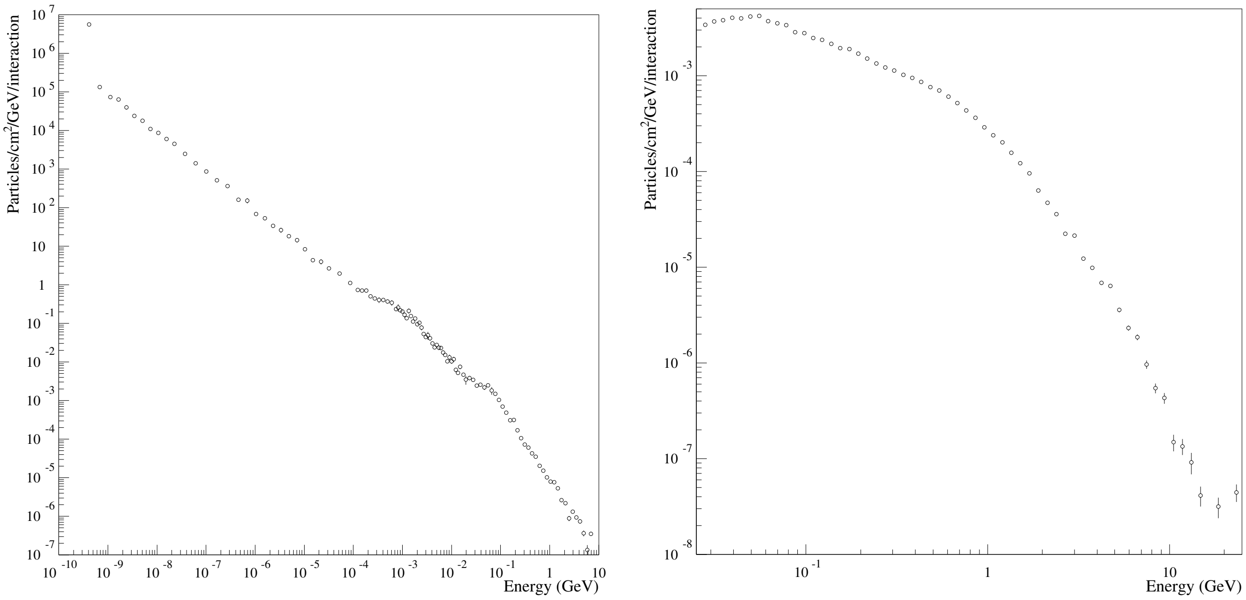
The programme for the LHC is to begin running with 3 years of low luminosity operation ($10^{33}cm^{-2}s^{-1}$) followed by 7 years of high luminosity operation ($10^{34}cm^{-2}s^{-1}$). One years running is assumed to be 10^7s, resulting in an integrated luminosity of $7.3 \times 10^{41}cm^{-2}$ over the LHC lifetime. Taking the proton-proton inelastic cross-section to be $81mb$[e] the expected number of interactions, N_{int}, is calculated to be 6.1×10^{16}[f].

The fluence [20] per interaction is evaluated by generating minimum bias particles using the DTUJET code and then modelling the subsequent interactions and decays of the particles

[e] $1Barn(b) = 1 \times 10^{-28}m^2$

[f] A small factor (70/67) is required to correct for differing proton-proton inelastic non-single diffractive cross-sections used in the DTUJET code ($67mb$) and assumed by ATLAS ($70mb$).

using the FLUKA transport code, taking into account the material distribution in the Inner Detector and Calorimeter. The resulting energy spectra for neutrons and charged particles are shown in figure 3.4.



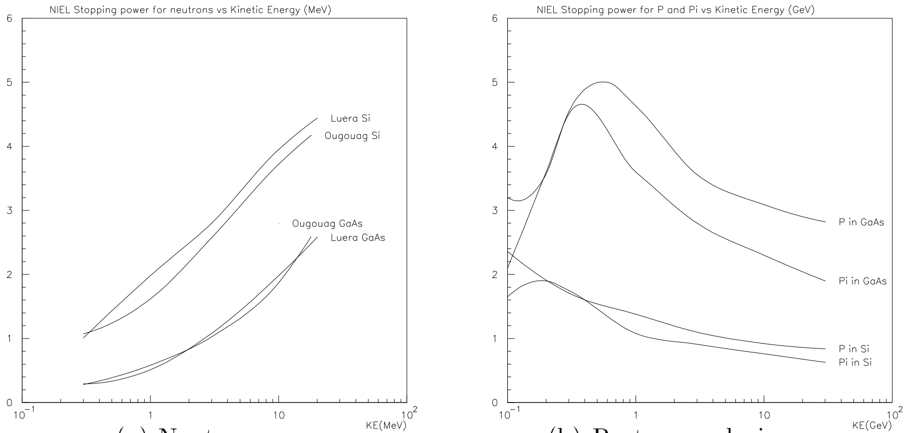
(a) Neutrons.

(b) Charged π

Figure 3.4: Energy spectrum at the 1^{st} SCT barrel for neutrons and charged π [20].

The damage in both Silicon (Si) and Gallium Arsenide (GaAs) devices is theorised to result predominantly from the non-ionising inelastic energy lost (NIEL) by the incident particles to the semiconductor material. This has been verified experimentally for different incident particles, over a wide incident energy, for Si and to a lesser extent GaAs [23–28]. Results from the irradiation of GaAlAs LEDs presented in chapter 7 also agree with the hypothesis. A simplification of the Inner Detector particle fluences to standard particle types can therefore be performed. The irradiation sources used to test devices have been either neutron sources with energy spectra centred close to $1 MeV$ or the $24 GeV$ proton synchrotron (PS) beam at CERN. Two standard particle types are therefore chosen: $1 MeV$ equivalent neutrons and $24 GeV$ equivalent protons. The NIEL stopping powers for various particle types are shown in figure 3.5.

The equivalent fluence of neutrons in terms of $1 MeV$ neutrons over the detector lifetime



(a) Neutrons. (b) Protons and pions.

Figure 3.5: NIEL stopping powers for neutrons and charged particles in Silicon and Gallium Arsenide [29].

is given by,

$$F_{1MeV\ n} = \sum_E \frac{dN}{dE} \Delta E \Big(\frac{NIEL(E)}{NIEL_n(1MeV)} \Big) N_{int}, \tag{3.1}$$

where, $\frac{dN}{dE}$ is the number of $neutrons/cm^2/GeV/interaction$, from figure 3.4,

ΔE is an appropriate energy interval,

$NIEL(E)$ is the NIEL for neutrons at energy E, from figure 3.5,

and $NIEL_n(1MeV)$ is the NIEL of $1MeV$ neutrons.

Similarly the equivalent fluence of charged particles in terms of $24GeV$ protons is given by,

$$F_{24GeV\ p} = \sum_E \frac{dP}{dE} \Delta E \Big(\frac{NIEL(E)}{NIEL_p(24GeV)} \Big) N_{int}, \tag{3.2}$$

where, $\frac{dP}{dE}$ is the number of $pions/cm^2/GeV/interaction$, from figure 3.4,

$NIEL(E)$ is the NIEL for pions at energy E, from figure 3.5,

and $NIEL_p(24GeV)$ is the NIEL of $24GeV$ protons.

The NIEL stopping powers for the standard particle types and those calculated for the energy spectra of the radiation sources used in the testing (see Appendix A) of the components are shown in table 3.2. The uncertainty on the values given for neutron data is $\sim 10\%$, while for proton and pion data an uncertainty $\sim 20\%$ is not unlikely.

Table 3.2: NIEL stopping powers for various particle types ($keV(gm\ cm^{-2})^{-1}$).

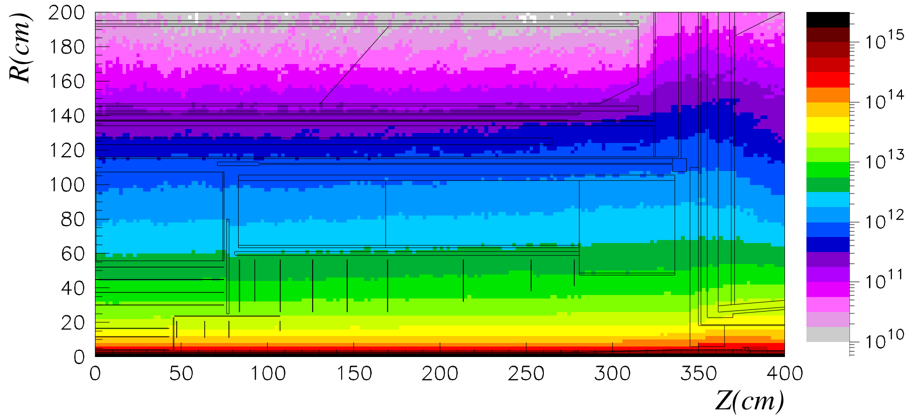
Irradiation	GaAs	Si
$1MeV$ neutrons	0.54	1.80
$24GeV$ protons	2.9	1.0
$300MeV$ pions	3.6	1.9
ISIS neutrons	0.81	2.17
Dynamitron neutrons	0.68	1.97

The 10 year integrated fluence at the 1^{st} SCT barrel due to charged particles, neutrons, and the total for all particles is shown in table 3.3. For each case the fluence is expressed in terms of $1MeV$ equivalent neutrons and $24GeV$ equivalent protons in Silicon, from which the photodiodes tested in chapter 7 are fabricated, and GaAs, which is used in the fabrication of both the multi quantum well modulators and the light emitting diodes (see chapters 6 and 7 respectively). Irradiation of components by both protons and neutrons is required for confidence. Irradiation by charged particles carries an increased likelihood of ionisation damage resulting in, for instance, surface charge effects.

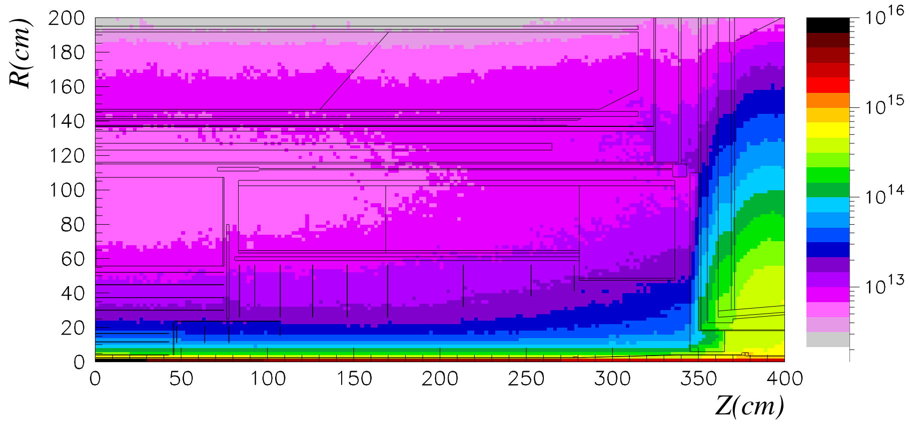
Table 3.3: Equivalent fluences at the 1^{st} SCT barrel.

Particle Type	F_n in Si ($10^{13}cm^{-2}$)	F_p in Si ($10^{13}cm^{-2}$)	F_n in GaAs ($10^{13}cm^{-2}$)	F_p in GaAs ($10^{13}cm^{-2}$)
Charged particles	6.3	11.3	39.7	7.4
Neutrons	6.7	12.1	12.9	2.4
Total	13.0	23.4	52.6	9.8

The fluence distribution throughout the Inner Detector cavity is shown in figure 3.6. The charged hadrons are mainly produced in direct pp interactions and are distributed approximately uniformly in rapidity and with low p_T. Beam losses and beam-gas interactions contribute an order of magnitude less radiation at high luminosity. Interactions in the calorimeter produce albedo neutrons which are a significant source of radiation and dominate at larger radii of the SCT.



(a) Fluence of charged hadrons.



(b) Total fluence in units of $1MeV$ equivalent neutrons in Si.

Figure 3.6: Annual radiation fluences expected in the Inner Detector cavity at high luminosity running ($cm^{-2}year^{-1}$) [20].

3.4.7 Thermal Properties

The SCT must be able to operate over the temperature range $-15^oC \rightarrow +30^oC$. This allows the commissioning and testing of the SCT at room temperature as well as the normal low temperature operation required by the Silicon microstrip detectors. Operating the detectors at low temperature reduces the increase in depletion voltage due to radiation damage and also reduces the leakage current and hence heat generation inside the detector substrate. It must function over $\pm 2.5^oC$ around the nominal running value with no need for recalibration.

3.4.8 Cost

A cost target for the data link of $\sim$ 360 CHF per SCT module is imposed for budgetary reasons. This price should include all the link components from the front-end drivers to the back-end receivers and amplifiers.

3.5 Pixels

The ATLAS collaboration would prefer a common solution for the readout of the Pixel Detector and the SCT. Many of the requirements are similar: the need to minimise the material and power budget, reliability, and the thermal operating conditions are the same as for the SCT. The data from the pixel modules can be grouped to optimise the bandwidth use of the links. This grouping will change with the different layers due to the increasing hit occupancy with decreasing radius.

The location of the pixels at smaller radii results in higher radiation fluences throughout the detector lifetime. The expected fluences for the barrel pixel layer located at $R = 11cm$ is $\sim$ 5 times higher than that of the 1^{st} SCT barrel. The b-layer barrel would experience a similar fluence before it and its associated electronics would require replacing.

Chapter 4

Data Link Architectures

4.1 Introduction

Three designs for the data link to read out the SCT modules have been investigated:

- optical analogue readout using a reflective optical modulator;
- optical binary readout using light emitting diodes and photodiodes;
- electrical binary readout over miniature twisted pairs.

Descriptions of the architecture and components of each of the systems are given in sections 4.2, 4.3, and 4.4. Evaluation of each of the links in turn is covered in chapters 6 to 8. The arguments for and the choice of the baseline link is discussed in chapter 9. It should be noted that the binary electronics option was chosen by the ATLAS community in early 1996 after consideration of the complete detector and readout package. The analogue optical link has therefore not been pursued to its conclusion. There then follows a description of future developments that are being investigated.

4.2 Analogue Optical Readout

The RD-23 [30] Project is a CERN research and development programme to investigate 'Opto-electronic Analogue Signal Transfer for LHC Detectors'. It is a collaboration including both academic and industrial institutions. The work on this option has been carried out as part of this project.

The readout chain is shown in figure 4.1. A hybrid front-end board connects to the detector. Mounted on this hybrid are the analogue front-end chips, readout controller (ROC) and optical readout package[(a)]:

- Each analogue front-end chip processes the signals from 128 strips of a detector as described in section 3.3.2. The data from two front-end chips are sent to one channel of the ROC;
- The ROC acts as an interface to the electro-optic converter, amplifying the signal from the analogue pipeline and providing an appropriate bias voltage for the electro-optic converter. It also distributes the trigger and control signals to the front-end chips;
- An eight channel bi-directional optical link transmits four channels of analogue signals from the ROC down optical fibre to the back-end readout module, and receives two channels of digital clock, trigger and control data. The analogue signals are transmitted at $40MHz$ and the digital signals at $40M\ bit/s$. It acts as the data interface for all information passed between a particular module and the control room.

The electro-optic converter used is of a passive multi-quantum well (MQW) type. It can be used either to modulate continuous wave light incident upon it or as a photodiode receiver. In the modulator mode the reflection coefficient of the device is altered by the voltage applied.

The back-end readout module contains two optical transmitters and four transceiver channels:

- The optical transmitters are lasers which are directly amplitude modulated to encode the data. Separate channels are used to send the clock and the trigger and control data;
- Each transceiver channel consists of a laser, to provide continuous-wave light that is modulated at the front-end hybrid by the MQW, and a photodiode to receive the

[(a)]The optical readout package may optionally be mounted on an individual hybrid.

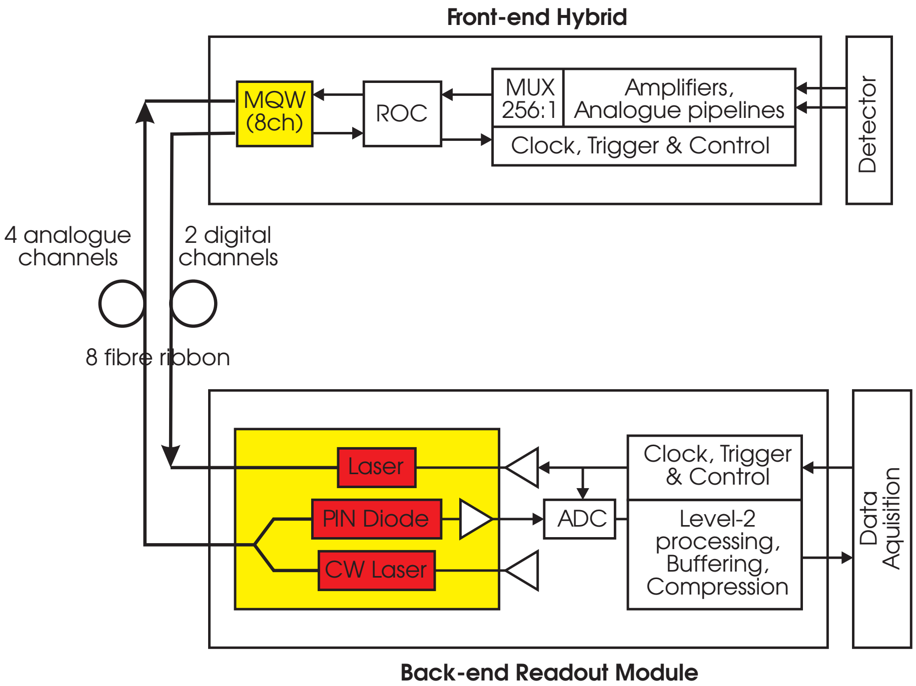


Figure 4.1: Schematic of the proposed analogue optical readout system for the ATLAS SCT.

reflected signal. An optical coupler connects both to the same fibre which carries the continuous-wave light to the MQW as well as the reflected modulated light back;

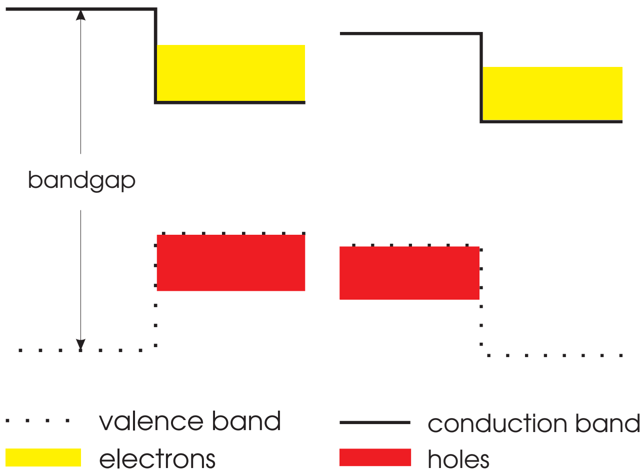
- The back-end module also contains electronics that digitise the analogue signals ready for processing by the trigger and storage by the data acquisition system.

Detailed descriptions of the link components follow.

4.2.1 Multi-Quantum Well Modulators

4.2.1.1 Quantum Wells

The choice of suitable semiconductor compounds allows the energy difference between the valence and conduction bands (the band-gap) to be altered. By fabricating semiconductors in very thin layers, with different band-gaps, the carriers can be confined in one dimension. There are two choices for the confinement of the carriers as shown in figure 4.2. In a Type I structure the electrons and holes are confined in the same layer of the structure, but, in a Type II structure the electrons and holes are held apart in different layers. We will consider the case of the Type I structures.



(a) Type I. (b) Type II.

Figure 4.2: Schematic diagram of Type I and Type II well structures.

The solution of the Schrödinger equation gives the wave-functions and energies for a particle of mass m trapped in a 1D potential well and can be found in all good quantum mechanics textbooks. For a potential well of depth V and extent $\pm\frac{1}{2}$L about z=0, as shown in figure 4.3, the even parity wave-functions are given by,

$$\Psi_n(z) = \begin{cases} Acos(k_1 z) & -\frac{1}{2}L < z < +\frac{1}{2}L \\ Bexp(-k_2(z-\frac{L}{2})) & z > +\frac{1}{2}L \\ Bexp(+k_2(z+\frac{L}{2})) & z < -\frac{1}{2}L \end{cases}, \tag{4.1a}$$

and the odd parity wave-functions are given by,

$$\Psi_n(z) = \begin{cases} Asin(k_1 z) & -\frac{1}{2}L < z < +\frac{1}{2}L \\ Bexp(-k_2(z-\frac{L}{2})) & z > +\frac{1}{2}L \\ Bexp(+k_2(z+\frac{L}{2})) & z < -\frac{1}{2}L \end{cases}, \tag{4.1b}$$

where,

$$k_1 = \frac{\sqrt{2mE_n}}{\hbar} \qquad k_2 = \frac{\sqrt{2m(V-E_n)}}{\hbar}, \tag{4.1c}$$

E_n is the energy for state n,

and A, B are constants such that the wave-function is normalised and continuous at the well boundaries.

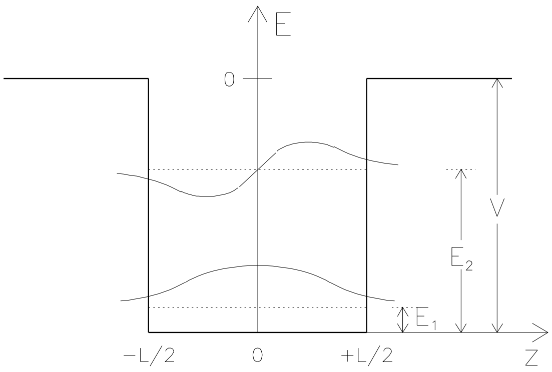


Figure 4.3: Eigen-energies and wave-functions for a finite quantum well.

For the case of holes the problem is more complex [31]. In the bulk semiconductor, propagation in a given direction can be described in terms of heavy and light-hole propagation. The quantum well potential acts to lift the degeneracy of the two mass states. The energy levels for a quantum well structure are shown in figure 4.4.The width of the barrier must be such as to prevent significant coupling between adjacent wells.

The lifting of the hole degeneracy leads to a more complicated form for the density of states. The infinite square well case is therefore discussed below.

The change from free carriers to carriers confined in 1 dimension has some important effects. Firstly the minimum photon energy to promote an electron from the valence to conduction bands is increased due to the quantisation of the energy levels. The density of states for the material also changes as shown in figure 4.5. In a bulk semiconductor there is a finite energy required to promote an electron from the valence band to the conduction band. The density of states is then proportional to $E^{-\frac{1}{2}}$, which will be filled from the lowest energy upwards. However, when the electron is confined in one dimension the density of states rises in steps as each of the allowed transitional energies is reached.

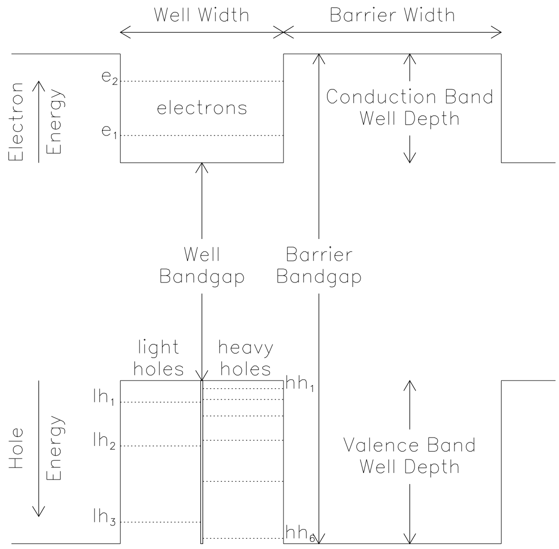


Figure 4.4: Band-gap diagram of a quantum well

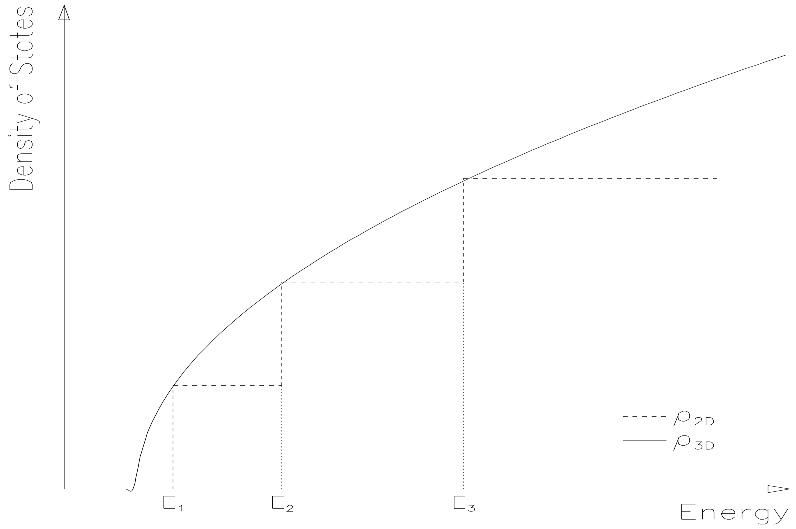


Figure 4.5: Density of states in bulk semiconductor and in a potential well.

4.2.1.2 Excitons

A bound state of an electron and hole known as an exciton, analogous to a hydrogen atom, is possible in bulk semiconductor materials. In fact the solutions to the Schrödinger equation for the hydrogen atom can be applied to the exciton to calculate its size and binding energy. The hole, effective mass m_h, takes the role of the proton, with the electron, effective mass m_e, in orbit. The binding energy for the exciton is given by the Rydberg constant [31],

$$R = \frac{2\mu e^4}{\hbar^2 (8\pi\varepsilon)^2}, \tag{4.2a}$$

where, μ is the reduced effective mass of the exciton,
e is the electronic charge,
and ε is the permittivity of the semiconductor.

The energy levels for the exciton are then,

$$E_n^{bulk} = \frac{R}{n^2}, \tag{4.2b}$$

where, n is the quantum number for the state.

The radius of the exciton is given by the effective Bohr radius [31],

$$r = \frac{4\pi\varepsilon\hbar^2}{\mu e^2}. \tag{4.2c}$$

The reduced effective mass given by,

$$\mu = \frac{m_e m_h}{m_e + m_h}, \tag{4.2d}$$

must be used since two particles of finite mass are orbiting about a common centre of mass.

In a potential well these values are altered slightly. The energy levels for the 2d case are given by the equation [31],

$$E_n^{2d} = \frac{R}{(n - \frac{1}{2})^2}. \tag{4.2e}$$

The quantum well also acts to confine the exciton in the z direction, this increases the binding energy and therefore lengthens the lifetime of the exciton which is short at room temperature in the bulk semiconductor due to ionisation by phonons. The overall effect is to increase the exciton peak height at the band-edge, and also to move it to higher frequency (energy).

The photon energy required to excite an electron into an exciton bound state is given by,

$$E = E_g + E_e + E_h - E_n, \tag{4.3}$$

where, E_g is the band-gap of the semiconductor,
E_e is the energy of the electron state,
E_h is the energy of the hole state,
and E_n is the energy of the n^{th} exciton state.

4.2.1.3 Quantum Confined Stark Effect

Applying an electric field across the quantum wells tilts the energy bands, this has two effects. Firstly the the energy gap between the bands is reduced, as shown in figure 4.6, resulting in a red shift in the absorption spectrum, this is known as the quantum confined stark effect [32], and secondly the overlap of the of the wave-functions is reduced resulting in a lower probability for absorption.

A measured absorption spectrum for bulk and quantum well material is shown in figure 4.7. The exciton peak in the bulk material appears as a small bump at the band edge. The MQW sample shows the exciton peak shift and enhanced absorption caused by confinement. The exciton peaks and plateaus for the n=1, 2 and 3 transitions are clearly seen. The light and heavy hole exciton peaks for the n=1 transition can also be resolved. Upon application of an electric field, the spectrum moves towards lower energies.

This effect can be used to build a modulator in which the absorption is controlled by the voltage applied and therefore the field across the device.

4.2.1.4 Use as a photodiode

The structure of the MQW modulator basically forms a photodiode. Although the design is not optimised in terms of the responsivity it is still possible to use the MQW to detect light.

For the application in the SCT the MQW would be required to receive digitally transmitted clock, trigger and control data. The MQW chips have been shown to perform as receivers for digital transmissions [30]. For digital operation the yield of channels is $\approx 100\%$ due to the lower signal to noise ratio that can be tolerated for acceptable operation.

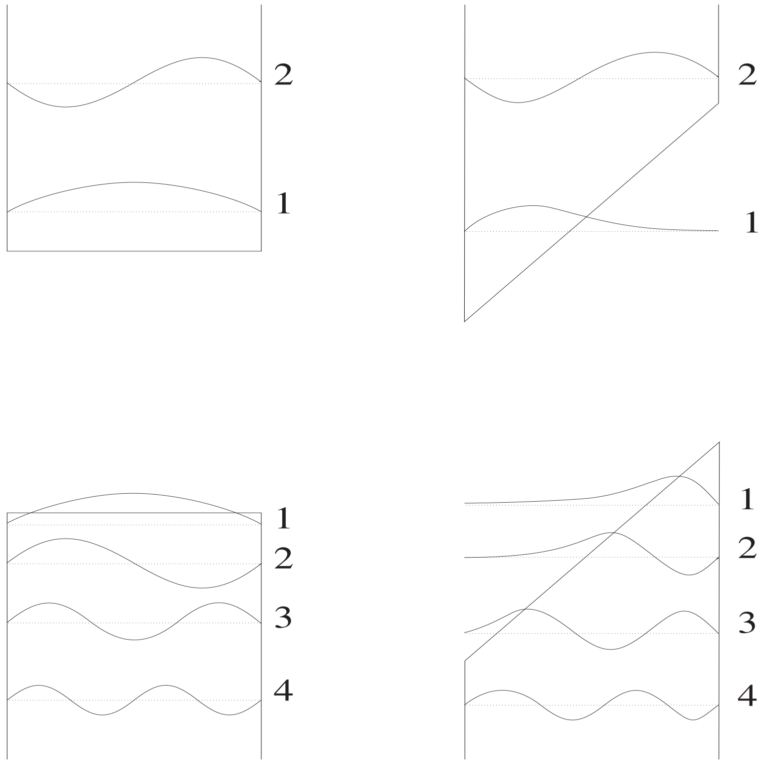


Figure 4.6: Schematic of the wave-functions in a quantum well with (right) and without (left) an applied electric field.

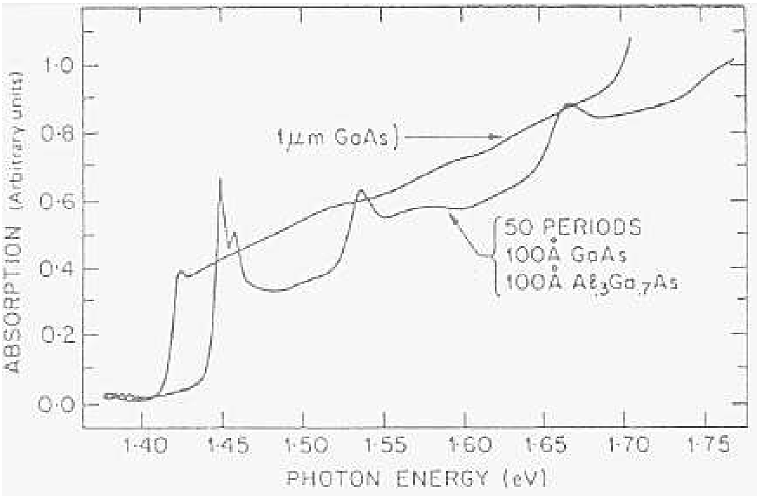


Figure 4.7: Comparison of the absorption spectra of bulk and 100Å MQW GaAs samples measured at room temperature [33].

4.2.2 Design and Fabrication

4.2.2.1 Growth Technique

The modulator wafers are grown epitaxially using a metal organic chemical vapour deposition (MOCVD) technique by GEC Marconi Materials Technology (GMMT). In this technique the growth of the structure is controlled by varying the the gas mixture which is passed over the wafer. The gas decomposes as it passes over the heated substrate, which is typically $600°C$, depositing material at the rate of ∼one mono-layer per second. This technique allows very pure crystals with no defects and highly controlled layer widths to be grown, all of which are essential for the performance of the quantum well devices.

4.2.2.2 Structure

Four rather than eight channel modulators have been fabricated for the test programme to make use of existing masks at GMMT. The modulator structure is shown in figure 4.8. The active region is made up of 50 periods of quantum wells, each well having a width of ∼ 100Å[(b)] and a barrier width of ∼ 100Å. The well material is $In_{.53}Ga_{.47}As$, which has a band-gap of $750meV$, and the barrier material InP, with a band-gap of $1350meV$.

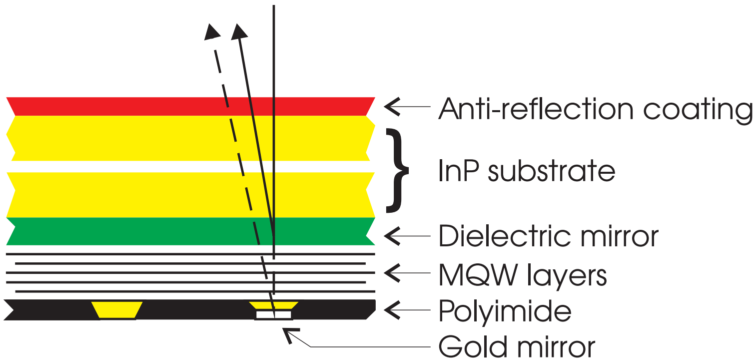


Figure 4.8: Schematic of slice through a MQW modulator wafer.

The active region is confined within an asymmetric Fabry-Perot cavity which is formed by two mirrors. The top dielectric mirror is grown from alternating bands of high and low refractive index layers to give a reflection coefficient of ≈ 30% [34]. The bottom mirror is a

[(b)] 10 Angstrom (Å)=$1 \times 10^{-9}m$

gold layer that is applied to the base of the completed wafer. It has a reflection coefficient of $\approx 95\%$. This results in light passing through the active region more than once, therefore giving a greater total absorption. It also gives rise to an additional wavelength dependence of the device since constructive interference of the reflections is required. The cavity is tuned to have an operational wavelength near that at which the first exciton peak lies.

Light enters the device through the substrate which has an anti-reflection coating applied. The voltage used to control the well absorption is applied via metallised p^+ contacts.

Figure 4.9 shows the theoretical reflection versus wavelength response for a MQW modulator. A laser emitting light at $1545nm$ is used to operate the device to match the tuning of the Fabry-Perot cavity.

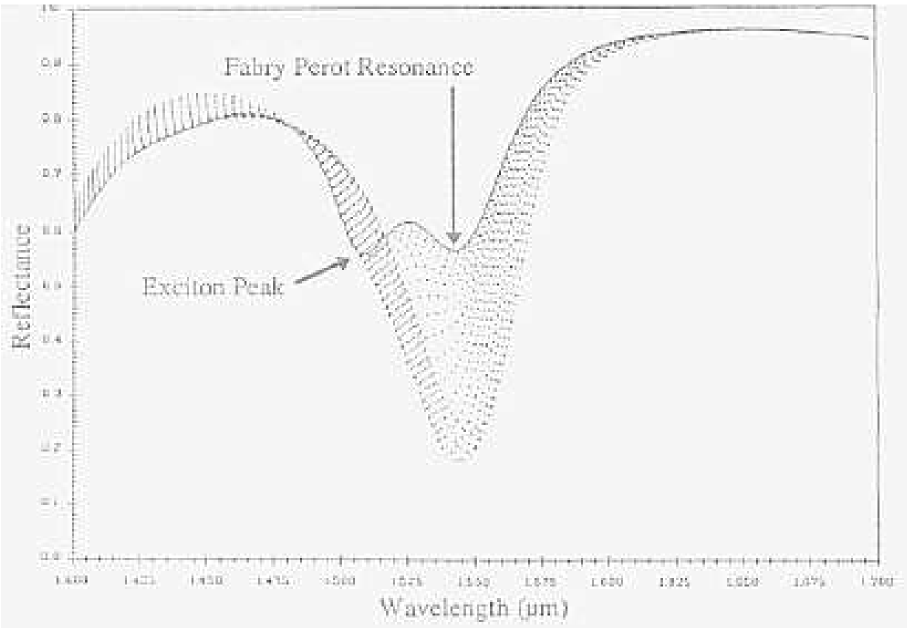


Figure 4.9: Calculated reflectivity spectrum of a MQW modulator [35]. The solid line is the reflectivity for $0V$ bias, while the dotted lines show the reflectivity at $\frac{1}{2}V$ increments up to $12V$. The modulator has a front mirror with reflectivity 0.2 and 37 quantum wells 65Å thick.

The wafer is 'diced' into chips each containing four modulator channels. The chips are then packaged using a passive technique as shown in figure 4.10. This involves solder-bump bonding the chip to a precision drilled carrier. In this process both the chip and carrier have metal bond pads which pair up. One set of bond pads has solder applied, the two parts are then brought into contact and heated. The surface tension of the heated solder acts to pull the pads into a very precise alignment and upon cooling provides the bond. The fibre ribbon is then inserted through the holes in the carrier to butt up to the substrate side of the modulator and glued with a radiation hard epoxy. Wire bonds are used to make electrical contacts to tracks on the carrier which are in turn connected to a lead-frame to which external connections can be made. Finally a lid is bonded over the chip for protection and a serial number badge added for identification. Devices packaged in this way are known as Phase 2 devices. Previous Phase 1 devices used an active alignment process with a lens array. This high cost process resulted in temperature sensitive devices, as the alignment of the chip, lens, and fibre would change, and was thus not suitable for use in the inner tracking detector.

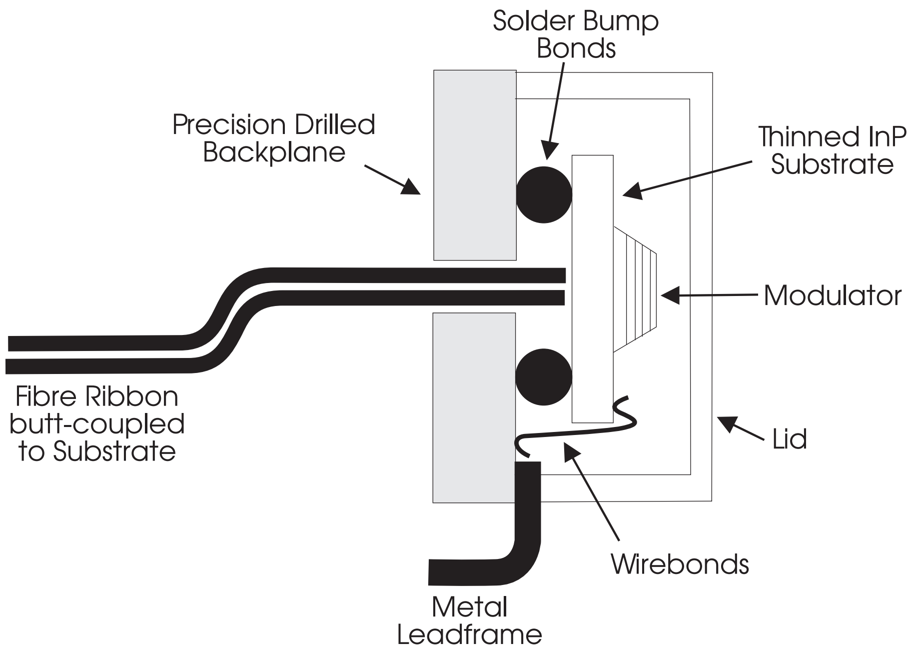


Figure 4.10: Schematic of a Phase 2 packaged MQW modulator.

4.2.3 Transceiver

The transceiver module performs two functions. It is the source of the continuous-wave light that is to be incident on the modulator, and also the receiver which converts the modulated light reflected back into an electrical signal. The basic concept is to have a semiconductor laser diode connected to the first of the pair of arms of a 2 to 1 coupler. The single arm is then attached to the fibre which transmits the light to the modulator. The reflected light returns down the same fibre to the coupler where the second of the pair of arms is connected to a photodiode. The photodiode gives a current response proportional to the power incident upon it. This is amplified using a high gain transimpedance amplifier to give a voltage signal suitable for later electronics.

The transceiver module was still under development when the choice to pursue binary readout was made. It would require significant additional work to progress from the test-bench modules (see section 6.2.6.2) currently available to a commercially produced package suitable for a readout system. An integrated waveguide structure, as shown in figure 4.11, was proposed as the next step in the transceiver development. The design has two lasers that can be modulated for transmission of the digital information. A further three lasers are then used in CW operation each split to power two analogue channels. Six photodiodes are used for reception. All of the active components are spliced onto the pigtailed passive waveguide structure.

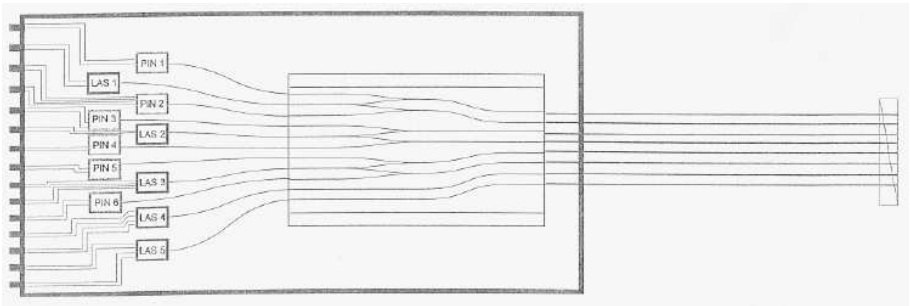


Figure 4.11: Design for integrated transceiver passive waveguide.

In addition, the transimpedance amplifiers would require conversion into more compact semiconductor devices suitable for use in the final transceiver.

4.3 Binary Optical Readout

The readout architecture for the binary optical link [20, 36] is shown in figure 4.12. The hybrid front-end board connects to the detector and holds the digital front-end chips. A separate opto-hybrid holds the optical components:

- Each digital front-end chip processes the signals from 128 strips of the detector as described in section 3.3.2. The data from six front-end chips (one side of a module) are sent to one channel of the LED driver circuit (LDC);
- The LDC contains two independent LED driver circuits. It provides a standing current to the LEDs to improve switching speeds. The output current will be controllable;
- The digital optical receiver integrated circuit (DORIC) receives and distributes the clock, trigger and control signals to the front-end chips;
- The two LED transmitters and PIN[c] photodiode receiver are specially selected to be radiation hard.

All three data links operate at $40M\ bit/s$.

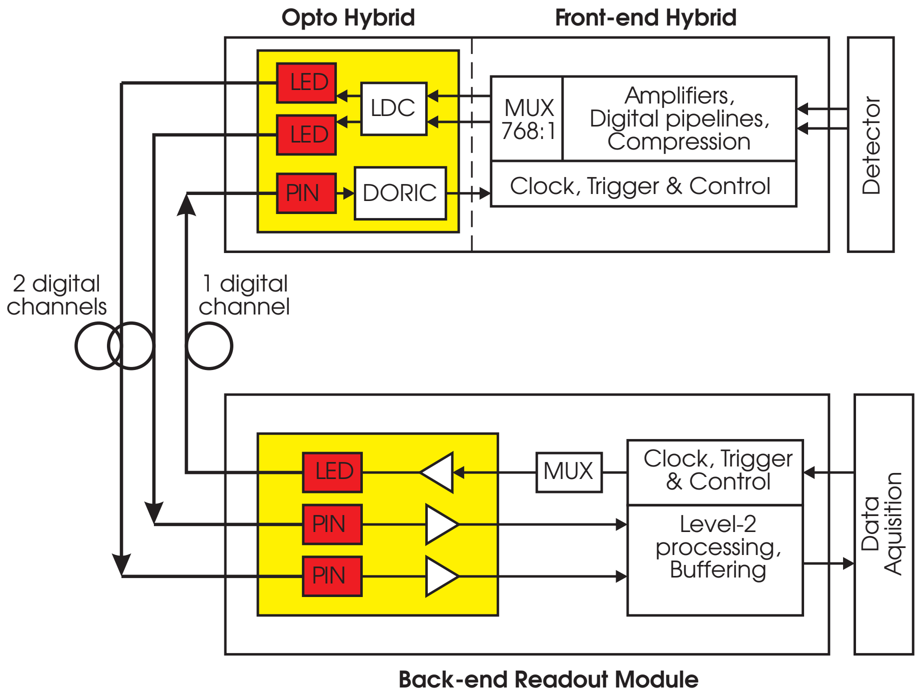


Figure 4.12: Schematic of the proposed binary optical readout system for the ATLAS SCT.

The back-end receiver module contains two receiver channels and one transmitter channel:

[c]PIN photodiodes differ from standard p-doped/n-doped structures by the inclusion of an intrinsic layer to give a wider depletion region resulting in an increased quantum efficiency (see section 5.2.1).

- The optical transmitter and receivers are standard optical components;
- The clock, and control and trigger signals are multiplexed (MUX) and transmitted using one optical channel;
- The back-end module also contains the level-2 trigger buffers required before data is sent to the data acquisition system.

The link components and implementation of the system are now described in more detail.

4.3.1 The Optical Package

The optical package used in the binary readout data link is shown in figure 4.13. It is a custom package manufactured by GEC Marconi UK[d]. It contains the two LEDs and the PIN photodiode as well as providing ground planes for the mounting of the LDC and DORIC chips. The packaging is specifically designed to be low-mass, non-magnetic and radiation hard to match the SCT requirements.

The GaAlAs LEDs produce light at $820nm$ wavelength, and $17\mu W$ optical power output at an operating current of $20mA$. The rise and fall times are $3.5 - 4ns$.

Two types of photodiode have been investigated for use in the package. Bulk Silicon photodiodes used for prototyping are inadequate for the final system. The response, $\sim 0.5A/W$, is sufficient, however the $10\% - 90\%$ rise and fall-times of $15ns$ at $-5V$ bias are inadequate for $40Mbit/s$ operation. The bulk Si diodes are made from a low-doped background n-type Silicon substrate. Dopants are then diffused into this substrate to form a thin p-type layer at the top of the device and a thin n-type layer at the bottom. The light is incident upon the device from the top. The chip thickness, and hence the depletion region thickness is $\sim 100 - 150\mu m$.

Epitaxial Silicon photodiodes exhibit a lower, but still acceptable, response ($\gtrsim 0.35A/W$) and the faster rise and fall-times ($1ns$ at $-5V$ bias) required for operation at $40Mbit/s$. The epitaxial Si diodes are made from an n-type substrate which then has a low-doped background n-type layer followed by a p-type layer grown epitaxially on top. The low-doped region which forms the depletion layer is only $\sim 15\mu m$ thick.

[d]GEC 520/1/02113/000.

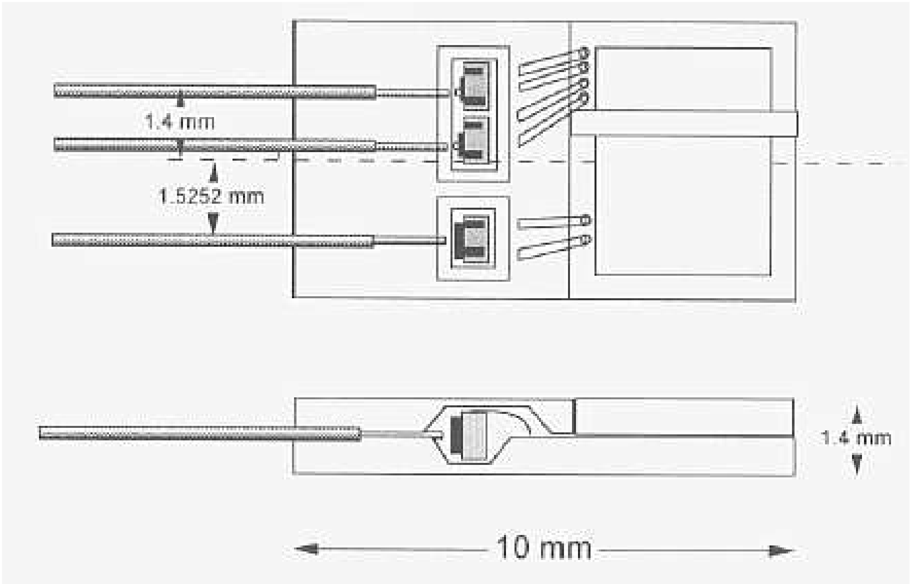


Figure 4.13: Diagram of the custom GEC LED/PIN package.

4.3.2 Redundancy Scheme

To avoid the loss of data due to single point failures, redundancy has been designed into the system. The architecture of a front-end module including the data lines and redundancy is shown in figure 4.14.

Redundancy is incorporated in the distribution of the clock by providing lines from the current module, m, to the next adjacent module, m+1. In the case of a failure in the transmission of the clock to module m it can use the clock provided by module m-1. An identical scheme is used for the redundancy of the trigger and control data.

The redundancy scheme for the transmission of the detector data is complicated by the increased number of serial components in the chain. Each front-end chip processes the data for 128 channels of the detector resulting in six chips for each side of the detector. Each side of the detector is read out using one optical channel. The chips on each side of the detector are linked together in a chain. The first chip in the chain is designated the master chip. On receipt of a level-1 trigger signal the master chip sends its data to the LDC and then passes a token to the next chip in the chain. Upon receipt of the token this chip sends its data to the master chip, which passes it to the LDC, and passes the token to the next chip. Each subsequent chip performs in a similar way sending its data to the previous chip in the chain

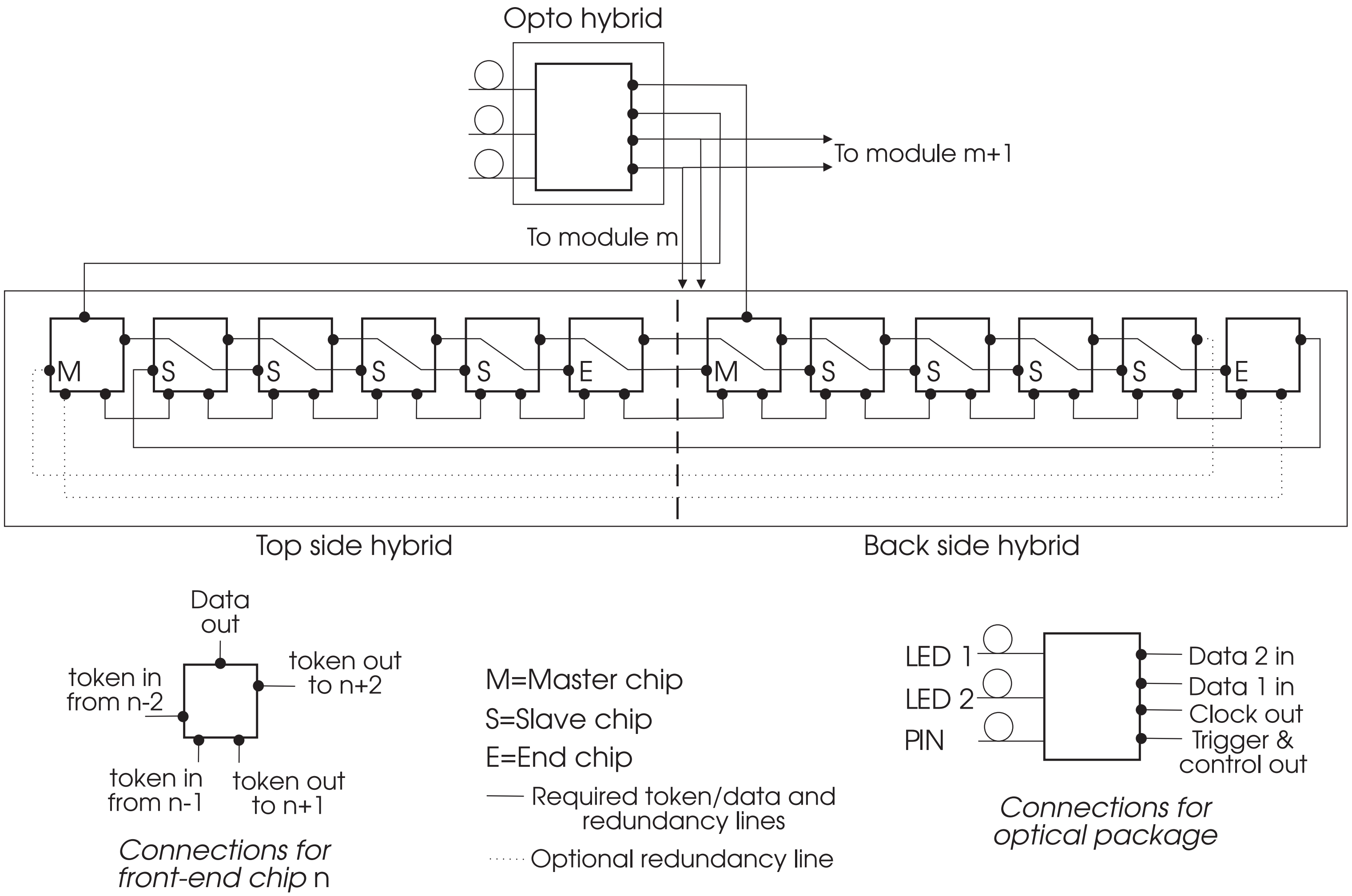


Figure 4.14: Binary front-end architecture and redundancy scheme.

so that it is eventually passed to the master chip and thus the LDC. The last chip in the chain, designated the end chip, attaches a unique data packet to the end of its data. When the master chip receives the end chip data packet it interprets it as the end of data and is then free to begin another readout cycle. This architecture is designed to minimise the number of connections required on the hybrid. The redundancy is incorporated by adding data and token links to the next but one chip in the chain. In the event of failure of any of the slave or end chips of a module the data is only lost from that particular chip and not the chips after it in the chain. In the event of a master chip or data link failure the data for the whole module is read out using the remaining master chip over a single optical link. This results in a slight increase in data loss at a given detector occupancy. The use of the redundant connections is governed by the slow control signals sent from the data acquisition system.

4.3.3 Clock, Trigger and Control Data Encoding

The clock and trigger and control data are transmitted over one optical channel resulting in a saving in terms of both cost and material. A bi-phase mark coding scheme is used to combine the bunch crossing clock (CLOCK) and first level trigger plus slow control signals (DATA) as shown in figure 4.15.

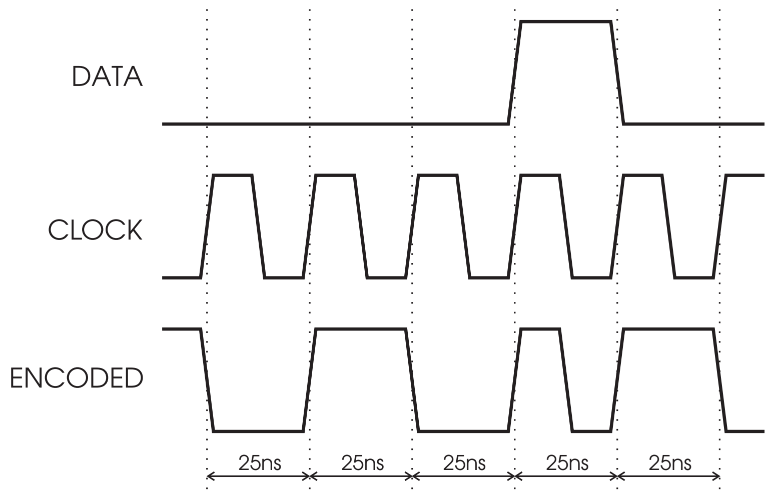


Figure 4.15: DATA signals transmitted as bi-phase mark encoding of CLOCK.

The encoding chip sits in the back-end readout module and therefore suffers from none of the stringent requirements listed in section 3.4. For this reason the design effort has been centred around the decoding chip (DORIC3) that is located at the detector. This chip has been fabricated in a non-radiation hard process, but with a design using the most tolerant components and circuitry that may result in sufficiently radiation-hard devices. Preliminary reports on the function of the DORIC chip are encouraging. It is scheduled to undergo irradiation followed by further testing throughout the summer of 1997.

4.4 Binary Electrical Readout

The link architecture for the binary electrical readout [37] is shown in figure 4.16. It is similar to the binary optical link in design and has been pursued as a safe alternative that has been used for data readout in the past. The digital front-end chip set-up is the same as for the binary optical readout, however the link set-up differs slightly:

- Two channels are used to transmit the data from the detector to the control room. Each channel transmits the data from one side of the detector;
- One channel is used to transmit the clock from the control room to the front-end hybrid;
- One channel is used to transmit the trigger and control signals from the control room to the front-end hybrid.

Each channel is constructed of a twisted pair cable terminated with low voltage differential signal (LVDS) transmitters and receivers. A repeater is necessary to boost and clean the signal part-way between the front-end and control room. The link is described in more detail below.

4.4.1 Low Voltage Differential Signals

The low voltage differential signal standard transmits the signal and its complement over a pair of cables. At the receiver the difference of the two signal is taken. Any electromagnetic noise pick-up on the cables will be of the same sign and magnitude if the cable pair are kept in close proximity along the length. The noise will therefore cancel at the receiver. Low

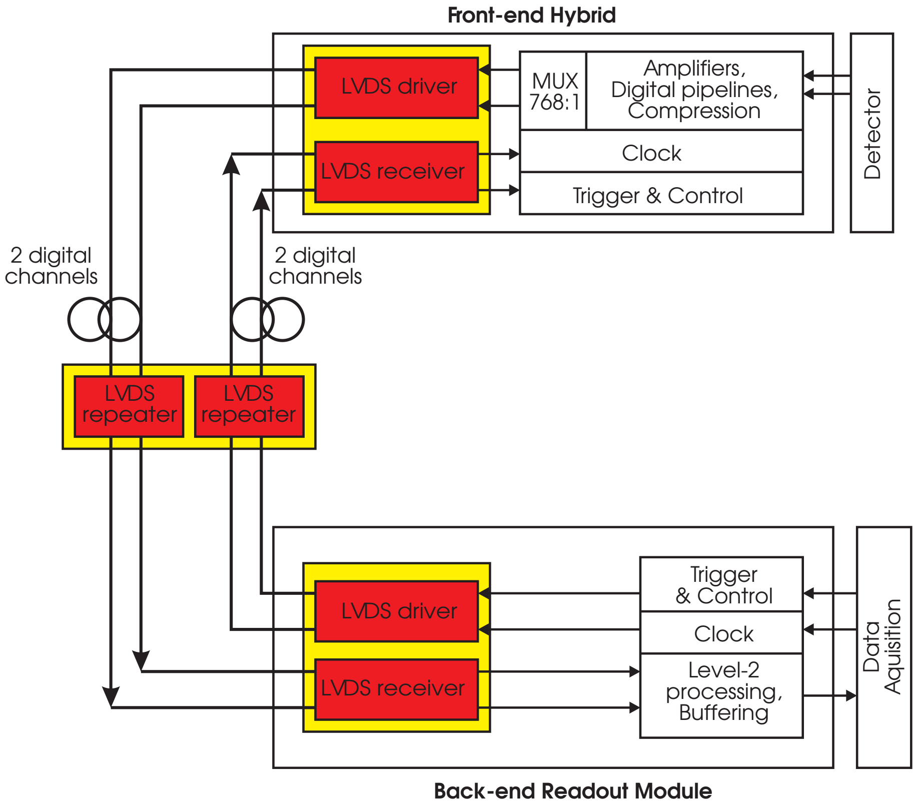


Figure 4.16: Schematic of the proposed binary electrical readout system for the ATLAS SCT.

voltage differences between the signal and its complement, of $\sim 350mV$, are used to reduce the power radiated to neighbouring cables and devices.

4.4.2 Shield Twisted Pairs

Fine gauge (AWG36) cables, as shown in figure 4.17, would be used within the central detector volume. The cables consist of two silicone insulated copper signal wires that are twisted together. A four strand drain wire is twisted around the pair in the opposite direction. An aluminised plastic outer shield surrounds the pair to give an outer cable diameter of $0.9mm$.

Once outside the ATLAS calorimeters standard (AWG28) multi-conductor twisted pair cables (typically 25 twisted pairs per cable) surrounded by a group shield would be used.

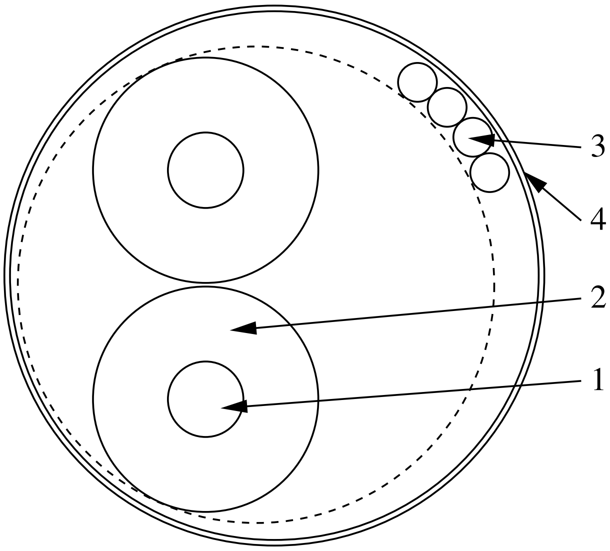


Figure 4.17: Cross-sectional view of AWG36 cable showing the copper conductors (1), silicone insulation (2), 4 strand drain wire (3), and aluminised plastic outer shield (4)[f] [37].

4.4.3 Repeaters

The signal degrades with distance along the cable due to both dispersion and attenuation. Repeaters will be necessary between the transmitter and receiver. The repeater is essentially

[f] Supplied by Habia Cable BV, Voorerf 12, 4824 GN Breda.

a receiver followed by a transmitter. The receiver recovers a clean TTL binary signal which is then converted back to LVDS for retransmission.

Chapter 5

Optical Link Components

This chapter describes the basic optical components that are used in the optical links and test setups in chapters 6 and 7. The discussion starts with the components used to generate the light, followed by the receivers and fibres. Connectors are then discussed and finally other components.

5.1 Light Sources

5.1.1 Light Emitting Diodes

In light emitting diodes (LEDs) [39, 40], light is generated by spontaneous emission. This occurs when an electron in a high energy conduction band combines with a hole in a low energy valence band as shown in figure 5.1. The difference in energy between the two bands is called the bandgap energy, E_g. The energy lost by the electron is released as a photon. The wavelength of the photon is related to the bandgap energy of the material used to make the LED by the formula,

$$\lambda = \frac{hc}{E_g}, \tag{5.1}$$

where, h is the Planck constant, $6.626 \times 10^{-34} Js$,

and c is the speed of light, $2.998 \times 10^{8} ms^{-1}$.

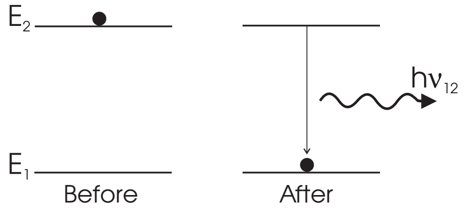


Figure 5.1: Spontaneous photon emission in semi-conductor material.

LEDs emit a broad distribution of wavelengths with typical full width at half maximum (FWHM) values in the range $20 - 80nm$.

5.1.2 Fabry-Perot Lasers

A laser [39,40] differs from an LED because it generates light mainly by the process of stimulated emission as shown in figure 5.2. In stimulated emission a photon triggers additional electron-hole recombinations, resulting in additional photons. A stimulated photon has the important properties that it has the same wavelength and phase as the photon that triggered its generation and will travel in the same direction.

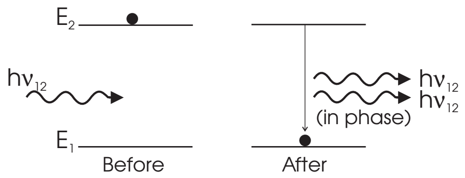


Figure 5.2: Stimulated photon emission in semi-conductor material.

To utilise stimulated emission to produce a laser an 'active' region of population inversion is required, where there are a large number of conduction band electrons and valence band holes in the same region available for stimulation. A cavity is then required to allow the build up of optical power as photons pass repeatedly through the active region. In a Fabry-Perot (FP) laser the cavity comprises semi-transparent mirrors formed by cleaving and polishing the ends of the semiconductor crystal. Only light for which the round-trip path in the cavity is an integral number of wavelengths will add constructively. As a result, the spectrum of a FP laser contains multiple discrete-wavelength components given by the characteristic resonant frequencies of the cavity, f_{res},

$$f_{res} = \frac{mc}{2ln} \tag{5.2}$$

where, m is an integer,

l, is the cavity length,

and n, is the refractive index of the cavity.

The power in each of the wavelength modes will be determined by the laser gain and mirror reflectivity at that wavelength.

5.1.3 Distributed Feedback Lasers

Distributed feedback (DFB) lasers differ from FP lasers in the fact that all but one of the spectral components is significantly reduced. In DFB lasers a grating is constructed along the active layer. Each ridge of the grating then acts as a reflecting surface. At the resonant wavelength all the reflections from different ridges add in phase, therefore only one resonant wavelength is in the region of laser gain. This results in a single lasing wavelength with the largest side modes between 30 and $50dB$ lower than the main spectral output.

5.2 Light Detectors

The main requirements for the photodetector in a fibre communications system are a high sensitivity, a short response time for high speed operation, minimum noise contribution to the signal, high stability and reliability, a small size for efficient fibre coupling, and low cost.

There are several types of optical detectors available, such as photomultiplier tubes, vacuum photodiodes and semiconductor photodiodes which work at the near-infrared region of the spectrum used for optical communications. However, the size of the various devices points to semiconductor photodiodes being the best choice.

5.2.1 Semiconductor Photodiodes

The basic detection process in a pn photodiode [39] is illustrated in figure 5.3. The device is reverse biased and the electric field across the pn junction sweeps the mobile carriers (holes or electrons) to their respective majority sides (p or n type material) forming a depletion region at the junction.

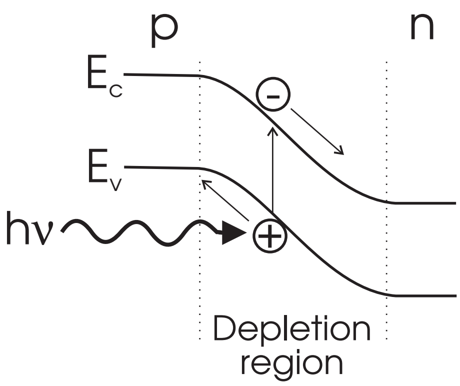


Figure 5.3: Photon absorption in a reverse biased pn junction.

A photon incident in the depletion region which has an energy greater than or equal to the bandgap energy, E_g, of the semiconductor material will excite an electron from the valence band into the conduction band, leaving a hole in the valence band. The photo-generated electron-hole pair are swept from the junction by the electric field giving rise to a reverse current. In practice not all the photons meeting the energy requirement will contribute to the photocurrent. The absorption of the photons follows an exponential law and the depletion region is of limited thickness, however it can be widened by using a *pin* junction structure. In this case an intrinsic region of semiconductor is grown between the doped layers. A tradeoff with the speed of the device, which decreases with increasing thickness, limits the size of the intrinsic region that can be used. Furthermore electron-hole recombination and absorption of photons outside the depletion region will reduce the quantum efficiency of the detector which is given by,

$$\eta = \frac{number\ of\ electrons\ collected}{number\ of\ incident\ photons}. \tag{5.3}$$

Of more practical use perhaps is the responsivity of the detector, $\Re$, which relates the output current, I_p, to the incident optical power, P_o, on the detector,

$$\Re = \frac{I_p}{P_o}. \tag{5.4}$$

The responsivity of the detector is related to the quantum efficiency by the equation,

$$\Re = \frac{\eta e}{h\nu}, \tag{5.5}$$

where, e is the charge on an electron,

h is the Planck constant,

and ν is the frequency of the incident photon.

The responsivity of the photodiode will therefore increase with the wavelength of the incident light.

5.3 Optical Fibres

Transmission of light using a dielectric waveguide structure was first proposed and studied at the beginning of the twentieth century. The first systems investigated consisted of a silica rod with refractive index ~ 1.5 surrounded by air. However, large losses due to discontinuities at the glass-air interface, and the required support structure made it an impractical waveguide. Refinements in the basic idea and technology of fibre construction have led to a cladded waveguide structure, figure 5.4, with losses $\lesssim 1dB\ km^{-1}$ in typical fibres [40].

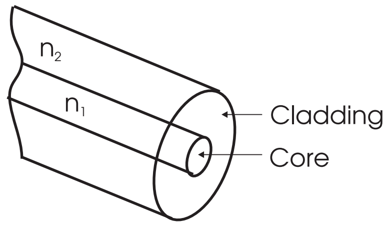


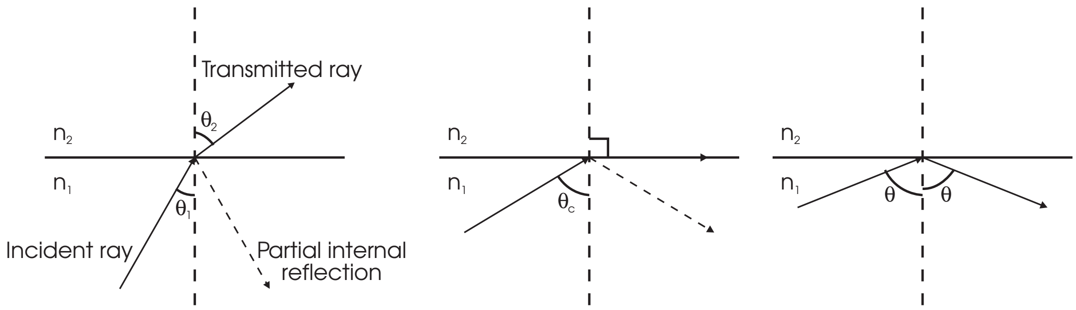
Figure 5.4: Optical fibre waveguide with core of refractive index n_1 surrounded by the cladding of slightly lower refractive index n_2.

The basic principle of light propagation in the optical fibre may be considered in terms of geometrical optics, as illustrated in figure 5.5. When a light ray is incident on a boundary between a dielectric with refractive index n_1 to one with lower refractive index n_2 at an angle θ_1 to the normal at the surface of the interface, then refraction is such that the transmitted light ray in the second dielectric is at an angle θ_2 to the normal, where $\theta_2 > \theta_1$. The angles of incidence, θ_1, and refraction, θ_2, are related to the refractive indices of the dielectrics by

Snell's Law of refraction:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{5.6}$$

A small amount of the light is reflected at the interface back into the originating dielectric (partial internal reflection).



(a) Refraction.

(b) Refraction at the critical angle θ_c.

(c) Total internal reflection for $\theta > \theta_c$.

Figure 5.5: Light rays incident on the boundary between high, n_1, and low, n_2, refractive index dielectrics.

A limiting case for refraction exists when the angle of refraction is 90^o and the refracted ray runs parallel to the interface of the dielectrics. In this case the angle of incidence, now known as the critical angle θ_c, will be less than 90^o and is given by:

$$\sin\theta_c = \frac{n_2}{n_1}. \tag{5.7}$$

At angles of incidence greater than the critical angle the light is reflected back into the original dielectric with high efficiency (total internal reflection). This mechanism can be considered to form the basis of the propagation of light down the optical fibre when it is incident on the core-cladding interface at a sufficiently shallow angle.

5.3.1 Single-mode fibre

Single-mode fibre is designed to allow the propagation of only one transverse electromagnetic mode (typically HE_{11}). The number of modes is restricted by the small core size, $\sim 9\mu m$. The distribution of power within the fibre is Gaussian-like centred on the fibre axis. A significant component of the power is therefore found in the tails of the distribution outside the physical region of the fibre core. The transmission of only a single optical mode offers the advantage of low dispersion (broadening of transmitted pulses). Single-mode fibre is therefore ideal for high bandwidth operations.

5.3.2 Multi-mode fibre

Multi-mode fibre offers distinct advantages over single-mode fibre due to the larger core size, $\geqslant 50\mu m$. These include the use of spatially incoherent sources, such as LEDs, which cannot be efficiently coupled into single-mode fibre, and lower tolerances for fibre alignment at both fibre-device interfaces and fibre-fibre interfaces in connectors.

The propagation of multiple modes limits the system bandwidth due to dispersion. However, in the ATLAS scenario the short distance of the link , only $\sim 100m$, means there will be no problem transmitting signals at the required $40Mbit/s$.

5.4 Fibre Connection

Practical issues, such as ease of device construction and handling, and the final installation of the optical link necessitates joints in the transmission medium. Individual fibres can be joined either permanently using fusion splices or temporarily with connectors. If the light from a fibre requires to be split to more than one fibre (or light from more than one fibre coupled into a single fibre) a device known as a coupler is used. These connection methods are now described in more detail.

5.4.1 Fusion Splices

The fusion splicing [40] of single fibres involves the heating of the two fibre ends with the application of axial pressure between the two fibres. The preparation and alignment of the fibre ends is crucial for a good low-loss joint. The fibre coating (any buffering layers added to the fibre for protection) is first removed and the fibre is then cleaved. The cleaving procedure involves scoring the fibre while it is held under tension. The surface scoring then propagates across the fibre resulting in a smooth square end-face. The fibres are then aligned using micro-positioners, generally with the aid of a microscope, to within a few μm. The fibres are then ready for the fusion step. The ends are heated to their fusing point and brought into contact. Any small misalignment will be corrected by the self-alignment of the fibres due to surface tension. After several seconds the fibres will have merged and the heating can be stopped. The most frequently used heat source is an electric arc which can be easily controlled to take into account the properties of different fibres, such as the fusing temperature and out-gassing temperature. Too low a temperature can result in improper fusing, while too high a temperature can cause out-gassing resulting in bubbles in the joint. With care joints with typical losses of $0.1 - 0.2dB$ can be achieved. The finished joint must then be packaged for protection because of to its lower strength, due to the removal of the fibre coating and effects of heating during the splicing. Multiple ribbon type fibres can be spliced in one operation with the correct cleaving and splicing equipment.

5.4.2 Connectors

Demountable fibre connectors require the same alignment tolerances as fibre splices, but must accomplish it in a removable fashion. The connector must allow for repeated connection and disconnection without loss in performance and must therefore provide reproducible accurate alignment of the optical fibres. Additional factors relevant to connector design are the protection of the fibre ends from damage, the stability and strength of the joint, and of course cost.

For single fibre connections cylindrical ferrule connectors allow joints with losses in the range $0.2 - 0.3dB$ to be made. The fibre is bonded into a precision bored ceramic ferrule. The fibre end will then be polished to give a smooth end face. The end surface is usually finished perpendicular to the fibre axis, however angled end faces can be used to reduce the reflection at the joint. The strength of the connector comes from the outer shell. Typical designs are the straight tip (ST) which is a bayonet type connector, and the fibre connector (FC) which is a screw fitting connector. The joint is made by inserting two connectors into a bulkhead assembly. The bulkhead contains an alignment sleeve which hold the ferrules in line. The two ferrules are butted up to one another to form a physical contact (PC) or, in the case of angled end-face connectors, angled physical contact (APC) connection.

Multiple fibre connectors exist for use with fibre ribbons. The construction techniques are similar to the single fibre connectors. The fibres are bonded into V-grooved silicon chips which include precision metal guide-rods and holes. The end faces of the fibres are then polished as before. A moulded plastic outer is then added to complete the multi-tip (MT) ferrule. Connection is made by butting two connectors together inside a rectangular guide sleeve. The alignment is provided by the guide-rods and the joint is held together with a simple metal clip. A more rugged design uses the MT ferrule mounted within a spring loaded MPO connector outer.

5.4.3 Couplers

Optical fibre couplers [40] distribute light from one fibre into another. Figure 5.6 shows the generic configuration for a coupler with n input ports and m output ports, along with two specific coupler examples. Combinations relevant to work carried out in this thesis are: three port couplers used either, with one input port and two outputs, as a splitter or, with the opposite port arrangement, as a combiner; 2×2 couplers; and 1×4 couplers. The various loss parameters associated with couplers are now discussed. The excess loss, in decibels (dB), which is defined as the ratio of the power input to the power output, is given by,

$$Excess\ loss = 10\log\frac{P_1}{\sum_{j=1}^{n} P_j}, \tag{5.8}$$

for the case of a light input power P_1 injected into input port 1 of the coupler, where we sum over all the output ports j. The insertion loss is defined as the loss obtained for a particular input port to output port path. Therefore,

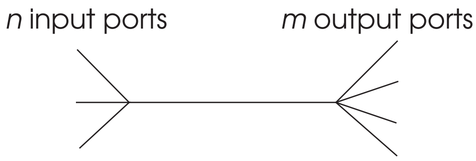
$$Insertion\ loss\ (port\ a\ to\ b) = 10\log\frac{P_a}{P_b}. \tag{5.9}$$

Similarly the directional isolation is defined as the loss obtained going from one input port to another. Hence,

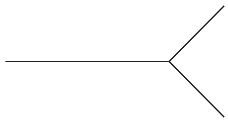
$$Directional\ Isolation(port\ a\ to\ a') = 10\log\frac{P_a}{P_{a'}}. \tag{5.10}$$

Finally, the splitting or coupling ratio describes the percentage division of power to an output port,

$$Splitting\ ratio\ (output\ port\ b) = \left(\frac{P_b}{\sum_{j=1}^{n} P_j}\right)\times 100\% \tag{5.11}$$



(a) Generic coupler.



(b) Splitter.

(c) 2×2 coupler.

Figure 5.6: Optical fibre coupler configurations.

Two fabrication methods for the construction of couplers are described below. In fused biconical taper (FBT) technique the fibres are twisted together and then spot fused under tension to form an elongated biconical taper structure. Normally the cladding of an optical fibre is of sufficient radius that the power at the edge of the fibre is zero, however the taper structure brings the fibre cores closer together. Optical power launched into one of the input fibres will slowly transfer into the other fibre. The total power transfer is a function of the separation of the fibre cores, which controls the rate of power transfer, the length of the tapered section and the wavelength of the light. Couplers fabricated from just two fibres, i.e. three port and 2×2 couplers, can therefore be fashioned to give any desired (wavelength dependent) power ratio at the output, the most common being 50 : 50 or 90 : 10. Couplers with more fibres are designed to give even power splitting between the outputs.

The second fabrication method is the optical waveguide coupler. This method utilises linked waveguides fabricated on a suitable substrate. Different technologies are available but two examples are the use of photo-lithographic methods followed by ion exchange to form buried waveguides within a glass substrate, and silica waveguides deposited on silicon substrates. Both methods offer the attraction of fabricating integrated arrays of couplers on the same substrate. Optical fibres from devices can then be joined to the end of the waveguides using v-groove alignment techniques.

5.5 Polarisation Controllers

The HE_{11} mode in single-mode fibres can be decomposed into two orthogonal polarisation states. Thermal and mechanical stresses or irregularities in the core cross-section can result in the modification of the polarisation state in the fibre.

Described below is a polarisation controller [41] which allows any input polarisation state to be altered to give any desired output polarisation state. The device utilises the birefringence change in the fibre when it is bent to delay the propagation of one polarisation mode relative to the other. A loop, as shown in figure 5.7, stresses the fibre in such a way as to produce a fast extraordinary axis xx' and a slow ordinary axis yy'. The changes in the refractive index relative to that of the originally isotropic medium are given by,

$$\begin{aligned} \Delta n_x &= n_e - n = \frac{n^3}{4}(p_{11} - 2\sigma p_{12})\left(\frac{r}{R}\right)^2, \\ \Delta n_y &= n_o - n = \frac{n^3}{4}(p_{12} - \sigma p_{12} - \sigma p_{11})\left(\frac{r}{R}\right)^2, \end{aligned} \tag{5.12}$$

where, n is the refractive index of the fibre,

p_{11}, p_{12} are the only two non-zero terms in the photo-elastic tensor p_{ijkl},

σ is Poisson's ratio,

r is the radius of the optical fibre,

and R is the radius of the fibre loop.

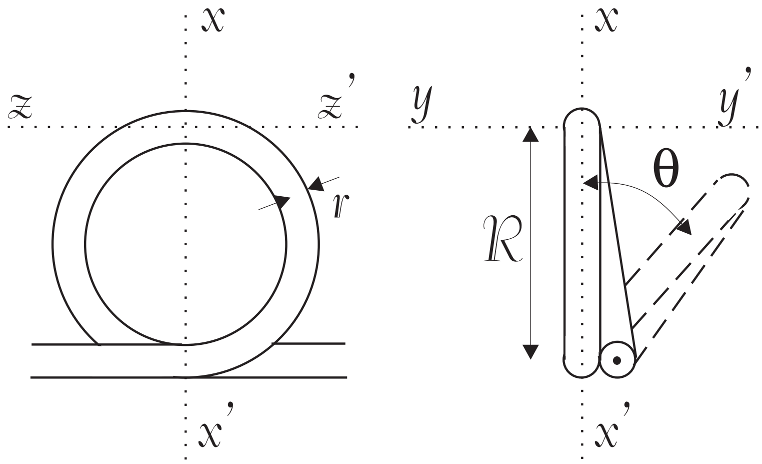


Figure 5.7: Definition of axes and dimensions of the fibre loop.

The refractive index difference between the two axes is given by,

$$\begin{aligned}\delta n &= \Delta n_x - \Delta n_y \\ &= \frac{n^3}{4}(1+\sigma)(p_{11}-p_{12})\left(\frac{r}{R}\right)^2 \\ &= a\left(\frac{r}{R}\right)^2.\end{aligned} \tag{5.13}$$

For example, in a silica fibre σ is 0.16, and at $633nm$ p_{11} is 0.121, p_{12} is 0.270 and n is 1.46. This results in a value for a of -0.133. This effect is small when compared to the birefringence effect seen in crystals, however it is integrated along the length of the fibre and so can become significant. After passing through length L of fibre the phase difference, ϕ, between the fast and slow modes will be,

$$\phi = \frac{2\pi}{\lambda}\delta n L, \tag{5.14}$$

where λ is the wavelength of the light.

Therefore, by combining equations 5.13 and 5.14, the phase delay for N turns of fibre can be calculated,

$$\phi = \frac{2\pi}{\lambda}a\left(\frac{r}{R}\right)^2 2\pi NR. \tag{5.15}$$

The values for N and R can therefore be chosen such that the phase delay between the modes is π or $\pi/2$, to form the interesting cases of half and quarter-waveplates respectively. The required conditions are given by,

$$\frac{\lambda}{m} = \frac{2\pi N a r^2}{R}, \tag{5.16}$$

where m is equal to 2 for a $\lambda/2$ waveplate and 4 for a $\lambda/4$ waveplate.

Rotating the plane of the fibre loop alters the orientation of the fast and slow optical axes with respect to the input polarisation state thereby altering the output polarisation state. A quarter-waveplate can be used to alter the ellipticity and therefore allows the input polarisation to be altered between linear and other states, and vice-versa. A half-waveplate can be used to rotate the orientation of linear polarisation states. A combination of two quarter-waveplates separated by a half-waveplate therefore allows total control of the output light polarisation.

5.6 Optical Isolators

Optical isolators are used to minimise the effects of optical feedback. This is usually of importance for optimal noise operation of a semi-conductor laser diode. The isolator has low loss in the forward direction, but high loss in the backward direction. Several methods to provide optical isolation are described below.

5.6.1 Quarter Wave-plate Isolator

The configuration of, and light polarisation through a quarter wave-plate isolator is shown in figure 5.8. The input polariser of the isolator picks the vertical polarisation state. The quarter wave-plate is oriented to give a left-hand circular polarised output. Upon reflection circular polarised light changes its handedness. Right-hand polarised light incident back on the quarter wave-plate after reflection will be changed to the horizontal polarisation state. The polariser will therefore block the returning light. This type of polariser however has several drawbacks,

- the input beam must be vertically polarised to avoid a large insertion loss,
- the subsequent components in the light path must work with circularly polarised light,
- there must be no de-polarisation of the light from its circular polarisation state.

The final flaw limits the use of the isolator in an optical fibre system due to the birefringence produced in stressed fibre. Expensive polarisation maintaining fibre would be required to enable this type of isolator to be used in a fibre system.

5.6.2 Faraday Rotator Isolator

In certain optical materials, containing magnetic atoms or ions, the circular polarisation states form the natural modes of propagation. When a magnetic field is applied to the material, along the direction of the light propagation, the two modes propagate with different phase velocities resulting in a rotation in the output polarisation state. However, the direction of rotation of the light is dependent on the direction of the applied magnetic field (or that of spontaneous magnetisation) resulting in a non-reciprocal rotation of the polarisation state for light travelling in the reverse direction.

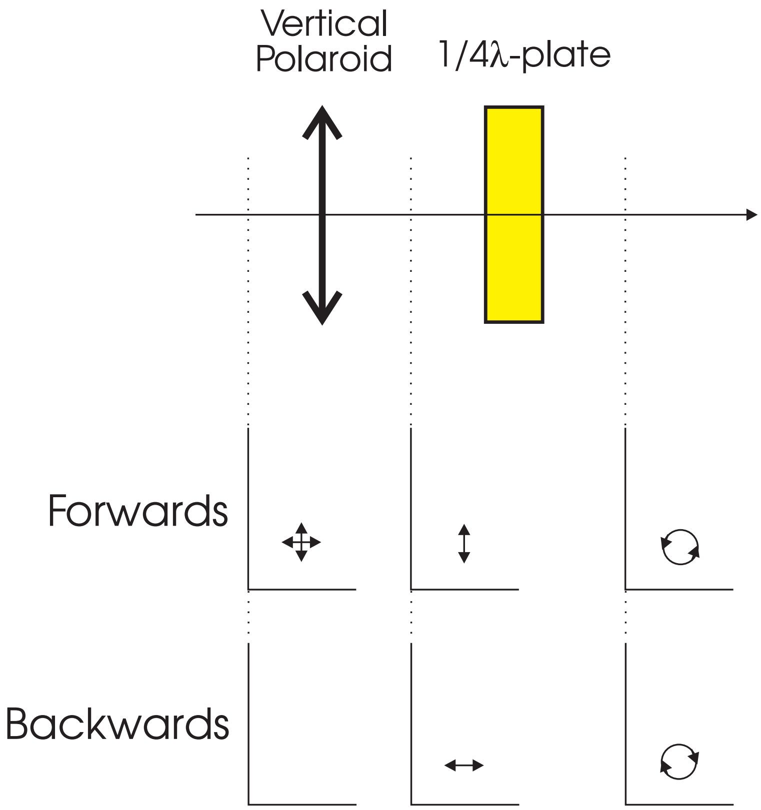


Figure 5.8: Configuration of a quarter wave-plate optical isolator.

An isolator can be constructed from a Faraday rotator and two polarisers, as shown in figure 5.9. The first polariser is oriented to transmit only vertically polarised light. The polarisation state is then rotated clockwise by $\pi/4$ by the Faraday rotator before passing through a suitably oriented output polariser. Reflected light, of any polarisation state, incident back upon the output polariser will be reduced to only the $\pi/4$ linear polarisation state. Passage through the Faraday rotator then produces an anti-clockwise rotation in the reflected light resulting in a horizontally polarised state at the input polariser which blocks the light. Although this design of isolator improves that of the quarter wave-plate device, it still suffers from a insertion loss dependence on the input polarisation state.

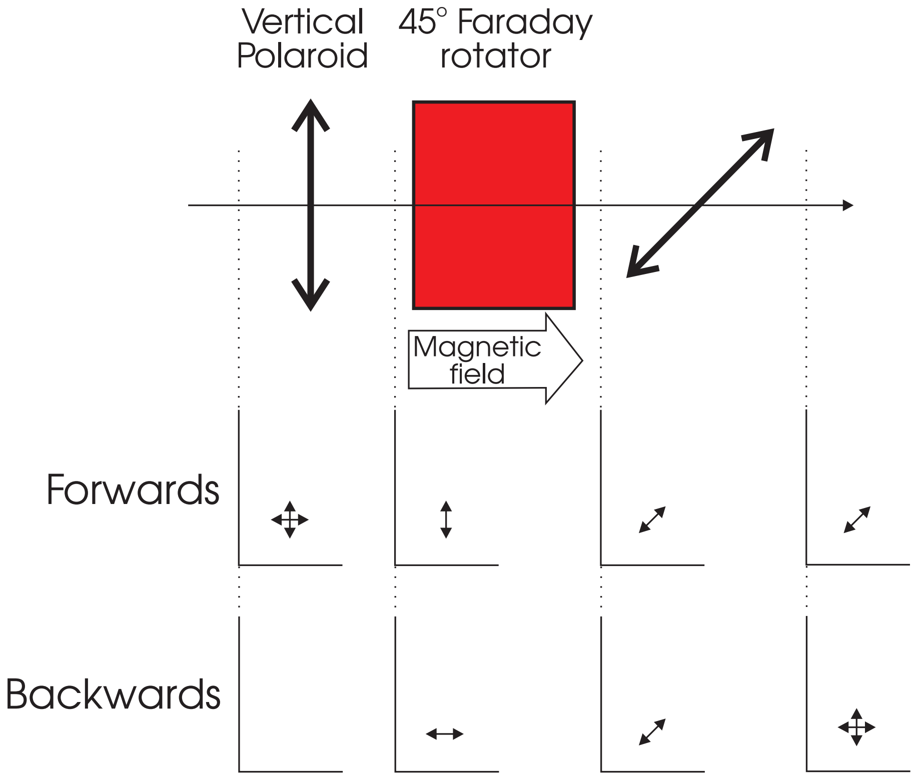


Figure 5.9: Configuration of a Faraday rotator optical isolator.

5.6.3 Polarisation Independent Optical Isolator

A number of designs for polarisation independent isolators exist. In each case the linear polarisation states are split into separate beams, which can then be operated on independently, and then recombined at the output. The separation of the polarisation states is usually achieved using a birefringent crystal to produce a spatial walkoff of one polarisation state, as shown in figure 5.10. Two configurations of optical isolators using spatial walkoff polarisers (SWP) are shown in figures 5.11 and 5.12. The polarisation states of the light throughout the isolator are shown in both cases.

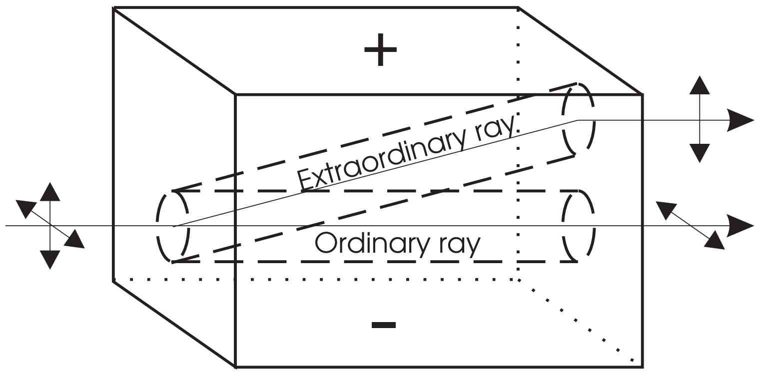


Figure 5.10: Spatial walkoff in a birefringent material. In the forward direction the extraordinary ray is separated from the ordinary ray. In the backward direction the two states would be brought back together.

The first isolator design provides isolation via the spatial position of the light [42, 43]. In the forward direction both polarisation states are brought back together to be coupled into the output fibre. However, in the reverse direction the light is shifted so as not to be coupled back into the input fibre. The non-reciprocal behaviour of the device is introduced by the Faraday rotator.

The second isolator provides its isolation by the use of polarisers [44]. An SWP is used to split the light into the two linear polarisation states. A half wave-plate is then used on the horizontal polarisation state to give a vertical polarisation state. The parallel, but separated beams are then incident on cascaded polarisation dependent Faraday isolators. A further half wave-plate is used to alter the polarisation state of one beam (the opposite one to previously) followed by a further SWP to combine the polarisation states into the output fibre. Reflected light will be stopped at the Faraday isolators.

The isolation achieved using the spatial technique is highly dependent on the alignment of the components. The second method is therefore preferred.

Whatever the isolator design it must also provide a low reflection at its input if it is to limit effectively the light power returned up the input fibre. Angle cut fibres with anti-

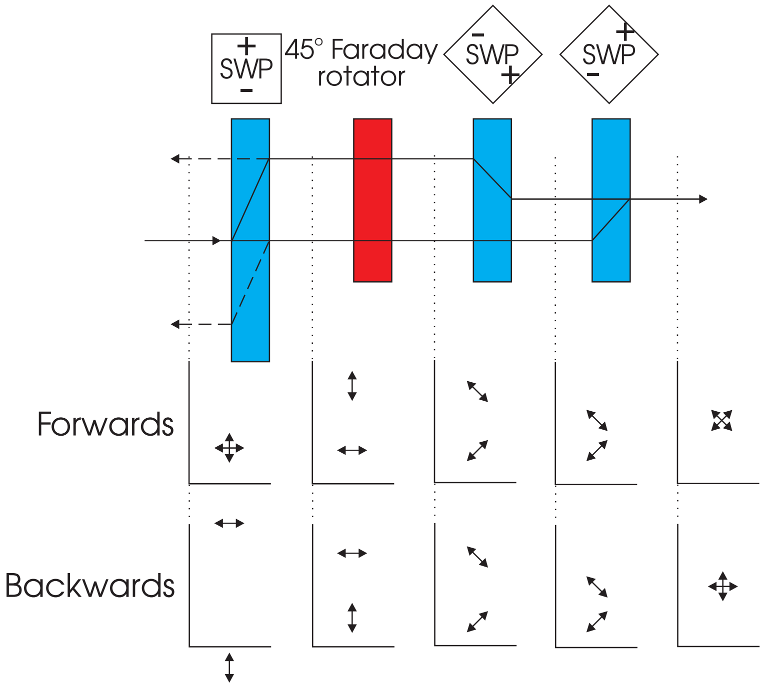


Figure 5.11: Configuration of a polarisation independent optical isolator using a spatial isolation technique.

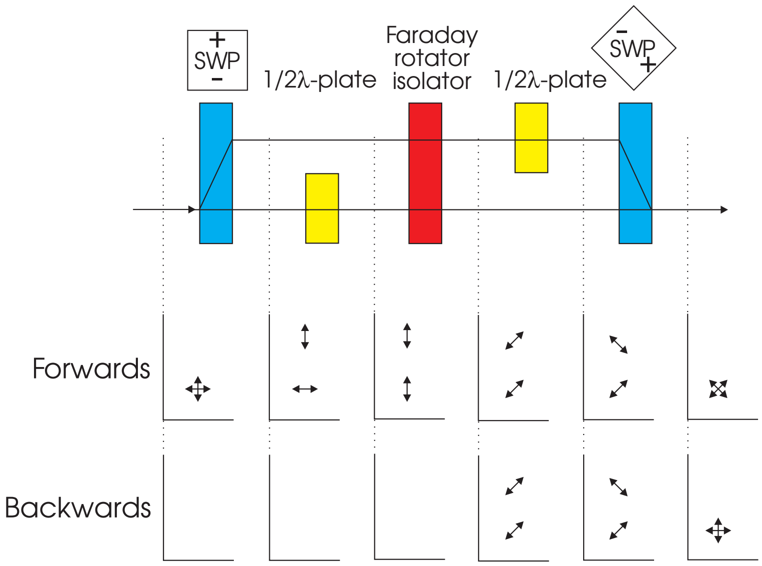


Figure 5.12: Configuration of a polarisation independent optical isolator using a polarisation isolation technique.

reflection coatings are therefore used.

Chapter 6

Analogue Optical Link

This chapter covers work carried out evaluating the analogue performance of the link, the radiation hardness of the MQW modulator, and presents results of the link being used to read out real data.

The present test setup is described and improvements expected in future generations of the system are discussed. A detailed investigation of the noise sources present in the system is described which has great importance for the implementation of a cost effective final system.

The performance of the MQW modulator during and after irradiation with both neutrons and protons is evaluated. Finally a report on the use of the system to read out Silicon microstrip detectors during an ATLAS beam test is presented.

6.1 Analogue Performance Tests

6.1.1 Link Parameters

The main parameters of interest in an analogue link are:

- Linear range:- This is the input voltage range over which the link can transmit a signal within a given distortion limit. It is found by measuring the transfer function (back-end transceiver voltage versus front-end modulator input voltage), fitting a straight line through the inflection point, and determining the input voltage range for a $\pm x\%$ maximum deviation from the fitted line;

- Link gain:- The slope of the transfer function at the inflection point;
- Noise:- The electrical noise after the receiver amplifier output.

The combination of these parameters allows the performance of the link to be described in terms of the signal to noise ratio per volt(SNR_{1V}) and peak signal to noise ratio (SNR). The SNR_{1V} is calculated by dividing the link gain by the noise, and the SNR is given by the product of the SNR_{1V} and the linear range.

6.2 Optical Link Noise Theory

There are several noise sources that manifest themselves in optical links, which can be split into three categories:

1. noise in the reception of the light: thermal noise and shot noise;
2. noise in the generation of the light: relative intensity noise (RIN), reflection induced RIN (RIR) and phase to amplitude noise (PAN);
3. noise due to refractive index discontinuities in the light path: RIR and PAN.

These noise components are described below. They are calculated in terms of the equivalent root mean square (RMS) current, $< i_T^2 >^{\frac{1}{2}}$, at the input of the receiver amplifier circuits in $AHz^{-\frac{1}{2}}$. The root mean square voltage noise at the output of the amplifier is given by,

$$< noise^2 >^{\frac{1}{2}} = < i^2 >^{\frac{1}{2}} \Delta f^{\frac{1}{2}} G, \tag{6.1}$$

where, Δf is the amplifier bandwidth,

and G is the amplifier transimpedance gain.

6.2.0.1 Thermal Noise

Thermal or Johnson noise [38] arises from the random thermal motion of charge carriers in resistive elements in the receiver amplifier circuits. Although the element as a whole stays neutral, random motion of electrons result in fluctuating charge densities and field gradients, which give rise to random noise currents and voltages. The root mean square thermal noise current is given by,

$$< i_T^2 >^{\frac{1}{2}} = \sqrt{\frac{4k_B T}{R_L}}, \tag{6.2}$$

where, k_B is Boltzmann's constant ($1.38 \times 10^{-23} JK^{-1}$),

T is the absolute temperature,

and R_L is the effective load resistance of the amplifier.

6.2.1 Signal Shot Noise

The shot or quantum noise [38] is a direct result of the discrete nature of light. The randomness of photon arrival times at the detector gives rise to a similar randomness in carrier generation. Photon absorption and carrier generation is also a statistical mechanism introducing further randomness into the detection process. These noise sources are statistically similar and can be combined into a single source of shot noise described by Poisson statistics.

The optical detection circuit can be regarded as a low-pass filter with integration time T. The net charge in the filter at a given time, t, is dependent on the number of electrons, n, accumulated during the previous integration time T. The instantaneous photo-current is defined as

$$I_p = \frac{ne}{T} \tag{6.3a}$$

where, n is a random function in time described by Poisson statistics

and e is the electronic charge.

The average photocurrent is then,

$$< I_p > = \frac{< n > e}{T}. \tag{6.3b}$$

Therefore the mean square current fluctuation due to shot noise is,

$$< i_s^2 > = < (I_p - < I_p >)^2 > = \left(\frac{e}{T}\right)^2 < (n - < n >)^2 > . \tag{6.3c}$$

However, for a Poisson distribution, with large n, we take the variance to equal the mean,

$$< (n - < n >)^2 > = < n >, \tag{6.3d}$$

therefore equation 6.3c becomes,

$$< i_s^2 >= (\frac{e}{T})^2 < n > . \tag{6.3e}$$

Using equation 6.3b we then obtain,

$$< i_s^2 >= \frac{< I_p > e}{T}. \tag{6.3f}$$

The noise equivalent bandwidth, Δf, of a filter with integration time T is $\frac{1}{2T}$. Therefore the mean square shot noise current, in terms of bandwidth is given by,

$$< i_s^2 >= 2 < I_p > e\Delta f, \tag{6.3g}$$

and hence the equivalent RMS current at the input to the amplifier is,

$$< i_s^2 >^{\frac{1}{2}}= (2 < I_p > e)^{\frac{1}{2}}. \tag{6.3h}$$

6.2.2 Dark Current Shot Noise

The dark current is the current in the photo-detector when there is no incident light. There are two components: the bulk dark current, I_b, and the surface dark current or leakage current, I_l. The former is due to thermally generated carriers in the bulk of the material, while the latter results from surface defects, dirt for instance, which provide free electrons to sustain a surface current. Both processes are shot-like following Poisson statistics and so take the same mathematical form. The RMS bulk dark current shot noise and surface leakage current shot noise are given by [38],

$$< i_b^2 >^{\frac{1}{2}}= (2 < I_b > e)^{\frac{1}{2}} \qquad < i_l^2 >^{\frac{1}{2}}= (2 < I_l > e)^{\frac{1}{2}}, \tag{6.4}$$

respectively.

6.2.3 Relative Intensity Noise (RIN)

The power output from lasers is always fluctuating. This may be due to temperature variations, acoustic vibrations, and other man-made causes. However even if these extraneous effects are eliminated, there remains a basic contribution that is due to spontaneous emission of radiation into the laser mode by atoms dropping from the transition level into lower levels.

The field due to these emissions is not coherent with that of the laser mode and so phase and amplitude fluctuations are produced. The RIN [38, 45] is defined as the relative fluctuation power in a $1Hz$ bandwidth,

$$RIN = \frac{<(P_L - <P_L>)^2>}{<P_L>^2}, \tag{6.5a}$$

where, P_L is the laser light intensity.

The value for the RIN is dependent upon the specific laser used. This noise decreases as the laser drive current, I_B, is increased according to the relationship,

$$RIN \propto \left(\frac{I_B}{I_{th}} - 1\right)^{-3}, \tag{6.5b}$$

where, I_{th} is the laser threshold current.

In the transceiver application the laser is run well above the threshold current so the RIN will be near its minimum.

Since the spontaneous emissions are random they are described by the same Poisson statistics as the shot noise. Following similar arguments the instantaneous photo-current will be,

$$I_p = \frac{e\eta P_L}{h\nu}, \tag{6.5c}$$

where, η is the quantum efficiency (electrons per photon) of the detector,
h is Planck's constant ($6.63 \times 10^{-34} Js$),
and ν is the frequency of the light.

The mean square noise due to RIN is therefore,

$$<i_{RIN}^2> = \left(\frac{e\eta}{h\nu}\right)^2 <(P_L - <P_L>)^2>. \tag{6.5d}$$

Substituting for 6.5a and taking the square root we obtain,

$$<i_{RIN}^2>^{\frac{1}{2}} = \frac{e\eta <P_L>}{h\nu} RIN^{\frac{1}{2}} = <I_P> RIN^{\frac{1}{2}}. \tag{6.5e}$$

6.2.4 Reflection Induced RIN (RIR)

Any light reflected back into the the laser can significantly increase the laser RIN [45]. In effect, an additional lasing cavity is produced which, for a reflection at distance L from the laser, gives rise to intensity fluctuation noise whose frequency spectrum shows periodic peaks. The frequency, f_p, of the fundamental component of the noise spectrum is nearly equal to the reciprocal of the round trip time in the fibre, as is given by,

$$f_p \simeq \frac{c}{2nL}, \tag{6.6}$$

where, c is the speed of light,

and n is the refractive index of the fibre.

Typically with back reflections of $\sim -70dB$ the laser RIN will be within a few $dB\ Hz^{-1}$ of its intrinsic level. As the intensity of the back reflection is increased up to $\sim -50dB$ the RIN also increases almost linearly, before saturating [45].

6.2.5 Phase to Amplitude Noise (PAN)

The interference of light in the direct path from the laser with that which has been delayed in time, by travelling a longer path, gives rise to noise by the following mechanism. The phase noise of the laser due to spontaneous emission, as discussed in section 6.2.3, means that the phase between the direct and delayed terms is continuously changing. The interference therefore changes the phase noise, which would not be detected at the receiver, into amplitude noise, which is detected.

The cases of a transmissive link and reflective link are shown in figure 6.1a and figure 6.1b respectively. R_1 and R_2 are the intensity reflection coefficients for two discontinuities in the refractive index leading to reflections. In the transmissive link both discontinuities will generally be due to connectors and, depending on the type, will be $-40 \rightarrow -60dB$. However in the reflective system the second case is also possible. Here R_2 is the mirror of the modulator, which generally will be $\sim -8dB$, making PAN a much larger contribution to the total noise of the system. The time for light to travel distance $2L$ in the fibre is τ. We look at the case for the reflective link here, but compare to that for a transmissive link [46,47].

Taking the electric field at the input to the system to be,

$$e_o(t) = e_o cos(\omega_o t + \phi(t)), \tag{6.7a}$$

where, e_o is the amplitude of the signal,

ω_o is the laser line frequency,

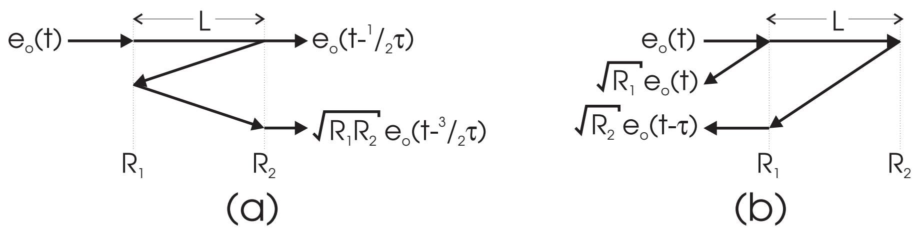


Figure 6.1: Signal terms leading to PAN in a) Transmissive b) Reflective optical links.

and $\phi(t)$ is the phase shift due to spontaneous emission of the light emitted at time t, the electric field at the receiver will be given by,

$$E(t) = e_o[\sqrt{R_1}cos(\omega_o t + \phi(t)) + \sqrt{R_2}cos(\omega_o(t-\tau) + \phi(t-\tau))]. \tag{6.7b}$$

The power ($= E(t)^2$) at the receiver will therefore be,

$$\begin{aligned} P(t) =& e_o^2[R_1cos^2(\omega_o t + \phi(t)) + R_2cos^2(\omega_o(t-\tau) + \phi(t-\tau)) \\ &+ 2\sqrt{R_1R_2}cos(\omega_o t + \phi(t))cos(\omega_o(t-\tau) + \phi(t-\tau))] \end{aligned}$$

$$\begin{aligned} =& e_o^2\Big[\frac{1}{2}R_1[1 + cos(2\omega_o t + 2\phi(t))] \\ &+ \frac{1}{2}R_2[1 + cos(2\omega_o(t-\tau) + 2\phi(t-\tau))] \\ &+ \sqrt{R_1R_2}cos(2\omega_o t + \phi(t) + \phi(t-\tau) - \omega_o\tau) \\ &+ \sqrt{R_1R_2}cos(\omega_o\tau + \phi(t) - \phi(t-\tau))\Big], \end{aligned} \tag{6.7c}$$

by using,

$$cos2\theta = 2cos^2\theta - 1 \qquad cos\alpha\ cos\beta = \frac{1}{2}[cos(\alpha+\beta) + cos(\alpha-\beta)]. \tag{6.7d}$$

Removing high frequency components we find that the current from the photo-detector ($= \frac{e\eta P(t)}{h\nu}$ as described in section 6.2.3) is given by,

$$i(t) \simeq \frac{e\eta}{2h\nu}e_o^2\Big[1 + 2\frac{\sqrt{R_1R_2}}{R_1+R_2}cos(\omega_o\tau + \phi(t) - \phi(t-\tau))\Big], \tag{6.7e}$$

where we have divided through by $R_1 + R_2$ to define the power in terms of that in the direct light path. We compare this to the transmissive link case where the current calculated in the same way is,

$$i(t) \simeq \frac{e\eta}{2h\nu} e_o^2[1 + 2\sqrt{R_2 R_1} cos(\omega_o \tau + \phi(t) - \phi(t-\tau))]. \tag{6.7f}$$

It can be seen that the size of R_1 must be much smaller in the reflective link to obtain the same PAN.

Equations 6.7e and 6.7f are of the general form for an interferometer with a DC term and a fringe term. To calculate the power spectrum we must first obtain the autocorrelation function for equation 6.7e and then Fourier transform, using a Lorentzian line-shape to describe the phase dependence [47].

There are two distinct cases for the power spectrum, the first is the coherent case where the reflected term has a fixed phase in relation to the direct term. The product of ω_o and τ fixes the bias point on the interferometer output, giving two extreme cases. When the interferometer is biased at its quadrature point ($cos(2\omega_o \tau) = -1$) we have,

$$P_Q(f) = \frac{4R_1R_2}{\pi(R_1+R_2)^2} \frac{\Delta\nu}{f^2+\Delta\nu^2} \Big(1 + e^{-4\pi\Delta\nu\tau} - 2e^{-2\pi\Delta\nu\tau} cos(2\pi f \tau)\Big), \tag{6.7g}$$

where, f is the conjugate variable of t and represents radio-frequencies,

and $\Delta\nu$ is the laser linewidth,

giving the maximum conversion to intensity noise. For biasing at the interferometer maximum or minimum ($cos(2\omega_o \tau) = +1$) we have,

$$P_M(f) = \frac{4R_1R_2}{\pi(R_1+R_2)^2} \frac{\Delta\nu}{f^2+\Delta\nu^2} \Big(1 - e^{-4\pi\Delta\nu\tau} - 2e^{-2\pi\Delta\nu\tau} \frac{\Delta\nu}{f} sin(2\pi f \tau)\Big), \tag{6.7h}$$

giving the minimum conversion to intensity noise.

The second case is for τ much longer than the coherence time of the laser. The phase between the reflected and direct light paths is therefore driven rapidly through all the interferometer bias points. In this case both equations 6.7g and 6.7h reduce to,

$$P_I(f) = \frac{4R_1R_2}{\pi(R_1+R_2)^2} \frac{\Delta\nu}{f^2+\Delta\nu^2}. \tag{6.7i}$$

The RMS current noise due to PAN is,

$$< i_{PAN}^2 >^{\frac{1}{2}} = < I_P > P(f)^{\frac{1}{2}}. \tag{6.8}$$

6.2.5.1 Suppression of PAN

The most obvious method to reduce the PAN in a system is to use low reflection components and connectors. The reflection from the MQW modulator is designed to be high, since it is a reflective device, so we cannot remove this reflection. However, the link has two connectors between the modulator and the transceiver to ease the installation into the final LHC detectors. Low reflection angle polished connectors are used in both cases to help minimise the PAN

A second way to suppress the PAN is to add a high frequency modulation to the optical carrier (known as dithering the laser) [48]. This process acts to redistribute the noise from the baseband to higher frequencies where it can easily be filtered out.

The noise power spectrum with phase modulation applied, $P_D^2(f)$, can be shown to be given by,

$$P_D^2(f) = J_0^2(A)P_I(f) + \Big[\sum_{n=1}^{\infty} J_n^2(A)[P_I(f-nf_0) + P_I(f+nf_0)]\Big], \tag{6.9}$$

where, $A = -2a\sin(\frac{\omega_0\tau}{2})$,

$P_I(f)$ is the PAN before the modulation is applied,

$J_n(A)$ are n^{th} order Bessel functions,

a is the phase modulation index (the fractional amplitude of the modulation over the DC term),

and $f_0 = \frac{\omega_0}{2\pi}$ is the modulation frequency.

The distribution of the noise among the different harmonics of the modulating signal frequency is shown in figure 6.2.

The part of the noise spectrum of concern to us is that which remains in the baseband even after the dithering is applied. This noise is reduced by a factor $J_0^2(A)$, which we define as the noise reduction factor (NRF). Practically the phase modulation is obtained by applying a small amplitude modulation to the CW laser light source, known as dithering. This alters the carrier density resulting in a refractive index change and hence phase modulation.

The NRF when dithering is used, shown in figure 6.3, can be seen to be a periodic function of $f_0\tau$. It is therefore dependent upon the separation of the reflection points, since $\tau \propto L$. The maximum reduction depends upon the modulation depth parameter a. It can be seen that large reductions are possible except for the unfortunate situation when $f_0\tau = k$,

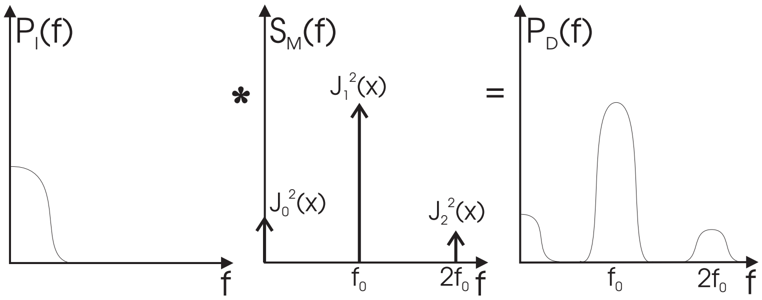


Figure 6.2: Distribution of the noise power among the different harmonics of the modulating signal frequency f_0.

where k is an integer. This means that a dithering frequency may not work for all the channels in an optical link.

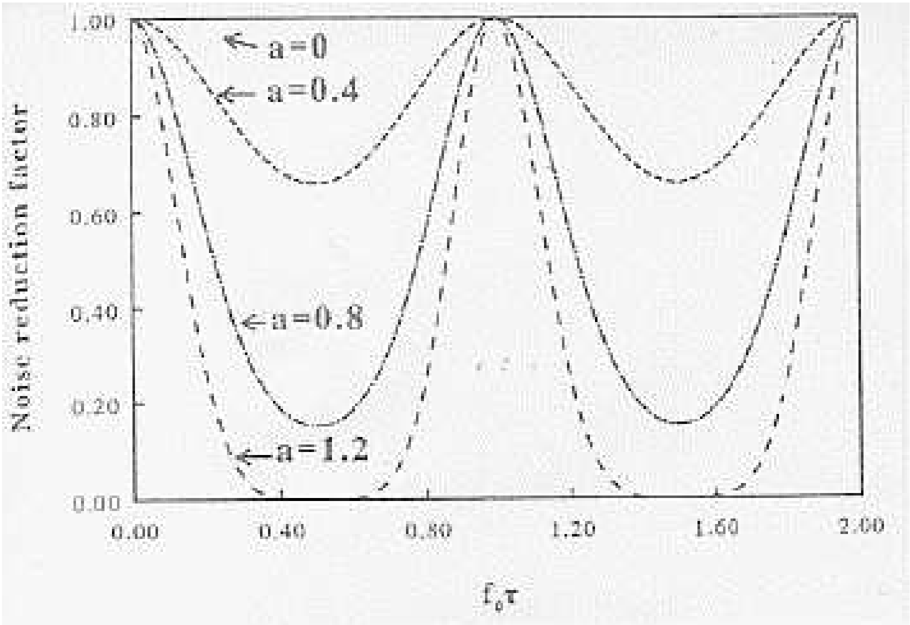


Figure 6.3: Noise reduction factor versus $f_0\tau$ for a directly modulated laser diode [48].

6.2.6 Test Setup

The link components that are being tested in the laboratory are not of the final design required for use in an LHC detector. The test components are described in more detail below.

6.2.6.1 MQW Modulator

The modulator has already undergone two iterations of development. Both the chip packaging and MQW structure were changed between the two phases. The change in the MQW structure was to optimise the reflectivity change of the device. However, a poor performance was achieved and it was later discovered that the new optimal design was located very close to a region where performance dropped rapidly. The tolerances in the growth stages resulted in a device that fell into this lower performance region. The final devices would return to the original phase 1 MQW growth. The low mass packaging design used in the second iteration would however be retained for the final devices. The test modulator has only four channels to reduce development costs by using existing masks for the fabrication of the wafers.

6.2.6.2 Transceiver

The design of the optics part of the transceiver is shown in figure 6.4. This design incorporates a 1×4 splitter which allows a single laser to be used to power all four of the optical channels. The laser is mounted externally and is connected via an optical isolator to the input of the transceiver. This arrangement allows different lasers to be tested easily using the same transceiver unit. This is useful since different types of lasers have different intrinsic noise levels, line-widths (which affects PAN), and stability in response to light reflected back into them.

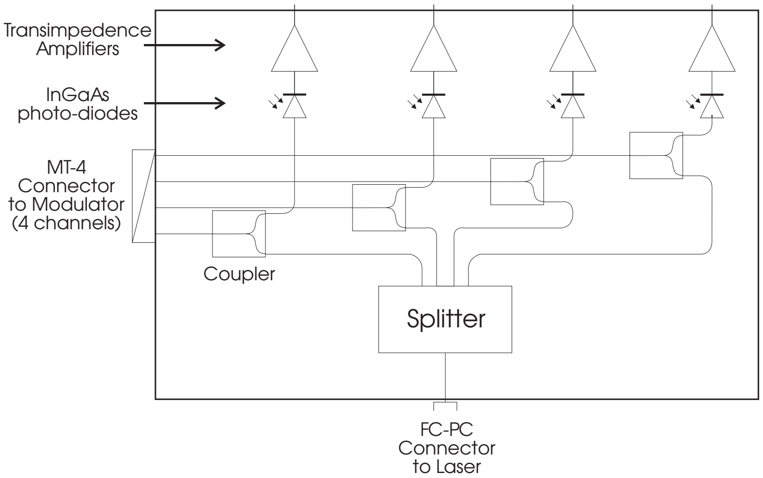


Figure 6.4: Schematic of optical part of transceiver module

All of the optical components (1×4 splitter, 4 off 2×1 couplers, 4 off pin diodes, FC (single fibre) and MT (multi-fibre) connectors) were of a pigtailed form allowing fusion spliced connections. Fused type couplers and splitters were used for this transceiver. Pigtailed waveguide type components have been used to prototype the integrated waveguide transceiver required in the final system. Comparable performance is seen. The transimpedance amplifiers used in the test transceiver use discrete components. Again integrated circuit versions would be required for the final system.

6.2.6.3 Lasers

The transceiver design for the final system uses one laser to power two optical channels. This arrangement is chosen to allow the required power in each channel to be met without the need for specially selected lasers, and to miniaturise the integrated transceiver design by replacing the 1×6 splitter with three parallel 1×2 splitters which require less room. Both should result in a cost saving.

The laser used in the test setup was a DFB laser manufactured by GEC Marconi Materials Technology type LD4909. The laser operated at $1540nm$ and produced a maximum power of $\sim 10\mu W$. The laser was operated in a dithered mode to reduce the effects of PAN. This should be avoided in the final system due to the complications it would involve for the installation.

6.2.6.4 Optical Isolator

An optical isolator is used in the test setup to remove reflections back into the laser. Cost considerations would require the removal of this component in the final system.

6.2.6.5 Connectors

The final system would require three connectors between the transceiver and MQW modulator to be compatible with the installation scheme of cables within the ATLAS detector. All other joints between components, such as laser-transceiver interfaces would be direct. The connectors used would be multi-fibre MT-ferrules in a high density carrier. The fibre end-faces would be angle polished to minimise reflections in the system.

The test setup incorporates several flat polished connectors between components and two angle polished MPO connectors between the transceiver and MQW modulator.

6.2.7 Modulator Tests

In May 1995 eight Phase 2 devices were delivered which nominally had all four channels operating up to specification. The devices were tested as a set under the same operating conditions in collaboration with the CERN group [30].

The transceiver unit at CERN uses 2 to 2 couplers in place of the 2 to 1 couplers in the Birmingham device. In normal operation the spare ends are index matched to minimise reflections, but this set-up allowed one transceiver to be connected in turn to each channel of the MQW modulator using a fan-out as shown in figure 6.5. This is still possible using the Birmingham transceiver, but would require 2 fan-outs leading to more connectors and thus more attenuation and noise. This setup allows the testing of the modulators without the need to correct for the different characteristics of each transceiver channel.

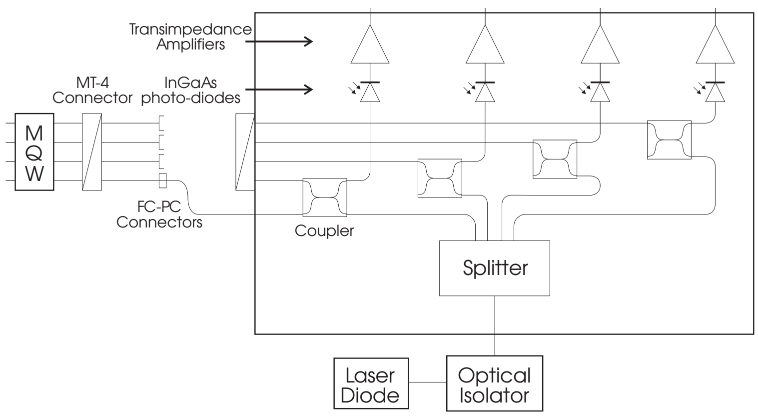


Figure 6.5: Set-up used for testing of phase 2 modulators.

The following parameters were measured:

- Bias voltage:- The DC voltage required to operate the device in the centre of its linear range (this is measured at a monitor point, the actual voltage on the chip is $\frac{10}{11}$ of this;
- Link gain:- Measured by applying a sinusoidal modulating voltage of 1V peak to peak and observing the peak to peak voltage at the transceiver. This gives a good estimate of the link gain if the bias voltage is set for operation in the centre of the linear range;

- Noise:- Measured in a $30MHz$ bandwidth and over a $1ms$ time period using the RMS function on a digital oscilloscope. The RMS (root mean squared) function calculates the RMS noise of the data points on the oscilloscope display;

- Linear range:- The transfer function is first measured by applying sinusoidal modulating voltage to the modulator of sufficient amplitude to cover the linear range of the device. A digitising oscilloscope is used to monitor the modulating voltage and the output of the transceiver. The trace data is transferred to a computer running LabVIEW using a General Purpose Interface Bus (GPIB). The trace data are plotted with the modulating voltage on the x-axis and the transceiver voltage, shifted by an appropriate phase difference on the y-axis to give the link transfer function. Examples of which are shown in figure 6.6.

  To obtain the linear range for the modulator channel a straight line is fitted through the inflection point of the transfer function. The gradient of this line is altered to give the largest input voltage range that gives $< 1\%$ deviation between the fitted line and the data.

The link gain, bias voltage and linear range for all the channels tested are histogrammed in figure 6.7. The spread in link gain is characteristic of an immature packaging technology. It is dependent on the device insertion loss and connector mating losses. The spread in bias voltage and linear range could be due to cavity inhomogeneities at the chip, package or system level. More understanding of the fabrication and packaging of the MQWs would be required to produce good device yields in production. The RMS noise measured for all the channels was $\sim 6mV$.

6.2.8 Link Tests

The peak signal to noise ratio for the optical link is expected to be dependent on the optical power incident on the modulator, since the signal at the transceiver rises linearly with the intensity of the light incident upon it, but the total noise due to various sources varies in a more complex way.

The signal and noise for various laser drive currents and therefore optical powers has

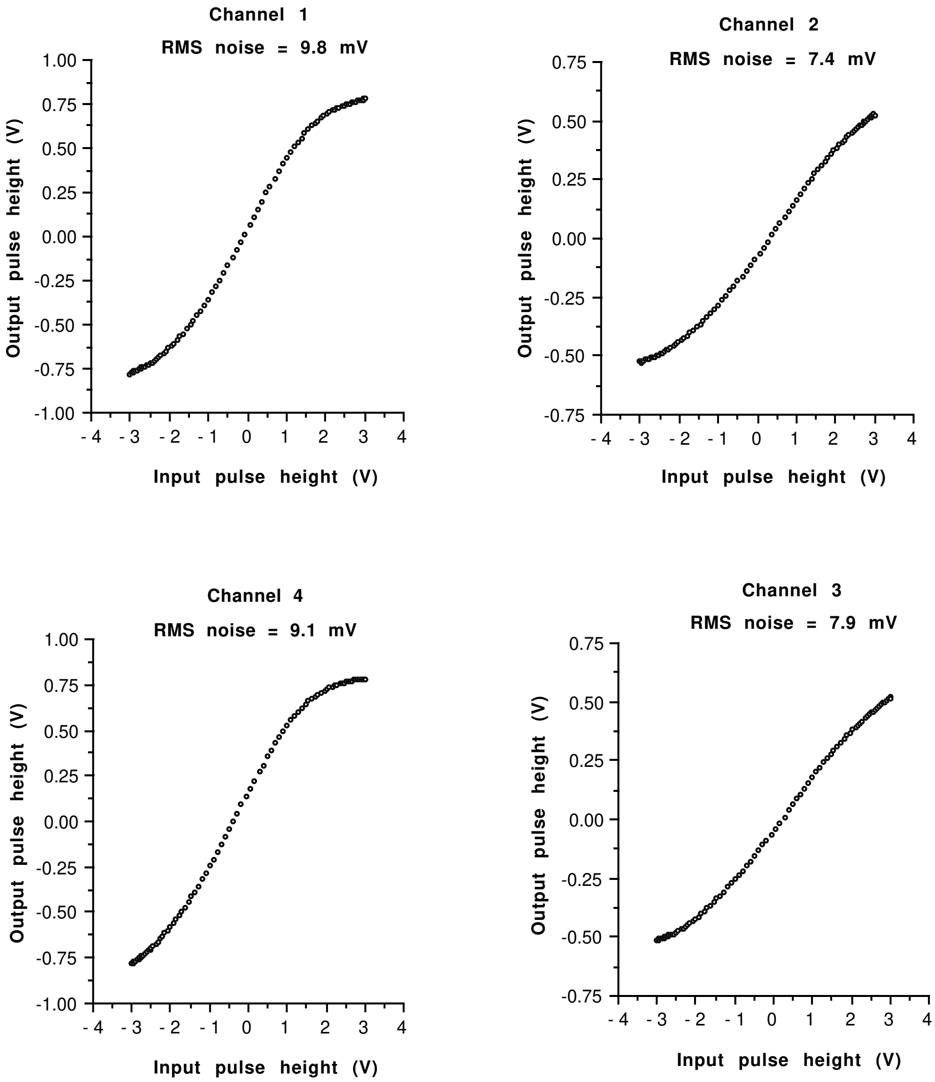
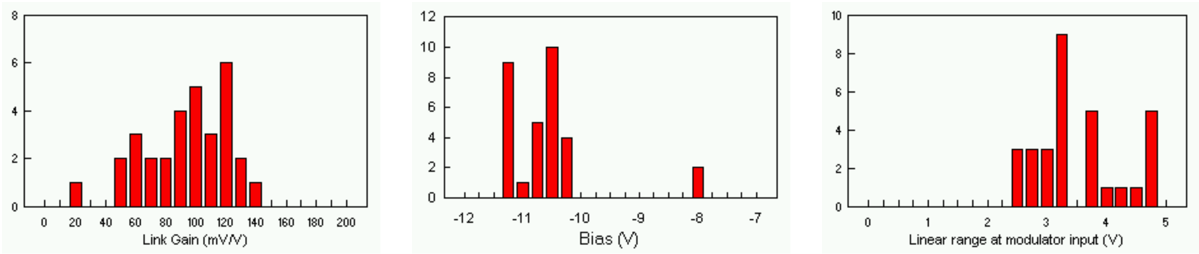


Figure 6.6: Transfer functions measured for a four channel MQW device (sn17)



(a) Gain. (b) Bias voltage. (c) Linear range.

Figure 6.7: Characteristics for 31 MQW modulator channels ($85\mu W$ in channel).

been measured for two modulator devices[a]: the link gain was measured using a 1V peak to peak input signal as described previously; the noise was measured using the RMS function on a digital oscilloscope, which takes the root mean squared value for the points displayed on the screen. The link was set-up with no modulating voltage applied, and the oscilloscope set to an appropriate time-base, $10ms$ on screen so that low frequency noise will be detected. A measurement bandwidth of $100MHz$ was used and an average of 10 RMS values taken.

The received signal as a function of laser current is plotted in figure 6.8. It can be observed that the signal for all four channels of device S/N17 rises linearly with increasing laser drive current, but that the gradient for each of the channels is different. Device S/N02 has a much lower link gain which doesn't vary linearly with laser drive current. This is due to bubbles in the glue used to hold the fibres in place. Altering the laser drive current has the effect of changing its wavelength slightly. This is due to the change in the population of electrons and holes in the active region, which affects the refractive index. The Fabry-Perot response of the glue cavity with wavelength is therefore seen. It is also found that warming the modulator has a similar effect, as the size of the cavity changes with temperature.

The difference in gains of the channels of S/N17 is due to different light powers in each optical channel. This arises in the transceiver due to uneven splitting in the couplers and differing losses in the splices between the components. The power in each channel as a function of laser drive current was measured. Figure 6.9 shows the signal (for 1V peak to peak modulation) to noise ratio versus laser power in each channel for device S/N17. All the channels show the same ratio which is still increasing up to the maximum power from the

[a] Phase 2 modulator packages: serial numbers (S/N) 02 and 17

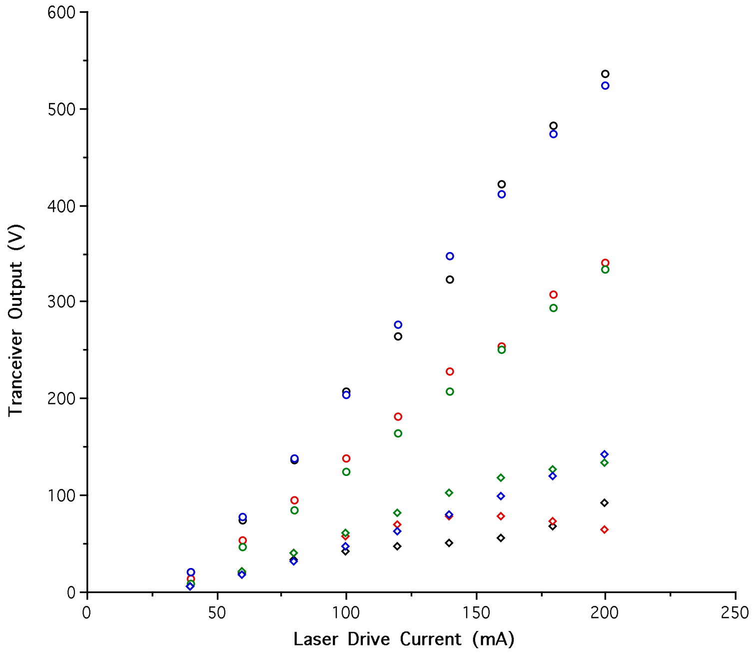


Figure 6.8: Received signal for all channels of devices S/N17 (○) and S/N02 (◊)

laser.

At very low powers the thermal noise, which is constant with laser power, dominates. In this region the signal to noise ratio increases steeply with laser power, since the signal increases linearly with the light intensity. At higher powers the shot noise, which increases as the square root of the light intensity, will dominate over the thermal noise. The signal to noise ratio will still increase, but at a slower rate.

At still higher powers RIN or PAN, depending on the components used in the link, will dominate over thermal and shot noises . Since these two sources of noise increase linearly with increasing light power, the signal to noise ratio of the system should approach an upper limit.

The set-up was changed so that a single optical channel was powered by the laser to allow higher powers to be reached. The signal, noise response and SNR_{1V} for S/N17 channel 4 are shown in figure 6.10. At the extreme limit of the laser power it is observed that both the signal and noise values begin to drop. On an oscilloscope the signal can be seen to distort. There are two possible sources for this effect, either the modulator has reached a point where it can no longer absorb the amount of photons required, or the transimpedence

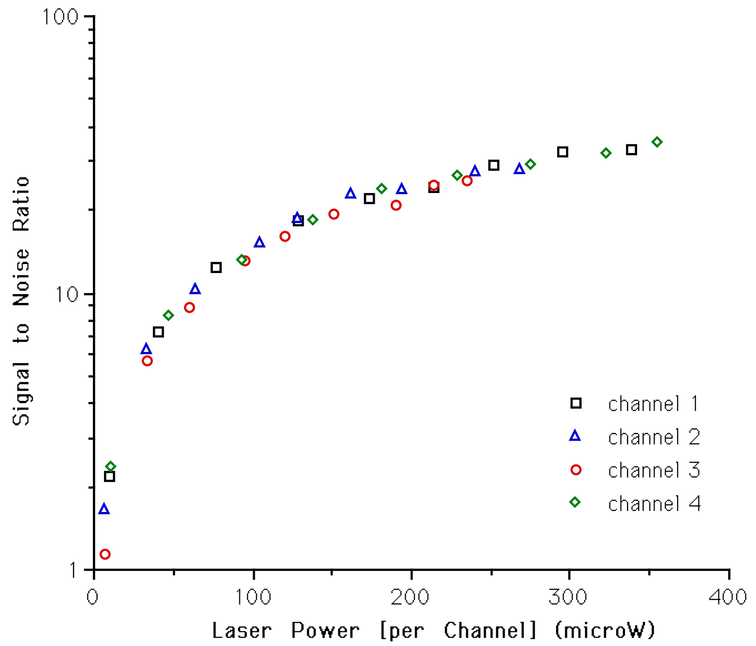


Figure 6.9: Signal to noise ratio for all channels of device S/N17

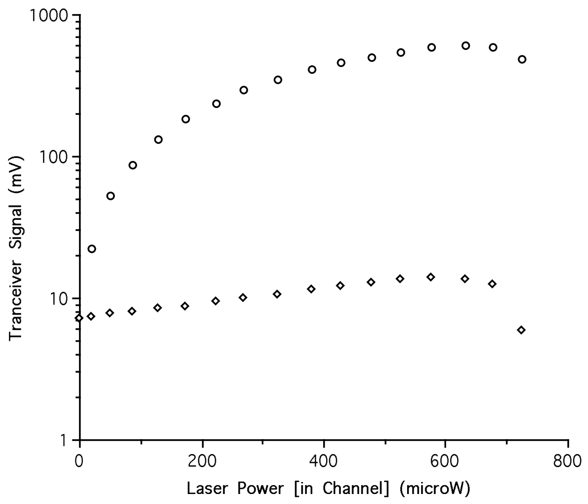
amplifiers are being saturated by the DC light levels. The nature of the distortion, the positive peak is clipped, and the fact that the noise becomes very low indicates that the amplifier is saturating.

The SNR_{1V} continues increasing up to $\sim 650\mu W$ in the channel when the amplifier saturates.

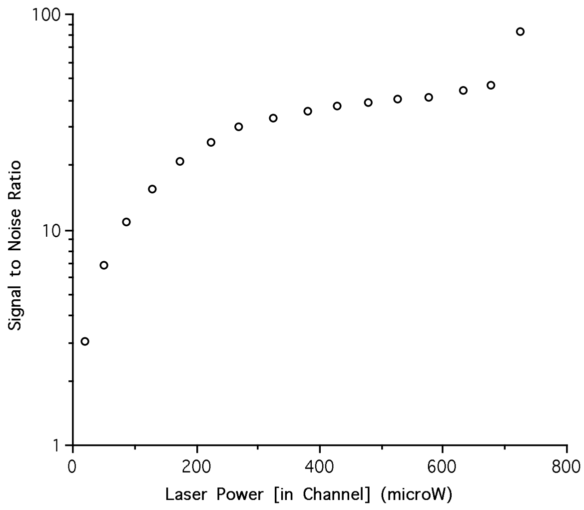
6.2.9 Analogue Performance Conclusions

In figure 6.11 the signal to noise ratio for a $1V$ modulating signal is plotted against the linear range for each channel of the phase 2 devices. The product of these two quantities is the signal to noise ratio. The lines on the plot are contours of constant signal to noise ratio, which increases to the right and to the top of the graph.

A factor 2.5 improvement in the SNR can be expected if the laser power in the channel is increased by a factor 3 to $250\mu W$ (see figure 6.9). The link gain, and hence the SNR, will be improved by using chips of a higher modulation efficiency (a factor 2 is expected to be possible), and decreasing the insertion loss of all optical components in the light path including the modulator (improved assembly process). These improvements mean that



(a) Received signal(∘) and noise (◊).



(b) Signal to noise ratio.

Figure 6.10: Received signal, noise and signal to noise ratio per volt for S/N17 (channel 4).

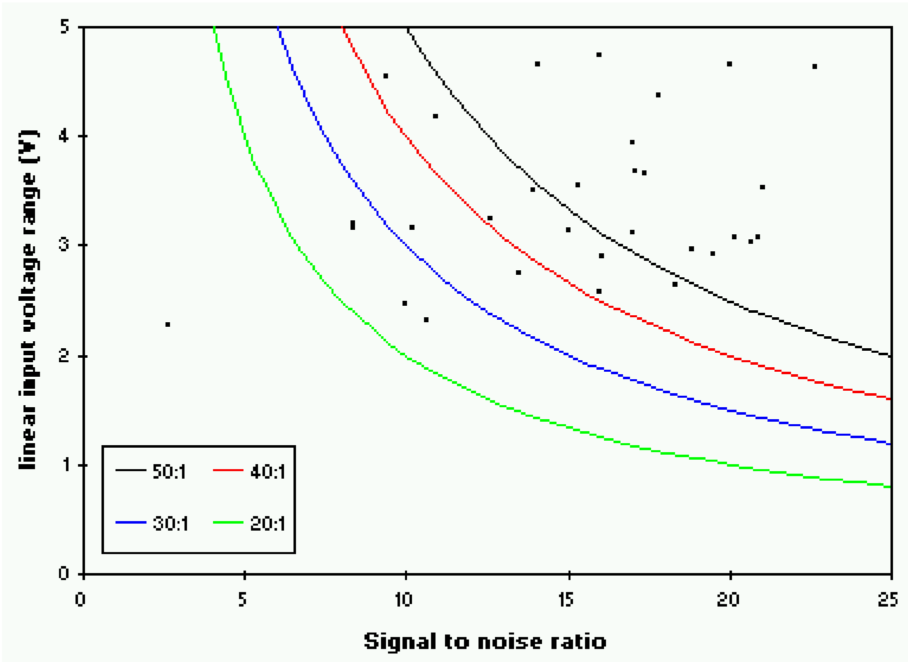


Figure 6.11: Signal to noise ratio$_1V$ and linear range distribution. Lines indicate peak signal to noise ratio contours.

channels measured to have a SNR $> 50 : 1$ would be foreseen to meet the required $> 250 : 1$ ratio for the final readout, as discussed in section 3.4. Of the present devices $\sim 60\%$ meet this requirement. However it should be noted that the noise figure has been measured using an isolated and dithered DFB laser and the signal levels with low insertion loss FC-PC connectors at one of the two breakpoints in the link. The final system would not include an isolator for cost reasons, and would use multi-fibre MT connectors. It would therefore not be expected to perform as well as the test set-up.

6.3 Optical Link Noise Study

The main aim of the practical noise study is to understand the PAN in the reflective link system.

The study involves two aspects; the first is the generation of a computer model of the noise based upon the theoretical study, and the second is the measurement of the noise in an optical test setup to check the model.

An overview of the test setup for studying the noise is followed by a description of the computer model and the methods used to obtain the parameters for the components described in the model. Measurements of noise in the setup and the modeled noise are then compared.

6.3.1 Test Setup

The test setup for studying noise in the optical link is shown in figure 6.12. The setup can be split into several linkable modules, which are all connectorised with FC/PC connectors:

- An isolated 1540nm DFB laser diode forms the light source;
- A PIN photodiode and transimpedance amplifier, of the same design as used in the transceiver, form the receiver;
- An optical power meter can be used to measure the light power at any of the connections in the setup;
- A 2×2 coupler;

- A polarisation controller can be used to change the polarisation of the light;
- Various lengths of fibres terminated with different connectors or components form the reflectors in the setup.

The basic setup used for the noise measurements is shown in the figure and consists of the light source and receiver connected to the input arms, **a** and **b**, of the coupler and the polarisation controller to the output arm **2**. The connectors at the end of the two output arms produce the two reflections required for the generation of PAN. The setup differs from the actual link in that both reflections are at the end of the optical path and can therefore be controlled and measured easily. The other modules can then be connected to the two output arms to alter the reflection magnitudes and relative distance to the detector. The spectral noise power is measured using a spectrum analyser connected to the receiver.

The lengths of all the components has been measured using a ruler and have an error $\pm 0.5cm$ except for the polarisation controller which has an error $\sim \pm 1cm$ due to the accuracy in measuring the diameter of the fibre coils.

6.3.2 Computer Model

The noise in the optical system has been modeled on the spreadsheet package Microsoft Excel. The contributions due to the sources discussed in section 6.2 are modeled. The relevant equations are listed below:

- Receiver noise:

$$< i_r^2 >^{\frac{1}{2}} = < i_r^2 >^{\frac{1}{2}}; \tag{6.10}$$

  which contains the contributions from the receiver thermal noise and photodiode dark and surface current shot noises which are all independent of the photocurrent.

- Signal shot noise:

$$< i_s^2 >^{\frac{1}{2}} = (2 < I_p > e)^{\frac{1}{2}}; \tag{6.11}$$

- RIN:

$$< i_{RIN}^2 >^{\frac{1}{2}} = < I_P > RIN^{\frac{1}{2}}; \tag{6.12}$$

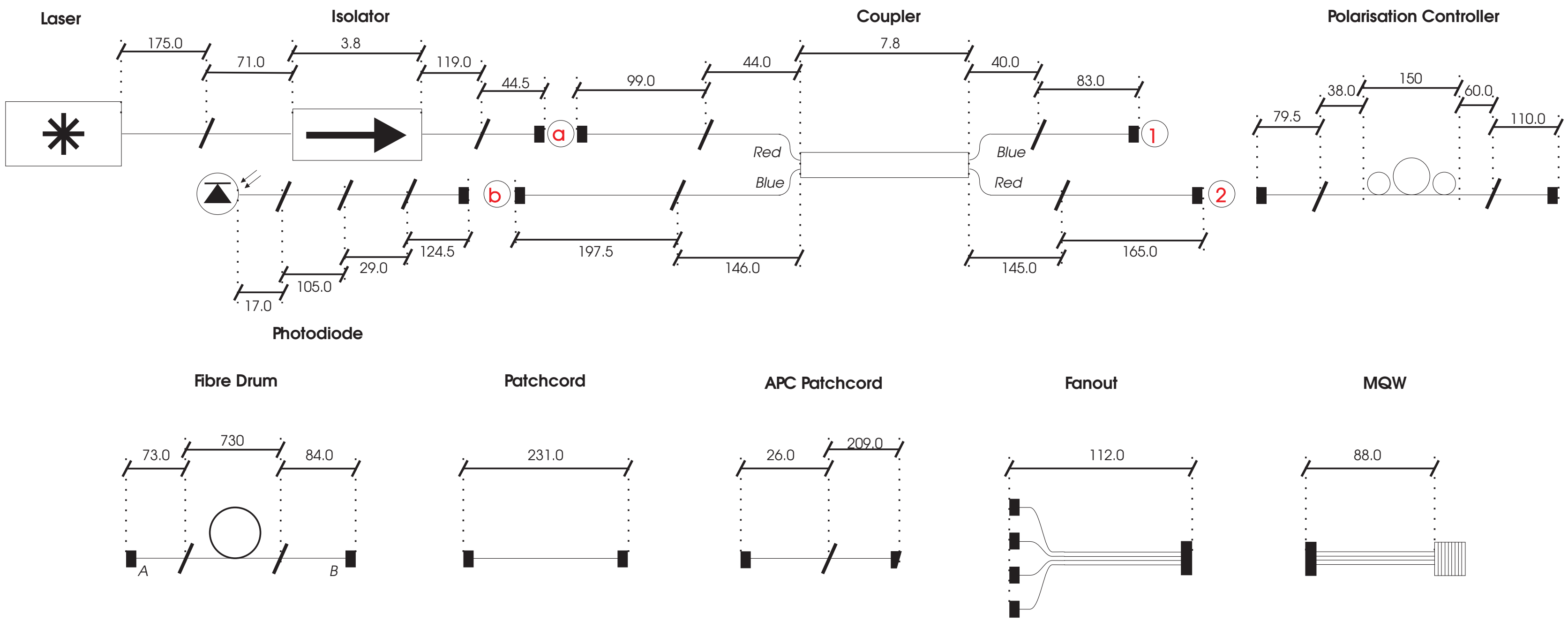


Figure 6.12: Schematic of set-up used for measurement of PAN.

- PAN at quadrature ($cos(2\omega_o \tau) = -1$):

$$P_Q(f) = \frac{4R_1R_2}{\pi(R_1+R_2)^2}\frac{\Delta\nu}{f^2+\Delta\nu^2}\Big(1+e^{-4\pi\Delta\nu\tau} - 2e^{-2\pi\Delta\nu\tau}cos(2\pi f\tau)\Big); \quad (6.13)$$

or at maximum or minimum ($cos(2\omega_o \tau) = +1$):

$$P_M(f) = \frac{4R_1R_2}{\pi(R_1+R_2)^2}\frac{\Delta\nu}{f^2+\Delta\nu^2}\Big(1-e^{-4\pi\Delta\nu\tau} - 2e^{-2\pi\Delta\nu\tau}\frac{\Delta\nu}{f}sin(2\pi f\tau)\Big); \quad (6.14)$$

The average is taken.

The periodic term in these equations leads to characteristic peaks and troughs in the noise spectrum. The separation of the peaks, f_{sep}, is given by,

$$f_{sep} = \frac{c}{2\Delta l n}; \quad (6.15)$$

where, Δl, is the separation of the two reflection points,
c is the speed of light,
and n is the refractive index of the fibre.

The RMS current noise due to PAN is,

$$< i_{PAN}^2 >^{\frac{1}{2}} = < I_P > P(f)^{\frac{1}{2}}. \quad (6.16)$$

The model calculates the noise power at 100 points between user defined start and stop frequencies. The noise power due to each of the noise sources is calculated by,

$$Power_{noise} = 10LOG((< i_{noise}^2 >^{\frac{1}{2}} G\Delta f_{res}^{\frac{1}{2}})^2/\Omega) + 30 \quad (6.17)$$

where, Δf_{res} is the resolution bandwidth of the spectrum analyser,
and Ω is the input impedance of the spectrum analyser.

The +30 (factor 1000) is required to change to the dBm used when taking measurements.

6.3.3 Parameter Measurements

6.3.3.1 Laser Parameters

The laser RIN value and wavelength used in the model are those given by the manufacturer for the DFB laser used. They are $-145dB/Hz$ and $1550nm$ respectively.

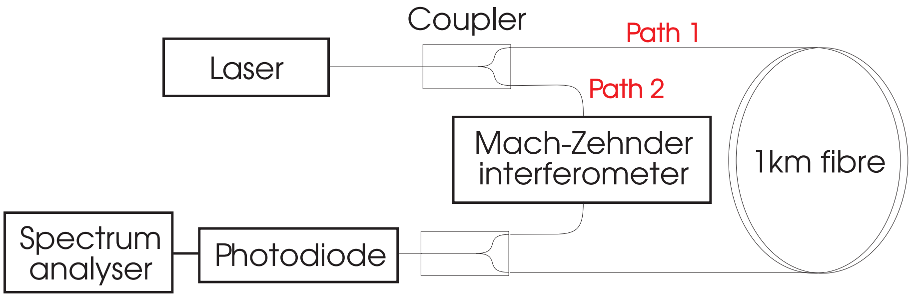


Figure 6.13: Schematic of the heterodyne system used for the measurement of laser linewidth.

The laser linewidth has been measured using a heterodyne scheme [49] as shown in figure 6.13.

A local oscillator at the correct frequency is obtained by using part of the laser output. This is achieved by dividing the laser output between two paths. The light through path 1 is delayed by time τ_d and is regarded as the local oscillator output. The signal through path 2 is frequency shifted using a Mach-Zehnder interferometer driven with a $50MHz$ saw tooth modulating input. The frequency shift is much larger than the spectral spread to be measured. The outputs of the two paths are mixed at a photodiode.

When τ_d is much longer than the coherence time of the laser, the throughputs of the two paths are not correlated. The noise spectrum of the mixer output exhibits a $3dB$ spectral spread somewhat wider than the original spectrum, since the signal and local oscillator output are equally as noisy. In semiconductor lasers, where the spectrum is close to Lorentzian, the spectral spread at the mixer will be twice that of the original.

The linewidth of the DFB laser was measured. The laser package contains a cooler and thermistor to allow the temperature to be controlled. Measurements were made for different thermistor resistance settings and therefore laser temperatures. The results are shown in figure 6.14. Two or three readings were taken at each thermistor resistance value and an average taken. The linewidth of the laser falls slowly as the thermistor resistance is increased corresponding to a decrease in temperature. The values measured agree well with the manufacturers specification ($\Delta\nu_{typ} = 2MHz,\ \Delta\nu_{max} = 10MHz$).

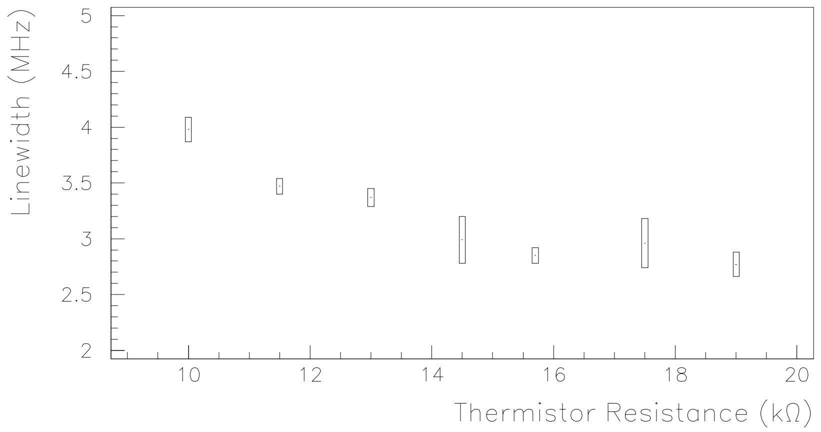


Figure 6.14: Measured linewidth of DFB laser.

6.3.3.2 Receiver Parameters

The photodiode responsivity is $0.8A/W$ at $1550nm$ as measured by the manufacturer.

The amplifier gain has been measured over the range $0 - 250MHz$. The technique used to measure the gain is to amplitude modulate the laser and thus the light incident on the amplifier photodiode. The amplitude of the modulated signal output from the amplifier, v_{out}, is then measured. The optical power modulation of the laser, P_{in}, is then measured, at low frequency, using a power meter. The transimpedance gain is given by,

$$G = \frac{v_{out}}{P_{in}\Re}. \tag{6.18}$$

The amplifier noise output power has been measured using the spectrum analyser when there is no light incident on the photodiode. The amplifier thermal noise and photodiode dark and surface shot noises will therefore be the only noise sources. The RMS noise current at the input to the amplifier is then adjusted in the model to match that measured.

6.3.3.3 Optical Parameters

The measurement of the optical power, with a power meter, was used to determine the magnitude of the reflections in the system. Measurements were made at the two output arms and the receiver arm of the coupler. The use of index matching gel to remove the reflection due to a particular connector allows the magnitude of the other reflection to be

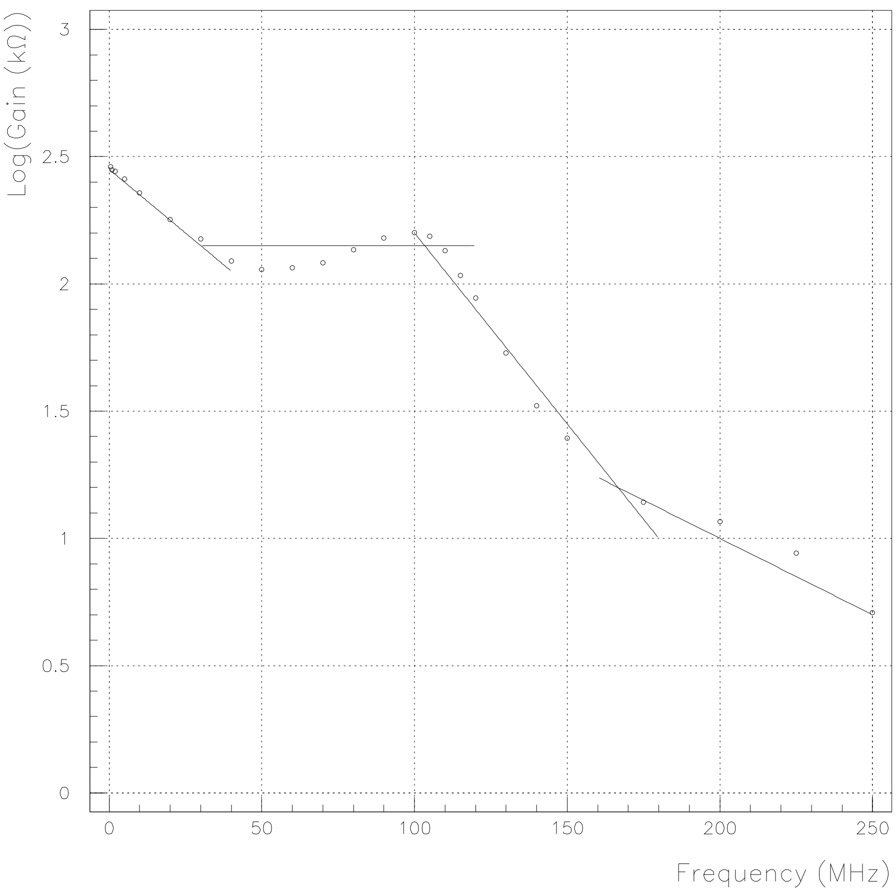


Figure 6.15: The transimpedance amplifier gain. The lines show the fit used in the noise model.

determined. The losses in the setup are parameterised in terms of the coupler loss ($-3dB$ for the 50/50 splitter used), reflection loss of the two connectors and a loss between the coupler and each of these connectors due to other components, for example, the polarisation controller or another mated connector pair. The adjustment of the connector reflections, losses between the coupler and connectors and the laser power launched into the system allow any system to be modeled.

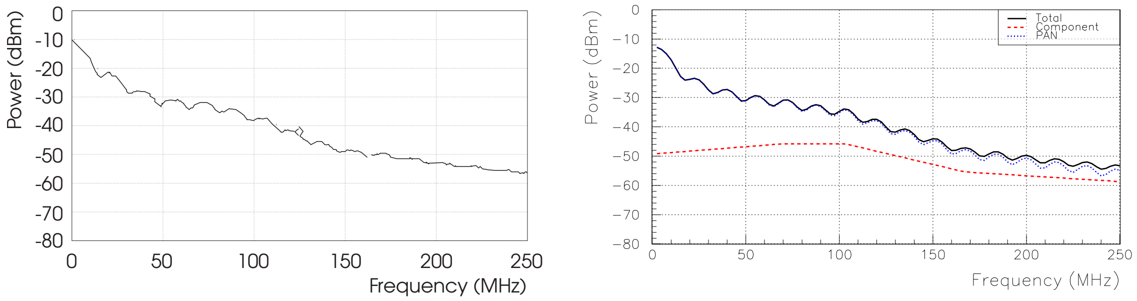
6.3.4 Noise Measurements

Comparative pairs of plots showing the noise measured in the test setup and that modeled follow. There are five parameters that govern the PAN generated in the system: the optical power in the system, the linewidth of the laser, the delay of the reflected signals relative to one another, and the magnitude of the two reflections. The first four sets of measurements are checks of the model. They look at simple reflections from flat polished connectors. The laser power and linewidth, and the separation of the connectors are altered. The noise in each case was measured for the polarisation controller setting giving the most noise. This corresponds to the two reflected light signals having the same polarisation at the receiver.

Measurements of more realistic systems are then made incorporating angle polished connectors and the MQW modulator to look at the effect of different magnitudes of reflections in the system.

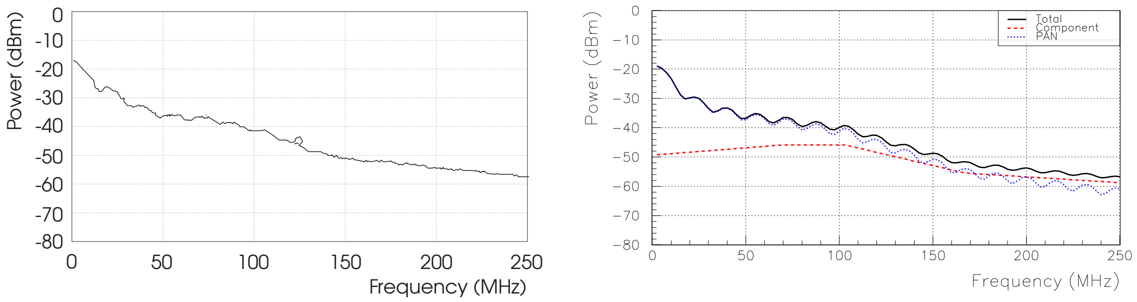
The first pair of measurements compare the noise observed at two optical powers. For these measurements the setup is the basic one described in section 6.3.1. The relative distance of $6.25m$ between the two connectors gives rise to characteristic peaks with spacing $16MHz$. The modeled noise shows the component noise due to just the amplifier and the PAN separately as well as the sum of all noise sources.

Figure 6.16 shows the noise for an optical power of $\sim 0.64mW$ launched into the system. The reflections from each of the connectors is $\sim -16dB$ which is slightly lower than the $-14dB$ predicted by theory. A loss of $3dB$ is experienced with each pass though the coupler as the light is split between the two possible output channels. A further loss of $\sim 4dB$ occurs in arm **2** with each pass through the polarisation controller. The modeled and measured noise agree well. Figure 6.17 shows the noise for half this laser power launched into the system. The optical power at the receiver due to each reflection undergoes a simple halving resulting in a factor 4, $6dB$, decrease in the PAN. The measured PAN agrees well with that modeled.



(a) Measured. (b) Modeled.

Figure 6.16: Noise plot 1: High laser power; 6.25m cavity formed by two FC/PC connector reflections.

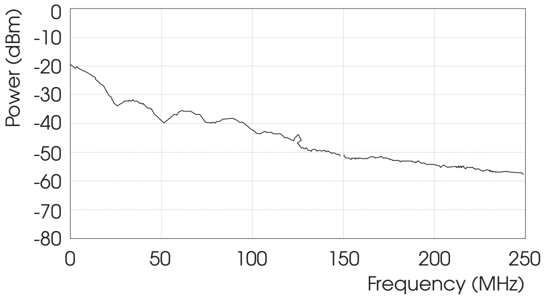


(a) Measured. (b) Modeled.

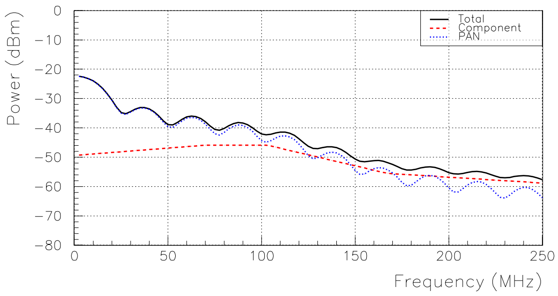
Figure 6.17: Noise plot 2: Low laser power; 6.25m cavity formed by two FC/PC connector reflections.

A $2.31m$ fibre patchcord is attached to arm **1** of the coupler changing the relative reflector separation to $3.94m$. Figure 6.18 exhibits the characteristic $25MHz$ peak spacing associated with this cavity size.

The linewidth of the laser has been altered by changing the laser temperature. In figure 6.19 the linewidth is $\sim 3.9MHz$ compared to $\sim 2.85MHz$ in the previous measurements. The effect is to slightly decrease the depth of the troughs as the coherence of the two reflected signals is reduced. A noticeable effect would be expected when the coherence length of the laser approaches the round-trip separation of the reflection points.

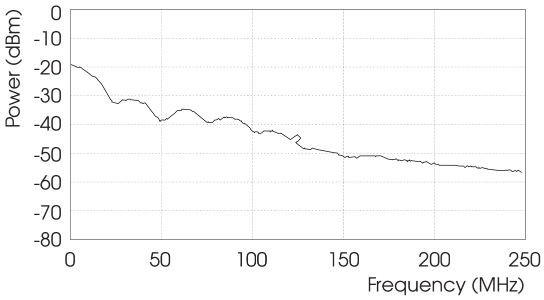


(a) Measured.

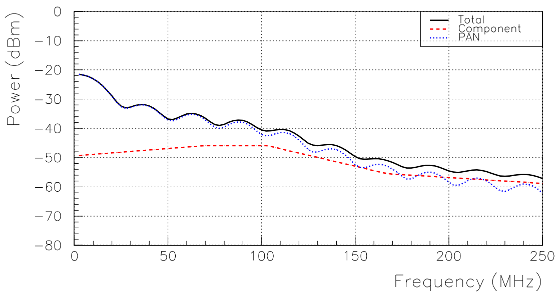


(b) Modeled.

Figure 6.18: Noise plot 3: Low laser power; 3.94m cavity formed by two FC/PC connector reflections.



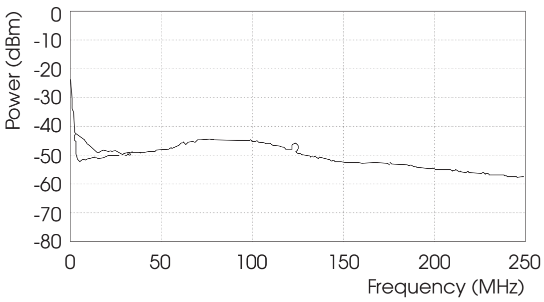
(a) Measured.



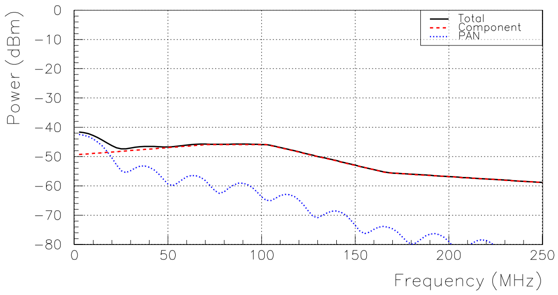
(b) Modeled.

Figure 6.19: Noise plot 4: Low laser power; 3.94m cavity formed by two FC/PC connector reflections; increased laser linewidth.

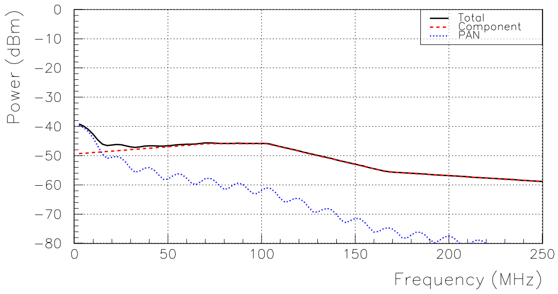
For these simple measurements the model predicts the measured system noise well. The effect of different reflectors is now studied. The laser linewidth has been returned to $\sim 2.85MHz$. The first change made to the system to to connect an angle polished connector to arm **1** of the coupler. The noise modeled in figure 6.20 shows results for two different reflector separations. The first, with a separation of $3.89m$, corresponds to the separation of the polished connector terminating arm **2** of the coupler after the polarisation controller and the angle polished connector terminating arm **1** of the coupler. The second plot, with a reflector separation of $6.25m$, corresponds to the polished connector on arm **2** and the polished connector pair on arm **1** that is used to connect the angled polished connector patchcord. The position of the first trough at $\sim 16MHz$ in the measured noise spectrum indicates that the reflection causing the noise is from the mated polished connector pair and not the angle polished connector. The angle polished connector was index matched and no change in the measured spectrum was observed confirming this hypothesis.



(a) Measured.



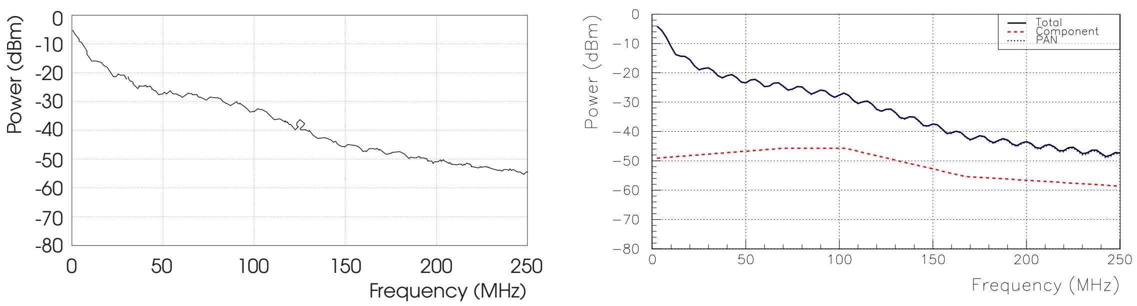
(b) Modeled (FC/APC).



(c) Modeled (PC-PC connection).

Figure 6.20: Noise plot 5: High laser power; cavity formed by i) FC/PC and FC/APC connectors or ii) FC/PC connector and PC-PC connection.

The MQW modulator was then attached to arm **2** after the polarisation controller using the fanout. The design of the MQW results in a large reflectance, approximately four times that of an open connector. Arm **1** of the coupler was terminated in a flat polished connector. The model matches the measured noise shown in figure 6.21 quite well. A discrepancy is observed at higher frequencies. This is thought to be due to the error in the measurement of the amplifier gain.



(a) Measured. (b) Modeled.

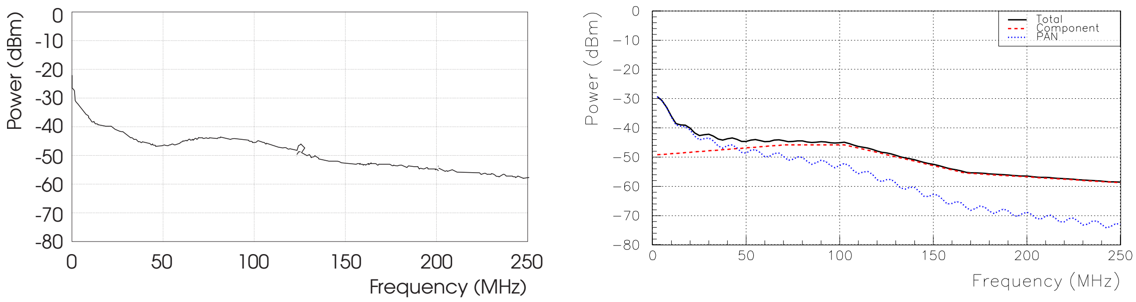
Figure 6.21: Noise plot 6: High laser power; 16.49m cavity formed by MQW modulator and FC/PC connector.

The flat polished connector was then replaced with the angle polished connector. The main reflection in the second arm was again found to be due to the mated connectors connecting the angle polished connector pigtail. This test setup now features a large reflection from a MQW and a small reflection from the mated connectors similar to those found in the real data link. It can be observed, in figure 6.22, that the majority of the noise is situated at low frequencies of the order of the data transmission rate of $40MHz$.

6.3.5 Noise Study Conclusions

The noise model highlights the problems of PAN in the analogue link. The noise is generated at frequencies comparable to the data transmission rate and therefore has a high impact on the link performance. This also prevents the removal of the noise via filtering.

The magnitude of the noise is dependent upon the size of the reflections in the system. To reduce the noise due to PAN to the level of the amplifier noise term the connector reflections must be reduced to $\sim -60dB$ from $\sim -40dB$. This may be possible via selection of the



(a) Measured.

(b) Modeled.

Figure 6.22: Noise plot 7: High laser power; 16.49m cavity formed by MQW and PC-PC connection.

connectors, but would involve additional labour and cost. The alternative is to reduce the reflection from the MQW. A reduction can be achieved by designing the MQW to operate in a normally off mode. At present the MQW design is optimised to work in a normally on mode. This means that the largest change in reflectance is around the point of greatest reflectance. However, optimisation around a lower reflectance point is possible. A $3dB$ drop in the reflectance could be achieved giving a similar improvement in the noise.

An alternative way to reduce the PAN is to reduce the mixing of the two reflected signals. This can be achieved by ensuring the two reflections are cross polarised at the receiver. This would require the use of polarisation maintaining fibre and a $\lambda/2$-plate at the MQW. The cost of polarisation maintaining fibre prevents the use of this scheme.

6.4 Irradiation of MQW modulators

The tolerance of the phase 2 type modulators to neutron and proton irradiation has been measured [50]. The standard test setup described in section 6.2.6 has been used for the tests. The modulators were located in the radiation area and the transceiver, connected by the ribbon fibre, in a barracks area where measurements were taken. During both irradiations two modulator packages were tested and two channels from each monitored. Channels 1 and 4 of modulator S/N17 and channels 2 and 3 of modulator S/N12 were used. Three properties were measured to assess the performance of the link:

- For the readout of data the important property is the link gain. During ATLAS readout operation a $40MHz$ data signal will provide the modulating voltage. However, for these tests only DC biases were applied to the MQW modulators. This avoided the problems associated with sending AC electrical signals over long distances to the irradiation area (where electrical equipment could not be used). A time varying light signal is required at the transceiver photodiode because the amplifiers are AC coupled. To achieve this the laser source was modulated.

  The amplitude of the time varying photodiode signal was measured as the DC bias voltage to the MQW modulator was varied in $1V$ steps from -7 to $-15V$. A third order polynomial is fitted to the data and the gradient at the inflection point found. The gradient is defined as the link response. The time varying signal at the photodiode is proportional to the modulation amplitude of the laser source and the reflectance of the MQW modulator. The link response is therefore proportional to the link gain. The factor of proportionality could not be measured as the signal before reflection from the modulator saturated the amplifier;

- The photocurrent is the current produced due to light incident on the MQW modulator, and is a good indicator of the opto-absorption of the device. It is determined by measuring the voltage across a known resistance, in series with the MQW modulator, at a particular incident light power. By varying the reverse bias voltage from -7 to $-15V$, the photocurrent gradient can be determined in the same way as for the link response.

  The photocurrent gradient is proportional to the link gain, because the opto-absorption determines the amplitude of the signal reflected back to the transceiver photodiode. However, the photocurrent should be a better guide to changes in MQW chip behaviour, mainly because it does not include uncertainties due to the receiver electronics;

- An important property of the MQW when used as a photodiode is its dark current, which determines the noise of the receiver electronics. The dark current is the current when there is no light incident on the MQW. It was determined by measuring the voltage across a known resistance, with the MQW biased at $-10V$ and the laser turned off.

6.4.1 Neutron irradiation

The modulators were irradiated in the Birmingham dynamitron (see Appendix A). The devices were positioned side by side $\sim 2cm$ below the beryllium target. Measurements were taken starting a few days before the irradiation to check the stability of the setup, during the irradiation, and for a few weeks afterwards to look for any recovery of the devices. The link response, photocurrent gradient, dark current, and received fluences for the measured channels are shown in figure 6.23. The fluence measurement was originally taken using one indium foil placed over both devices. A later test for homogeneity, using separate smaller foils, indicated that a different fluence was received by each device and that the original indium foil was in fact an alloy and not pure indium. This resulted in a high error on the fluence measurements.

Decreases in link response of $\sim 20\%$ can be seen for channels 3 and 4, corresponding to neutron fluences of $\sim 3.3 \times 10^{14}$ and $6.5 \times 10^{14} cm^{-2}$ respectively. Channel 1, however, exhibits no obvious radiation induced degradation (within the experimental uncertainty of $\sim 10\%$). Results are not shown for channel 2 because it was disconnected a number of times to allow monitoring of the light power in the system. This remained constant to within the measurement capabilities of the optical power meter ($0.1dBm$).

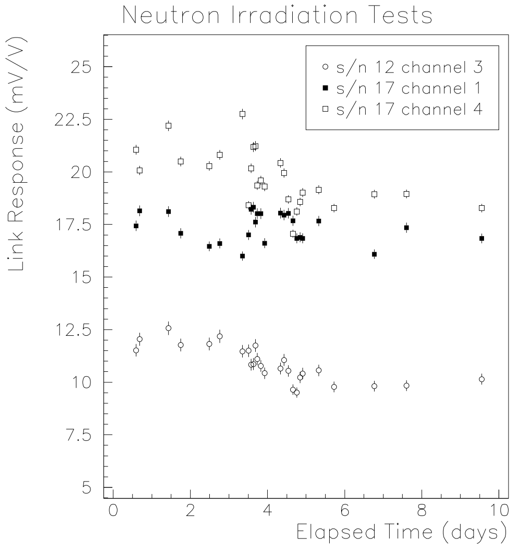
Clear degradations in the photocurrent gradient of the order of 10% can be seen for all three channels. The negative values correspond to an increase in photocurrent as the bias becomes more negative. Again only three channels are plotted for the reasons stated previously.

The dark current for all four channels increases dramatically during the irradiation period (from several nA to $\sim 1\mu A$). The two channels of MQW device S/N17 exhibit a larger increase in dark current corresponding to the higher neutron fluence received.

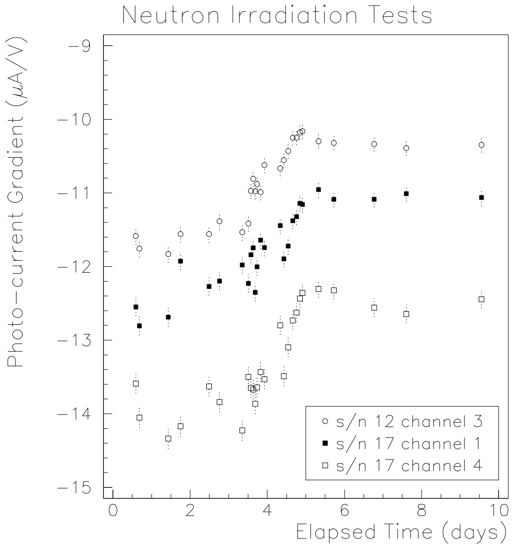
6.4.2 Proton irradiation

The experimental setup used for the neutron irradiation study was transferred to the CERN PS (see Appendix A). The same four modulator channels were monitored. The results are shown in figure 6.24.

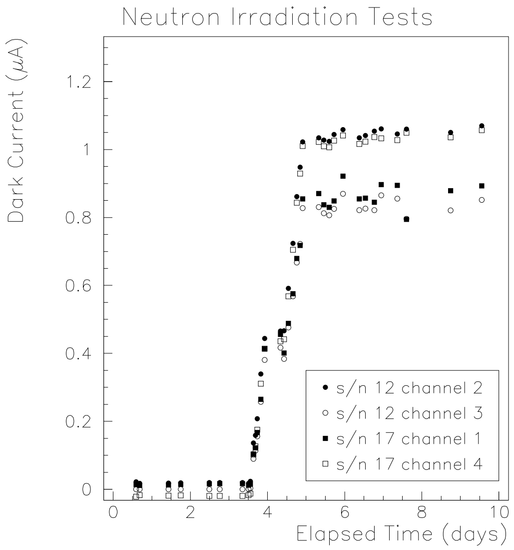
The link response measurements taken prior to the irradiation show a spread in their



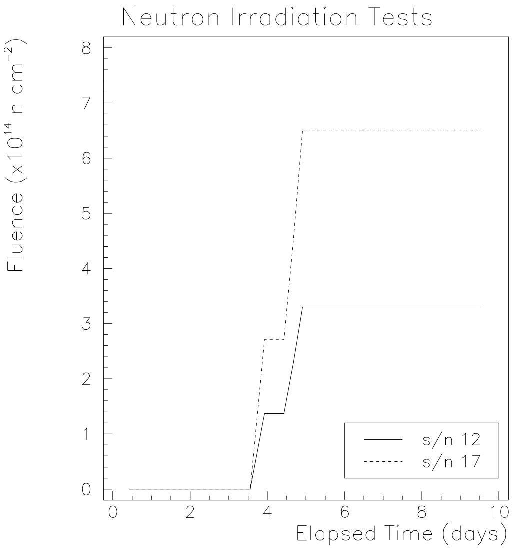
(a) Link response at inflection point.



(b) Photo-current gradient at inflection point.

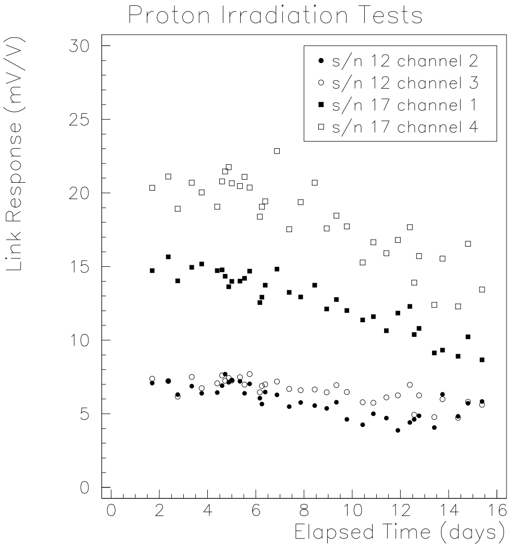


(c) Dark current at -10V.

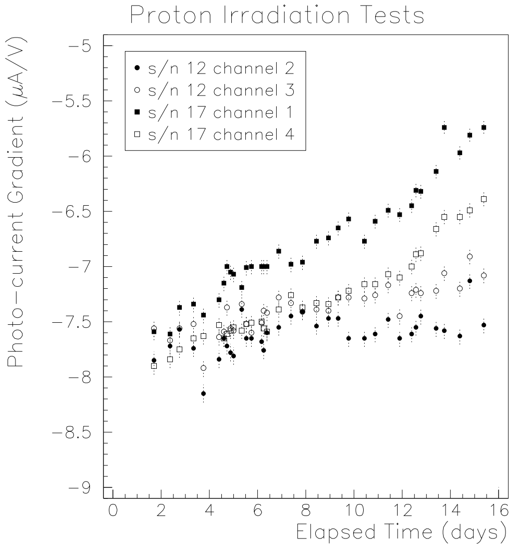


(d) Neutron fluences for the two MQW packages; the measurement uncertainty is $\sim 50\%$.

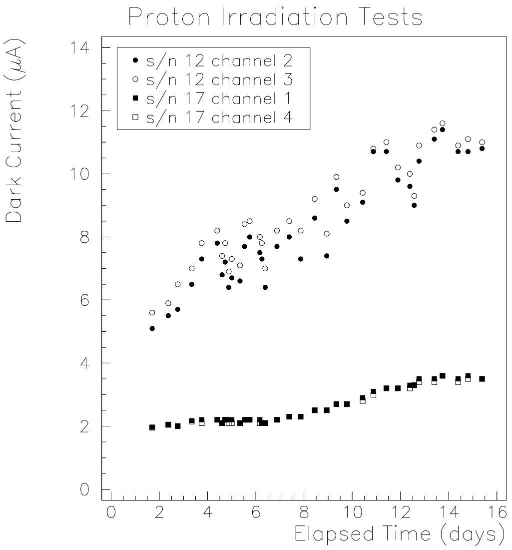
Figure 6.23: Neutron irradiation test results of MQW modulators.



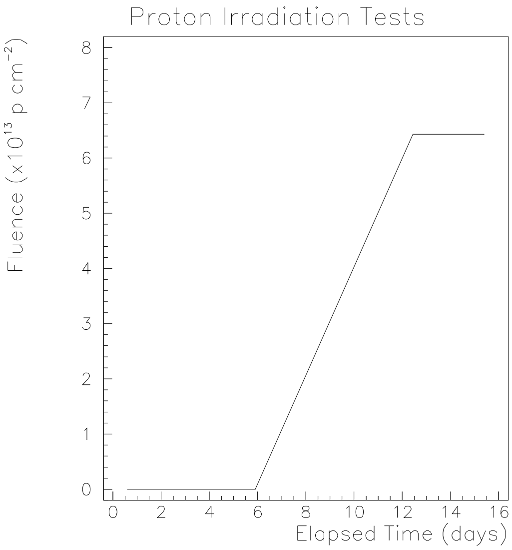
(a) Link response at inflection point.



(b) Photo-current gradient at inflection point.



(c) Dark current at -10V.



(d) Proton fluence for the two MQW packages; the measurement uncertainty is $\sim 20\%$.

Figure 6.24: Proton irradiation test results of MQW modulators.

values of typically $\sim 5\%$. This is an estimate of the stability of the link, for the setup used in these tests. Degradations are observed on all channels as a result of irradiating with proton fluences $\sim 6.4 \times 10^{13} cm^{-2}$. The size of the degradation in the worst case is $\sim 30\%$.

The photocurrent gradients of modulator S/N17 exhibit large degradations that are continuous throughout the test period, although there is no obvious correlation with the proton irradiation period. In the worst case of channel 1 the degradation is $\sim 20\%$. For MQW device S/N12, it is not clear that any degradation has occurred within the uncertainties associated with the stability of the link.

The dark currents for the two modulators also exhibit different responses. For device S/N17 the dark currents are steady during the pre-irradiation period (but at a higher level than at the end of the Dynamitron (neutron) tests due to the higher ambient temperature ($27 \pm 1^{o}C$ cf. $15 \pm 2^{o}C$)). An increase in dark currents correlated to the proton irradiation period is observed, from $2\mu A$ to $3\mu A$. The two channels of modulator S/N12, however, show dramatic increases in dark currents compared to the values measured at the end of the Dynamitron tests. Throughout the proton tests they continued to increase steadily. When measuring the dark currents a time dependence was observed after switching off the light incident on the MQW. This is the reason for the larger statistical spread in the measured values. The cause is thought to be due to passivation problems at the modulator mesa [51].

6.4.3 Conclusions

MQW-modulators have been irradiated with neutrons from the Birmingham Dynamitron to fluences of $\sim 5 \times 10^{14} cm^{-2}$, followed by $24GeV$ protons to fluences of $\sim 6 \times 10^{13} cm^{-2}$. Deteriorations of the link gains, due to the neutron irradiation, are less than $\sim 20\%$ for all the channels measured. The corresponding deteriorations in the photocurrent gradients are less than $\sim 10\%$. For the proton irradiation, deteriorations of up to $\sim 30\%$ and $\sim 20\%$ have been measured in the link gain and photocurrent gradient respectively. The lower proton fluence, compared to the neutron fluence, suggests that $24GeV$ protons are more damaging than $1MeV$-equivalent neutrons. This agrees with the NIEL hypothesis for devices fabricated from GaAs. The exact ratio of the damage cannot be calculated due to the large uncertainties in the measurements.

Dramatic increases in dark currents are observed, on all the MQW channels, during both irradiation tests. All the channels exhibited stable behaviour for the neutron tests, but during the interval before the proton tests the channels on one of the MQWs showed a large increase in dark currents and became less stable. This is believed to be due to problems with the passivation of the device.

The results obtained for the MQW in these irradiation tests suggest that a performance margin of up to $\sim 50\%$ may be required at the start of LHC operation, to ensure acceptable data readout throughout the LHC lifetime. Test-beam results ([52] and section 6.5) indicate that the MQW based link is satisfactory for transmitting analogue data before irradiation damage. One can conclude that even after irradiation damage it would be acceptable for the transmission of digital data, but, would require considerable care when used for analogue data transmission.

6.5 ATLAS SCT Beamtest

6.5.1 Setup

The optical link has been integrated into a complete readout chain in the August 1995 ATLAS SCT H8 beamtest at CERN [52, 53]. A silicon detector module in a 'keystone' geometry, designed for use in the forward region of the detector, with one APV5 front-end chip [54] was kindly made available for use with the readout by the Liverpool particle physics group.

The module included driver circuitry which was modified, but not optimised for use with the MQW readout, due to the late and very limited arrival of the APV5 chips. The driver/amplifier circuit used for the MQW is shown in figure 6.25. The value of the resistor R_1 alters the DC offset of the signal or bias voltage of the MQW. The value of the resistor R_2 alters the gain of the second amplifier stage, this changes the modulating voltage applied to the MQW for a one MIP signal from the detector, but also changes the bias voltage. The value of the capacitor, C_1, must be chosen so that the time constant R_2C_1 is long compared to the clock speed. The APV5 chip suffered from several performance shortfalls from the expected ATLAS operation, one of which was the large spread in the pedestal values (equivalent to $\sim 4MIPs$ for the 128 channels) resulting in the requirement for a much larger input range to be covered by the readout.

Since only one APV5 chip was available for the module it meant that only one quarter

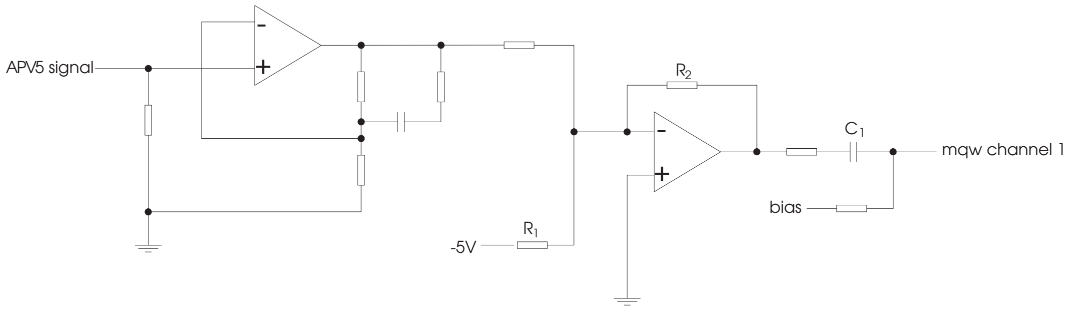


Figure 6.25: Driver circuit for MQW readout used in H8 test-beam.

of one side of the module was read out, and that only one of the the four optical channels was used.

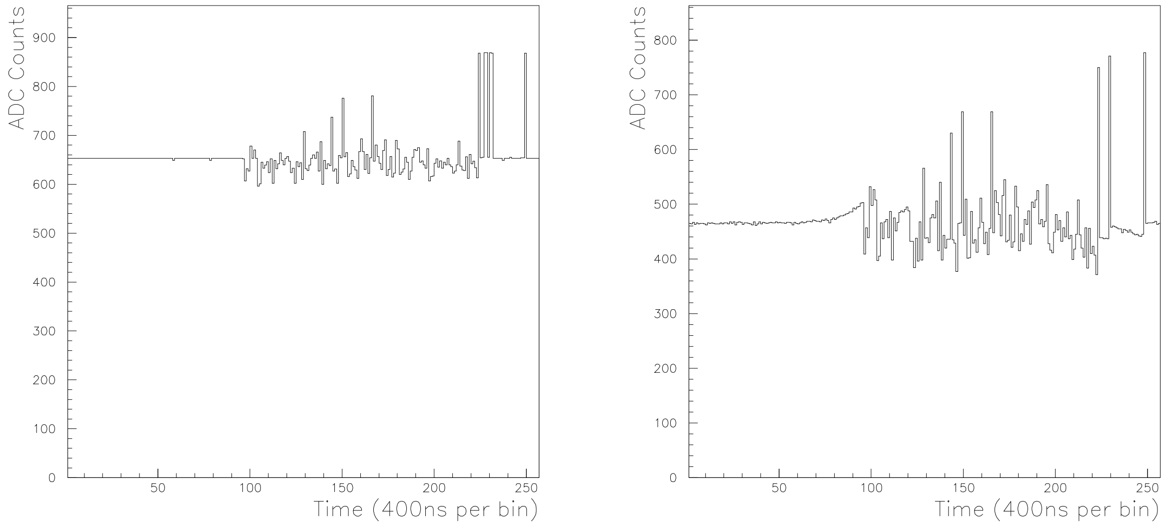
Trigger and clock signals for the APV5 chips were sent from a special VME module in the barracks. The trigger signal was generated from the logical AND of two scintillators placed in the beam before and after the detector.

In the barracks, the signal from the transceiver was digitised using a SIROCCO analogue to digital converter (ADC) and then saved to Exabyte tape.

Data had previously been read out from the same module using a LED type optical link. This link was not suitable for LHC conditions in terms of radiation hardness and power consumption, but had a very large peak signal to noise ratio so that no degradation in the signal due to the link would be expected. The results have been compared.

6.5.2 Analysis

The data on the exabyte tapes were analysed off-line using a Fortran program written by the Off-line Group at the test-beam. This was modified for use on MQW read out data. Examples of raw data from the APV5 chip read out using LED's and MQW's are shown in figure 6.26. The data are transferred from the SIROCCO to the data acquisition system (DAQ) in reverse order, so time runs from right to left. The major features on both plots are a single header bit to show the start of the data, followed by three 8 bit digital addresses which refer to which of the 160 bin long pipeline the data were stored in. The pipeline is designed for use in the LHC. It allows time for a decision to be made, using data from the other parts of the detector (ATLAS or CMS). If the event is interesting the inner tracker data are read out, otherwise the data are just discarded. In the peak mode of the APV5 chip used at the test beam only one address is output. This is followed by the 128 channels of analogue data multiplexed in a non-sequential order.



(a) LED readout.

(b) MQW readout.

Figure 6.26: Test-beam raw data for LED and MQW readouts.

The basic program performs the following major functions:

- unscrambles the order of the analogue data;
- decodes the digital time bin address;
- performs a pedestal subtraction on the analogue data;
- performs a common mode subtraction on the analogue data;
- calculates the RMS noise for each detector channel;
- finds hits in the detector;
- produces histograms.

It was found that the APV5 chips suffered from significant variations in pedestal values between time bins used to store the data as well as between the channels in one time bin.

For this reason the pedestal and RMS noise values were calculated for each channel and each time bin, resulting in two 128 by 160 arrays.

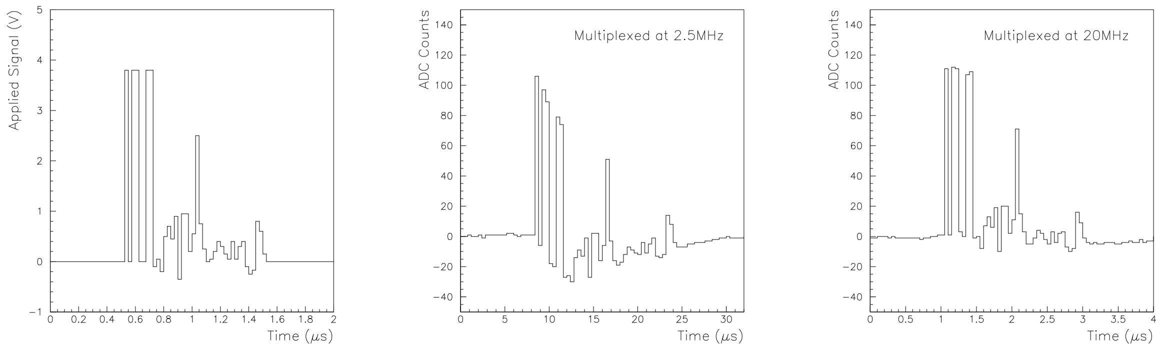
6.5.2.1 AC coupling

The MQW link is AC coupled at both the driver and receiver ends. The driver end is designed in this way to allow the required bias voltage for operation in the devices linear region to be applied. Designing the receiver end in this way is useful due to the nature of the signal that is to be amplified, a small amplitude modulated light signal on top of a large CW background. In theory AC coupled links are perfectly acceptable, but in the beam-test problems were encountered due to another of the APV5 design faults. It turned out only to operate when clocked at $2.5MHz$ rather than the $40MHz$ for which the link circuitry was designed. It proved possible to alter the drive circuitry to operate at this frequency, but the finely tuned design of the transimpedance amplifiers meant that the receiver time constant could not be changed in the short time available.

A test-bench setup at Birmingham, running under LabVIEW, uses a sequencer and digital to analogue converter (DAC) to produce pulse trains for input into the link. This setup can be clocked at 40, 20, 10, 5 or $2.5MHz$. The transceiver output is fed through an ADC back to the computer. A simulated signal similar to that from the APV5 chip, shown in figure 6.27, is used as the input to the link. The link output for a clock rate of $2.5MHz$ displays the same characteristic AC undershoot as seen in the MQW test-beam data. However, for a clock rate in excess of $20MHz$ the undershoot disappears. This shows that the AC coupling will not be a problem in the final system.

The AC undershoot caused several problems in the data analysis:

- The values for the first ~ 16 analogue channels read out were severely dependent upon the specific digital address preceding them in the data stream. This was corrected for since the pedestal values were calculated for each time bin anyway;
- The same undershoot occurs in the channel read out after a channel that has been hit. This will increase the noise measured somewhat;
- The actual value read by the ADC is also affected by the AC coupling. The ADC was set to sample each data value after about $\frac{3}{4}$ of its $400ns$ period. The transimpedance



(a) Simulated APV5 chip signal.

(b) Output for $2.5MHz$ clock.

(c) Output for $20MHz$ clock.

Figure 6.27: Test-bench measurement of AC coupled system using a simulated APV5 signal.

amplifiers had a time constant of $2\mu s$ and so the sampled value will be 7% below the actual value. However, the noise is the same at all points of the signal, resulting in a degraded signal to noise ratio.

6.5.3 Results and Conclusions

The signal to noise measured on adjacent strip pairs for the LED and MQW readouts are shown in figure 6.28. Each histogram entry is calculated by summing the pedestal subtracted pulse height information from two adjacent strips and dividing by the noise from one of those channels. The noise peak at zero is removed by a 2.5σ cut. A Gaussian fitted to the data yields signal to noise ratios of 9.8:1 for LED read out data and 9.4:1 for MQW read out data.

Other data sets have been studied when the MQW bias voltage was changed from the default setting of $-9.0V$ to -8.5, -9.5, and $-10.5V$. This has the effect of making some channels operate in the non-linear region. However, the large spread in the channel pedestal values prevents detailed studies. The signal to noise ratio for the MQW was reduced by 10% at most.

The peak signal to noise ratio for the MQW channel used for the readout was measured to be $\sim 160:1$ in the laboratory. The driver circuit was adjusted so that the 4 MIP pedestal range plus an additional 2 MIP signal was accommodated in the modulator's linear range.

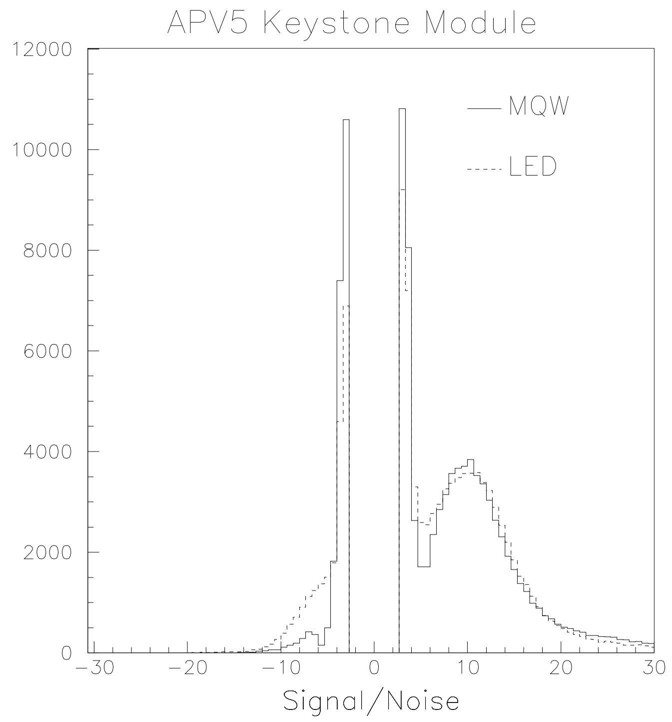


Figure 6.28: Pulse height distributions showing a comparison of LED and MQW based readout.

This results in a $\sim 27:1$ signal to noise ratio per MIP. The degradation in the signal to noise ratio from that measured for the 'perfect' LED based link to the MQW modulator link is consistent with this.

Chapter 7

Binary Optical Link

This chapter covers work carried out evaluating the radiation tolerance of the front-end optical package used in the binary optical link. Any degradation of the link components due to radiation must be measured to ensure that the performance after 10 years operation in LHC exceeds the specifications given in section 3.4.

Results of neutron and proton irradiation of LEDs and photodiodes are reported. The tests by the Birmingham group have concentrated on devices mounted in the GEC optical package. The LED measurements complement tests carried out by the Bern group [55, 56] on $\sim$ 150 bare die devices. The photodiode results are the first irradiation tests carried out on these devices. Testing packaged devices allows online measurement of the device characteristics throughout the irradiation period utilising the fibre connection. The radiation hardness of the packaging is also tested.

Two stages of tests have been carried out. For the first set of tests four optical packages were made available. However, previous laboratory tests on the devices resulted in subsequent problems in the mounting and wire bonding of the packages for the irradiation tests. The laboratory tests also showed that the photodiodes chosen for the prototyping of the package would be inadequate for use in the final package due to slow rise and fall-times of $15ns$ [20]. A yield of four working LED channels split between two packages and three photodiode channels were obtained for the irradiation tests. In the case of the LEDs this proved sufficient, when complemented by the Bern results, to allow the radiation tolerance to be characterised. The performance of the photodiodes however was erratic and impossible to interpret with the limited number available. The optical packages were irradiated with protons in the PS beam at CERN in November 1996 followed by neutron irradiation in the Birmingham Dynamitron in February 1997.

A further study concentrating on the photodiodes is being performed. Six optical packages have been tested. Three photodiodes were mounted in each package for the purpose of these tests. Two different types of photodiodes were tested: three packages containing bulk Silicon photodiodes, of the same kind used in the first tests, and three packages containing epitaxial Silicon photodiodes. The epitaxial devices are expected to be more radiation tolerant due to their smaller ionisation sensitive volume [57]. The tests on these packages are ongoing and only the first results of neutron irradiation performed in the Birmingham Dynamitron in May 1997 are discussed. Future work in evaluating the photodiode performance is planned by the Birmingham group. Higher neutron fluences have been studied to test the device's suitability for the readout of the pixel detectors. The second neutron irradiation was carried out during June and July 1997 in the ISIS radiation hardness test facility [58] which excluded the option of biasing and monitoring the devices. Measurements were made before and after the exposure. Proton irradiation of the diodes will be carried out during August and September 1997 in the CERN PS. However, time dictates that the results from these tests cannot be included in this thesis. Following these irradiations the rise and fall-times of the devices will be evaluated. This will require the remounting of the devices and therefore has not been attempted between the various irradiation stages. A further one package of each type of photodiode have been kept un-irradiated as a control sample for these tests.

7.1 Radiation Damage Theory

The damage process begins with the collision of a high energy incident particle with the semiconductor crystal lattice. The low displacement cross-section of the semiconductor atoms and small volume of semiconductor devices results in a uniform distribution of primary collision sites throughout the material. Furthermore, the recoil neutron will generally exit the device before it undergoes a second interaction. The struck atom (known as the primary knock-on atom, PKA) will be knocked from the lattice. It then rapidly loses energy due to both ionisation and the displacement of additional atoms. The division between the ionising and non-ionising damage is well known [59, 60]. The cascade results in a tree-like damage structure of crystal defects.

The first stage of the damage process is concluded when the PKA comes to rest. The vacancies and interstitials (known together as a Frenkel pair) produced will then diffuse through the semiconductor crystal until they form stable complexes. Many will recombine ($\sim 95\%$) leading to an immediate repair of the lattice. However, some will combine to form stable defects such as di-vacancies, vacancy-impurity complexes, vacancy-dopant complexes, and larger clusters. These defects form effective recombination and trapping centres resulting in a decrease in the minority carrier lifetime, carrier density and carrier mobility.

7.1.1 LEDs

The physical mechanism that results in the degradation of the light output from LEDs is the creation of non-radiative recombination centres which compete with the radiative centres in the recombination of excess carriers [61, 62]. The initial total minority carrier lifetime, τ_o, before irradiation is given by the combination of the lifetime associated with radiative processes, τ_{o_r}, and that for non-radiative processes, $\tau_{o_{nr}}$,

$$\frac{1}{\tau_o} = \frac{1}{\tau_{o_r}} + \frac{1}{\tau_{o_{nr}}}. \tag{7.1}$$

The lifetimes for the radiative and non-radiative processes can be related to the carrier capture cross-section, σ, and density, ρ, of the centres by the equations,

$$\frac{1}{\tau_{o_r}} = \sigma_r \nu_{th} \rho_{o_r}, \quad \text{and,} \qquad \frac{1}{\tau_{o_{nr}}} = \sigma_{nr} \nu_{th} \rho_{o_{nr}}, \tag{7.2}$$

$$\tag{7.3}$$

where ν_{th} is the thermal velocity of the minority carriers.

Following irradiation the total minority carrier lifetime, τ, is given by,

$$\begin{aligned} \frac{1}{\tau} &= \frac{1}{\tau_{o_r}} + \frac{1}{\tau_{o_{nr}}} + \sigma_{inr} \nu_{th} \rho_{inr} \\ &= \frac{1}{\tau_o} + \sigma_{inr} \nu_{th} \rho_{inr}, \end{aligned} \tag{7.4}$$

where the subscript inr refers to radiation induced non-radiative centres.

Assuming that the total number of radiation induced defects is small compared to the size of the crystal, so that damage caused by individual incident particles will not overlap[(a)], the density of radiation induced non-radiative centres can be theorised to be proportional to the radiation fluence, ϕ. It is given by,

$$\rho_{inr} \equiv C\phi, \tag{7.5}$$

where C is constant giving the density of defects produced per unit radiation fluence[(b)].

A damage constant κ can then be defined,

$$\kappa \equiv \sigma_{inr}\nu_{th}C, \tag{7.6}$$

which, when substituted into equation 7.4, yields the relationship between the minority carrier lifetime and the radiation fluence,

$$\frac{\tau_o}{\tau} = 1 + \tau_o\kappa\phi. \tag{7.7}$$

The damage constant conceals the detailed phenomenological information relating to the physical interactions between the incident particles and the semiconductor material.

The minority carrier lifetime must be related to the properties of the LED that are measured to determine its performance i.e. the light output power, P, as a function of the forward current. For an LED whose light output is due to a current which is diffusion controlled, the light output is related to the minority carrier lifetime by the equation,

$$P = C_1\tau e^{qV/kT}, \tag{7.8}$$

where, C_1 is a constant that is independent of τ and T,
q is the charge on an electron,
V is the forward bias voltage across the LED,
k is the Boltzmann constant,
and T is the absolute temperature of the LED.

(a) For the neutron irradiation of silicon it has been estimated that non-linear effects should not occur before fluence values exceed $\sim 2.5 \times 10^{15} ncm^{-2}$ [63].

(b) In practice more than one type of defect will be produced and a sum over the capture cross section and density of all the types must be made

Substitution into equation 7.7 gives the relation between the relative light output power and the radiation fluence for the LED when measurements are taken at a constant forward bias voltage,

$$\frac{\tau_o}{\tau} = \frac{P_o}{P} = 1 + \tau_o \kappa \phi. \tag{7.9}$$

However, the degradation of the light output power at constant operating current is a more useful relationship. This essentially tells us the efficiency of the device which is equal to the light output divided by the current. In addition the operation of the LEDs in the ATLAS readout will be controlled by the LDC which supplies a designated current to the LED.

To obtain an equation relating the LED light output power degradation to the minority carrier lifetime at constant current the dominant current flow mechanism in the LED must be known. For a total current dominated by diffusion, the total current density, J, is given by,

$$J = \frac{C_2}{\tau^{1/2}} e^{qV/kT}. \tag{7.10}$$

Solving this equation for $e^{qV/kT}$ and substituting into equation 7.8 gives a relationship between the light output power and the LED drive current,

$$P = C_3 \tau^{3/2} J, \tag{7.11}$$

where, C_2 and C_3 are constants.

Substituting into equation 7.7 as before yields the equation,

$$\frac{\tau_o}{\tau} = \left(\frac{P_o}{P}\right)^{2/3} = 1 + \tau_o \kappa \phi, \tag{7.12}$$

for the case of diffusion dominated radiative and total currents.

Similarly for a space-charge recombination dominated total current, the total current density is given by,

$$J = \frac{C_4}{\tau} e^{qV/2kT}, \tag{7.13}$$

which can be shown to give a degradation,

$$\frac{\tau_o}{\tau} = \left(\frac{P_o}{P}\right)^{1/3} = 1 + \tau_o \kappa \phi, \tag{7.14}$$

when measured at constant current.

In each case the degradation becomes significant when the product $\tau_o \kappa \phi$ becomes comparable with unity. The damage constant κ is dependent on the NIEL of the incident particles, as discussed in section 3.4.6, and the semiconductor material of the device and is therefore essentially fixed. However, by heavily doping the optical region of the LED the initial radiative lifetime, τ_{o_r}, and therefore the pre-irradiation minority carrier lifetime, τ_o, can be decreased resulting in a more radiation hard device. This has the added beneficial side-effects of increasing the LED efficiency and speed.

7.1.2 Photodiodes

The performance of a photodiode as a detector can be characterised by three parameters:

- The dark current is the leakage current when there is no light incident on the photodiode,
- The responsivity of the photodiode describes its efficiency in converting light into an electrical signal,
- The rise and fall time of the photodiode limits its speed of operation.

Radiation damage produces defects which can result in effects such as [64]:

- Increased dark currents as defects act as centres to increase the generation bulk current,
- Degraded responsivity as defects act as electron or hole trapping centres for the photogenerated pairs, resulting in an increased probability of recombination,
- Degraded rise and fall-times due to de-trapping or a reduction in the carrier mobility.

The increase in the device dark current is expected to be the major change in thin junction devices such as photodiodes. The changes in the device response and rise and fall-times are expected to be small, but still require measuring.

The increase in dark current can be related to the minority carrier lifetime of the semiconductor if the generation-recombination is dominated by mid-band levels caused by defects. Another source affecting the dark current could be ionising damage to the surface of the device, which would generally lead to a catastrophic rise in the current. The dark current, I_{dc}, for a device having depletion depth w and active area A is given by,

$$I_{dc} = \frac{qAwn_i}{2\tau}, \tag{7.15}$$

where, q is the electron charge,
n_i is the intrinsic carrier concentration,
and τ is the minority carrier lifetime.

The relationship of the minority carrier lifetime to the incident radiation fluence on the device is described by equation 7.7 as,

$$\frac{\tau_o}{\tau} = 1 + \tau_o \kappa \phi.$$

Rearranging this equation and then substituting equation 7.15 yields the relationship,

$$\Delta I_{dc} = \kappa \phi, \tag{7.16}$$

for the radiation induced change in the dark current.

7.2 Annealing of Radiation Damage

A significant proportion of radiation damage has been observed to recover under forward bias (injection annealing) [61, 62, 65] and at increased temperatures $\sim 200^oC$ (thermal annealing) [62, 66]. In GaAs p^+n junctions the annealing has been studied under forward bias [67, 68] using Deep Level Transient Spectroscopy [69]. The annealing process was found to be directly related to recombinations at the defect.

7.3 Irradiation of LEDs

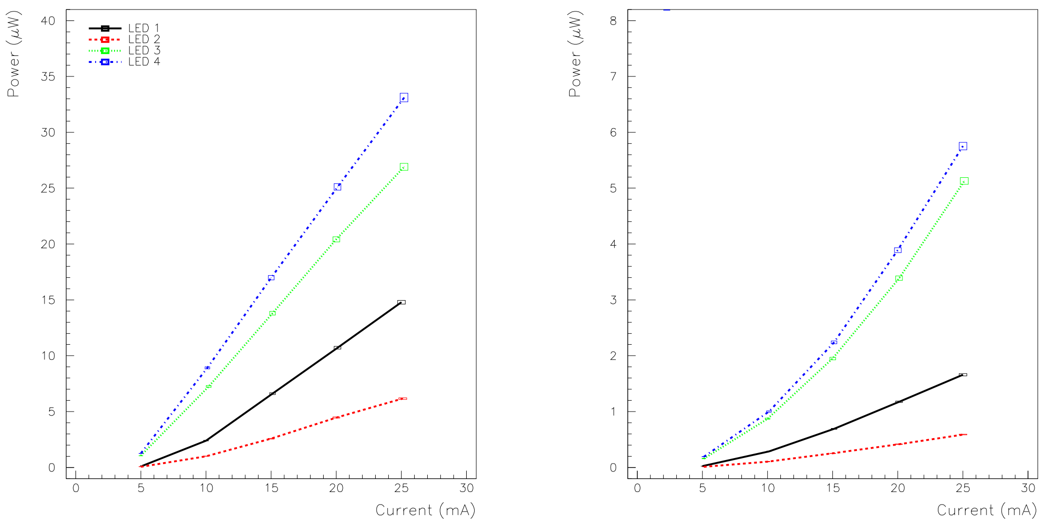
The LEDs used in the GEC optical package are fabricated from GaAlAs and emit light at $850nm$. Four of these LEDs have been monitored throughout proton irradiation and a subsequent annealing period followed by neutron irradiation and an annealing period. The power output of the LEDs has been monitored, at intervals, at forward currents of $5-25mA$ in steps of $5mA$ using an optical power meter. A 3^{rd} order polynomial fit is made to the data points which can then be used to calculate the light power at any given drive current. During the irradiation periods the LEDs were continuously forward biased at $10mA$, except for the short periods required for the measurements. For the annealing periods the effects of increased bias currents of up to $50mA$ were investigated.

7.3.1 Proton Irradiation

The optical packages were installed in the coolbox used for the irradiation of ATLAS Silicon microstrip detectors. The radiation hard fibre pigtails attached to the devices were connected to $\sim 40m$ of normal multimode fibre leading to the barracks area. The current to the LEDs was controlled using a simple emitter follower circuit attached to each LED by twisted pair cable. The coolbox was maintained at a nominal temperature of -7^oC. However, technical difficulties with the liquid Nitrogen cooling system led to temperature fluctuations from $\sim +30^oC$ to -40^oC.

The power output as a function of the drive current for the LEDs is shown in figure 7.1. The response before irradiation and after a $24GeV$ proton fluence of $\sim 0.4 \times 10^{14}\ cm^{-2}$ is shown. It should be noted that the response becomes non-linear after irradiation. This is due to a change in the dominant total current control mechanism in the device as the drive current is increased.

The results of the proton irradiation are shown in figure 7.2. The first plot shows the relative light output power of the four LEDs plotted against time. The light output of each LED for a $20mA$ drive current is calculated from the fit to the data points. The data have then been normalised to the data point taken just before the start of the proton irradiation. The first two data points are taken before the Nitrogen cooling was started. A clear negative temperature dependence of the light output of the LEDs is observed of magnitude $\sim -1\%/^oC$. A degradation in the relative light output correlated with the irradiation period is clearly seen. This is highlighted by the plot of relative light output against fluence. The step in light output at $\sim 30 \times 10^{12} p\ cm^{-2}$ is associated with a halt in the radiation while the liquid Nitrogen dewar was refilled. The temperature of the box increased dramatically in this period resulting in a decrease in the LED efficiency. LED 4 shows a pronounced drop in power, over that expected by the radiation degradation, for a short period. The subsequent recovery back to the expected level is due to the connector being cleaned. The fluence was



(a) Pre-irradiation.

(b) After $\sim 0.4 \times 10^{14} p\ cm^{-2}$.

Figure 7.1: The optical power output of LEDs before and after irradiation. The LEDs 1 and 2 are from package S/N 04 and LEDs 3 and 4 from S/N 11.

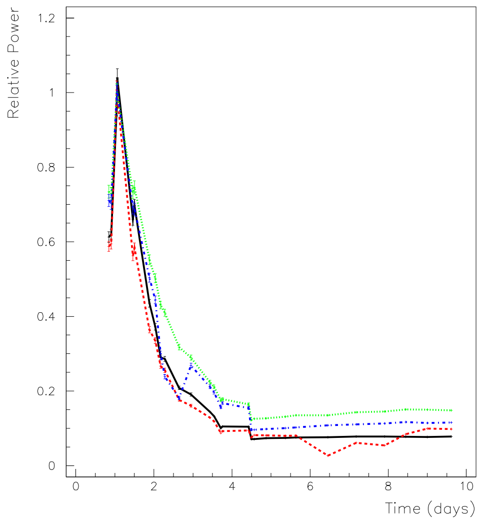
measured at the corners and centre of the $6cm \times 6cm$ scanning area. Variations of up to 50% were seen between the different foil locations. An average of the foil values was taken for the final fluence. The fluence value can be expected to have a systematic error $\sim 10\%$ from the individual foil measurements. The difference in fluences received by the different packages is estimated at $\lesssim 10\%$ as the devices are located towards the centre of the scanning region.

The initial minority carrier lifetime has not been measured in these tests. A new damage constant K is therefore defined as the lifetime-damage constant product,

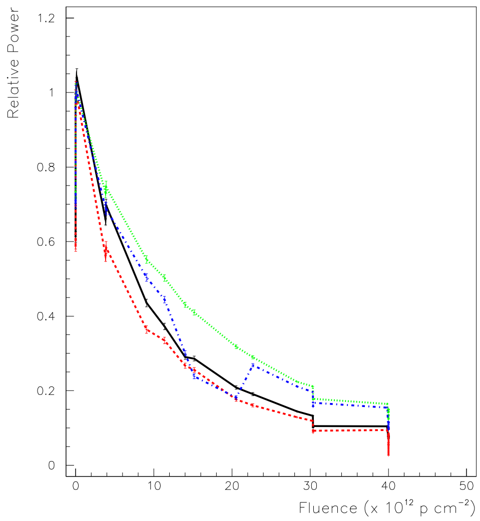
$$K \equiv \kappa \tau_o. \tag{7.17}$$

Before the proton damage constant, K_p, for the LEDs can be calculated the current control mechanism must be established (see section 7.1.1). Figure 7.3 shows plots of $\log((\frac{P_o}{P})^{2/3} - 1)\ vs\ log\phi$, for the diffusion dominated current case, and $\log((\frac{P_o}{P})^{1/3} - 1)\ vs\ log\phi$, for the space-charge recombination dominated current case. A slope of unity is expected for the case that describes the current control mechanism of the LED. These graphs indicate that both the radiative and total currents of the LEDs are diffusion controlled at a $20mA$ drive current.

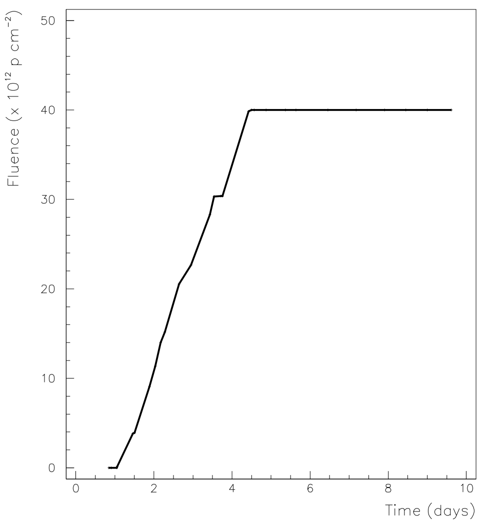
A fit has been made to determine the damage constant, as shown in figure 7.4, yielding



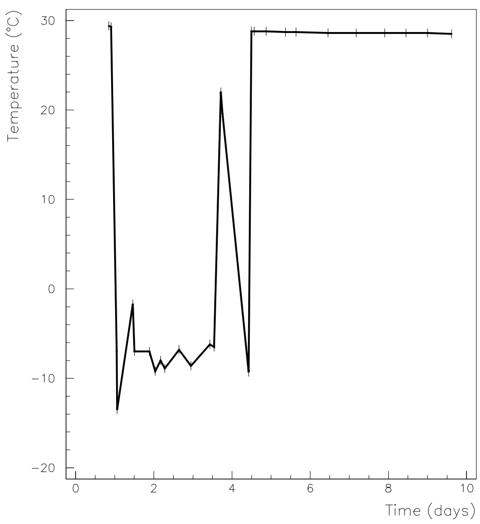
(a) Relative Power vs Time.



(b) Relative Power vs Fluence.

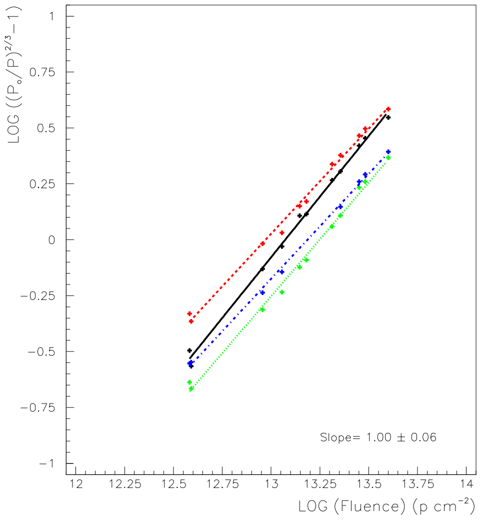


(c) Fluence vs Time.

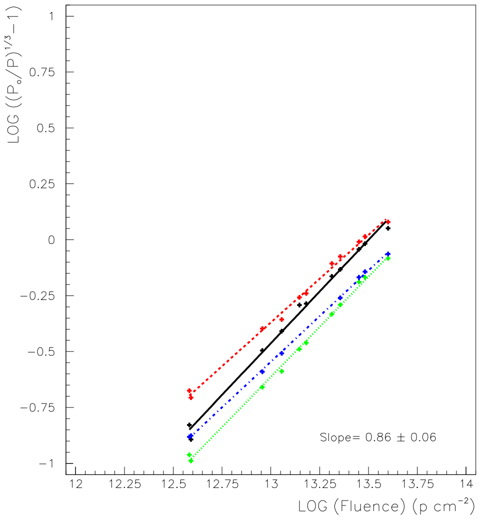


(d) Temperature vs Time.

Figure 7.2: Results of LED proton irradiation.



(a) Diffusion dominated.



(b) Space-charge recombination dominated.

Figure 7.3: Determination of the dominant current flow mechanism in the LEDs.

the following results:

LED 1 $9.04 \pm 0.14 \times 10^{-14} cm^2$,

LED 2 $10.13 \pm 0.15 \times 10^{-14} cm^2$,

LED 3 $5.77 \pm 0.11 \times 10^{-14} cm^2$,

LED 4 $6.31 \pm 0.13 \times 10^{-14} cm^2$.

A systematic error of $\sim 10\%$ can be ascribed to these results due to the errors associated with the radiation fluence measurement.

A difference in the radiation tolerance calculated for the LEDs in the separate packages can be seen. This could be due to the packages receiving a different fluence. The estimated difference in the fluence received by the different packages of $\lesssim 10\%$ does not account for the differences in the damage constants. It would therefore point to a real difference in the devices themselves. This could be due to either some annealing of the damage even at the low currents used during the irradiation period, or due to a lower starting minority carrier lifetime due to different doping between the devices.

The values can be compared with results taken by the Bern group [56] for LEDs that were either unbiased or pulsed with a $20mA$ forward current during proton irradiation also carried out in the CERN PS. The damage constants reported were $K_p = 9.3 \pm 0.2 \times 10^{-14}cm^2$ for unbiased LEDs and for those pulsed a lower value of $K_p = 6.8 \pm 0.2 \times 10^{-14}cm^2$ was obtained. The first value agrees well with that obtained for LEDs 1 and 2 in the packaged device tests. The lower values obtained for LEDs 3 and 4 appear consistent with there having been some annealing during the irradiation.

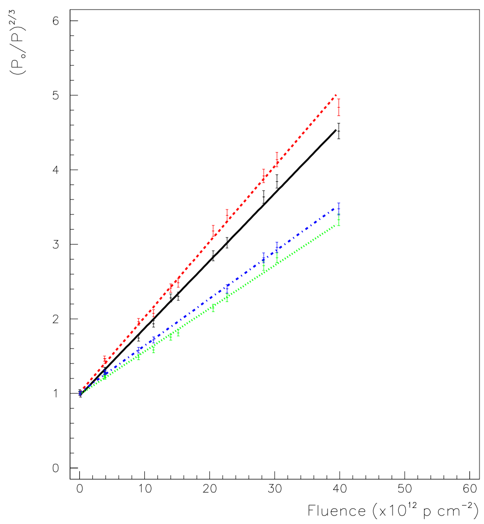
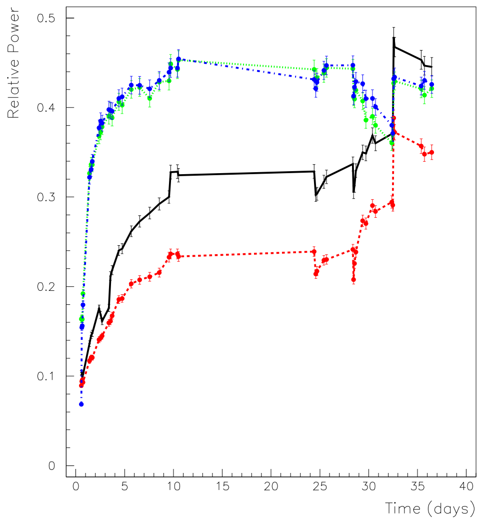


Figure 7.4: Plot of $(\frac{P_o}{P})^{2/3}$ vs ϕ to determine the proton damage constant for the LEDs.

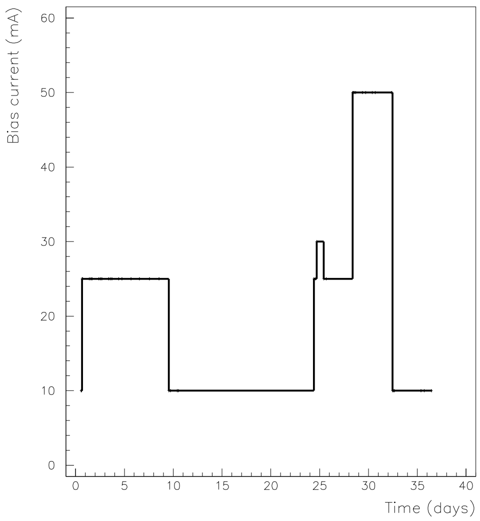
7.3.2 Post-proton Irradiation Annealing

Following the proton irradiation the LEDs were monitored for an extra five days while still in the beam area. A small amount of annealing of LEDs 3 and 4 is apparent, but no annealing of LEDs 1 and 2 can be detected. The devices were then stored unbiased for a week, to allow the induced activity to fall to acceptable levels, before transportation back to Birmingham for further studies. At Birmingham the effect of higher forward currents on the annealing rate was investigated. Figure 7.5 shows the results of the tests.

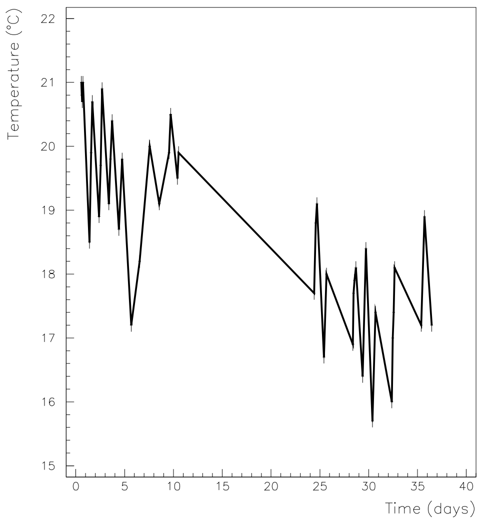
Measurements of the LED power versus current response were made in the same way as previously. However, the bias current of the LEDs between the measurements was altered throughout the tests. A preliminary set of measurements when the bias current had been



(a) Relative power.



(b) Ambient temperature.



(c) Bias current.

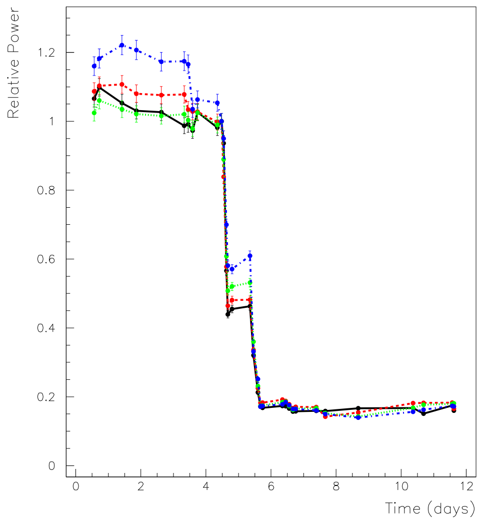
Figure 7.5: Results of post-proton irradiation annealing.

set to $10mA$ for several hours was first taken. This allowed the data to be normalised to the relative light output measurement taken before the equipment was dismantled for transportation. The bias current was then increased to $25mA$, resulting in significant recovery in the light output of all the LEDs over a period of several days. It should be noted that the rate of recovery is faster for LEDs 3 and 4. Another period at the lower bias current of $10mA$ highlights the negligible annealing at this current. Further running at $25mA$ followed by $50mA$ results in a plateauing of the recovery with the higher power LEDs even showing signs of an anti-annealing process. It is interesting to note that although the rate of annealing of the LEDs is different the final value obtained is quite consistent. A recovery to $> 30\%$ relative light output is observed for all the LEDs.

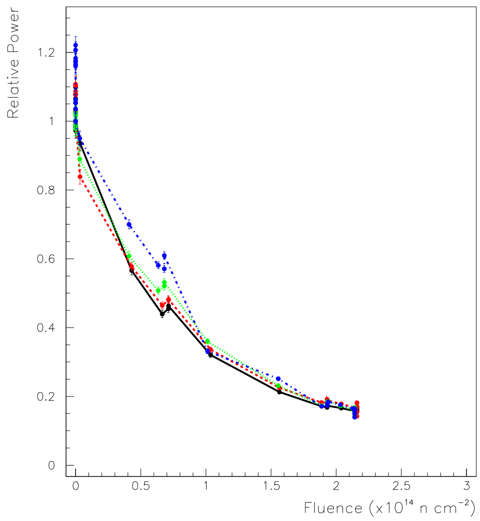
7.3.3 Neutron Irradiation

The same optical packages were subsequently irradiated in the Birmingham Dynamitron. The devices underwent three days of testing which consisted of ~ 8 hour irradiation periods during the day, but no irradiation overnight. The same monitoring technique was used as for the proton irradiation. Figure 7.6 shows the results of the neutron irradiation. The relative light output is normalised to the the last light power value measured just before the start of the irradiation. In fact a small period of beam was incident on the devices the day before, however the run had to be terminated due to problems with the Beryllium target. A small drop in the light outputs due to this exposure is observed. The fluence received by each of the optical packages was measured using individual foils, although in fact each device received almost exactly the same fluence. The low fluence achieved on the final day of running was due to a change in the position of the devices. They were placed lower (further from the target) to allow the irradiation of some devices for the Bern group. The degradation observed in the relative light output of the LEDs is clearly correlated with the irradiation period.

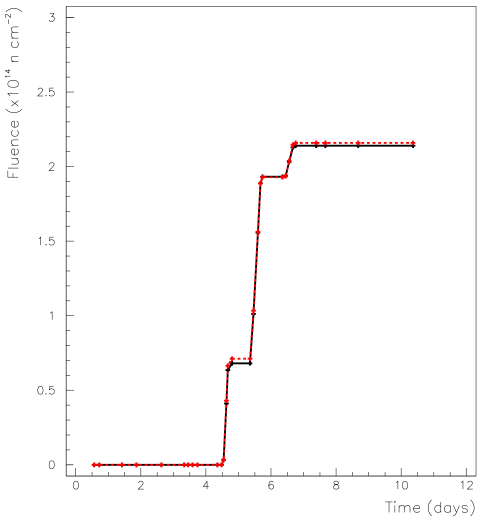
The damage constant for the neutron irradiation has been determined assuming a diffusion dominated radiative and total current as previously. Figure 7.7 shows two cases for the $(\frac{P_o}{P})^{2/3}\ vrs\ \phi$ fit used to determine the damage constant. In the first case the relative output power of each LED has been normalised to the light output power measured at the start of the neutron irradiation (i.e. the same normalisation as used for figure 7.6). The resulting



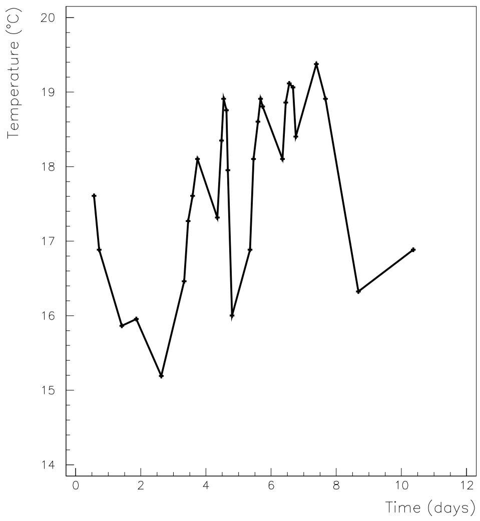
(a) Relative Power vs Time.



(b) Relative Power vs Fluence.



(c) Fluence vs Time.



(d) Temperature vs Time.

Figure 7.6: Results of LED neutron irradiation.

damage constants in this case are:

LED 1 $1.13 \pm 0.02 \times 10^{-14} cm^2$,

LED 2 $1.08 \pm 0.02 \times 10^{-14} cm^2$,

LED 3 $1.09 \pm 0.02 \times 10^{-14} cm^2$,

LED 4 $1.12 \pm 0.02 \times 10^{-14} cm^2$.

plus $\sim 10\%$ systematic error.

However, it should be remembered that the damage constant, K, is defined in equation 7.17 as the initial minority lifetime-damage constant product. Since the neutron irradiation follows prior irradiation with protons there will already have been an increase in non-radiative defects and therefore a decrease in the minority carrier lifetime. The second case for the determination of the damage constant is a fit when the relative light output of the LEDs is normalised relative to the original power output of the LEDs before any irradiation. This is achieved in practice by scaling the relative light output values calculated for the neutron irradiation to the relative light output values obtained after the post-proton irradiation annealing. The values for the neutron damage constant, K_n, obtained by this method are:

LED 1 $1.92 \pm 0.02 \times 10^{-14} cm^2$,

LED 2 $2.17 \pm 0.02 \times 10^{-14} cm^2$,

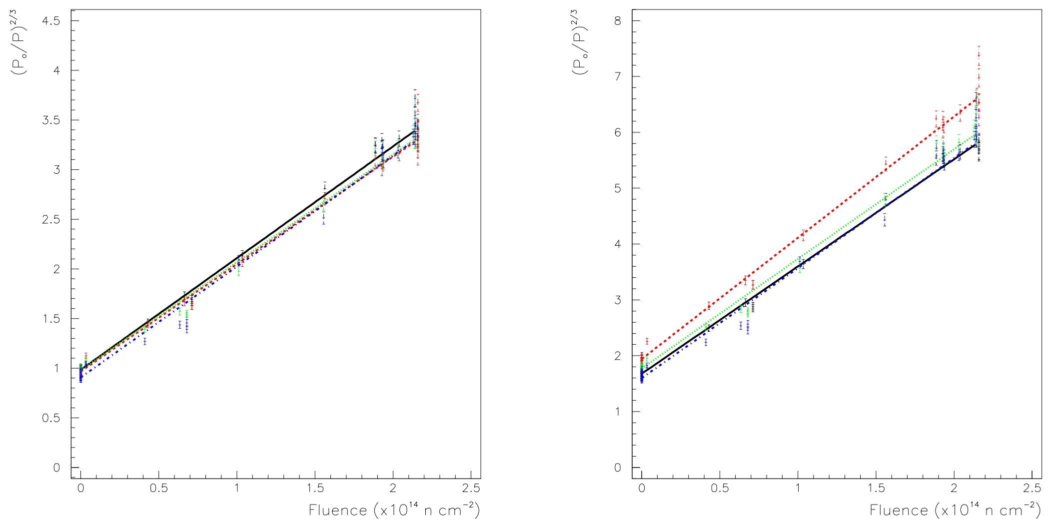
LED 3 $1.96 \pm 0.02 \times 10^{-14} cm^2$,

LED 4 $1.97 \pm 0.02 \times 10^{-14} cm^2$.

plus $\sim 10\%$ systematic error.

The results can again be compared with the tests performed by the Bern group [56] which resulted in a value $K_n = 2.95 \pm 0.03 \times 10^{14} cm^{-2}$. The Bern irradiations were performed in the ISIS facility and so assuming the damage scales with the NIEL of the neutrons a value of $K_n = 1.97 \pm 0.2$[c] $\times 10^{14} cm^{-2}$ would be expected for the Dynamitron neutron energy distribution. This value is consistent with the damage constants calculated by the second method for the packaged LEDs.

[c]The larger error is due to the large uncertainty in the values for the NIEL of $\sim 10\%$



(a) Normalised to pre-neutron irradiation light output.

(b) Normalised to pre-proton irradiation light output.

Figure 7.7: Plot of $(\frac{P_o}{P})^{2/3}$ vrs ϕ to determine the neutron on damage constant for the LEDs.

7.3.4 Post-neutron Irradiation Annealing

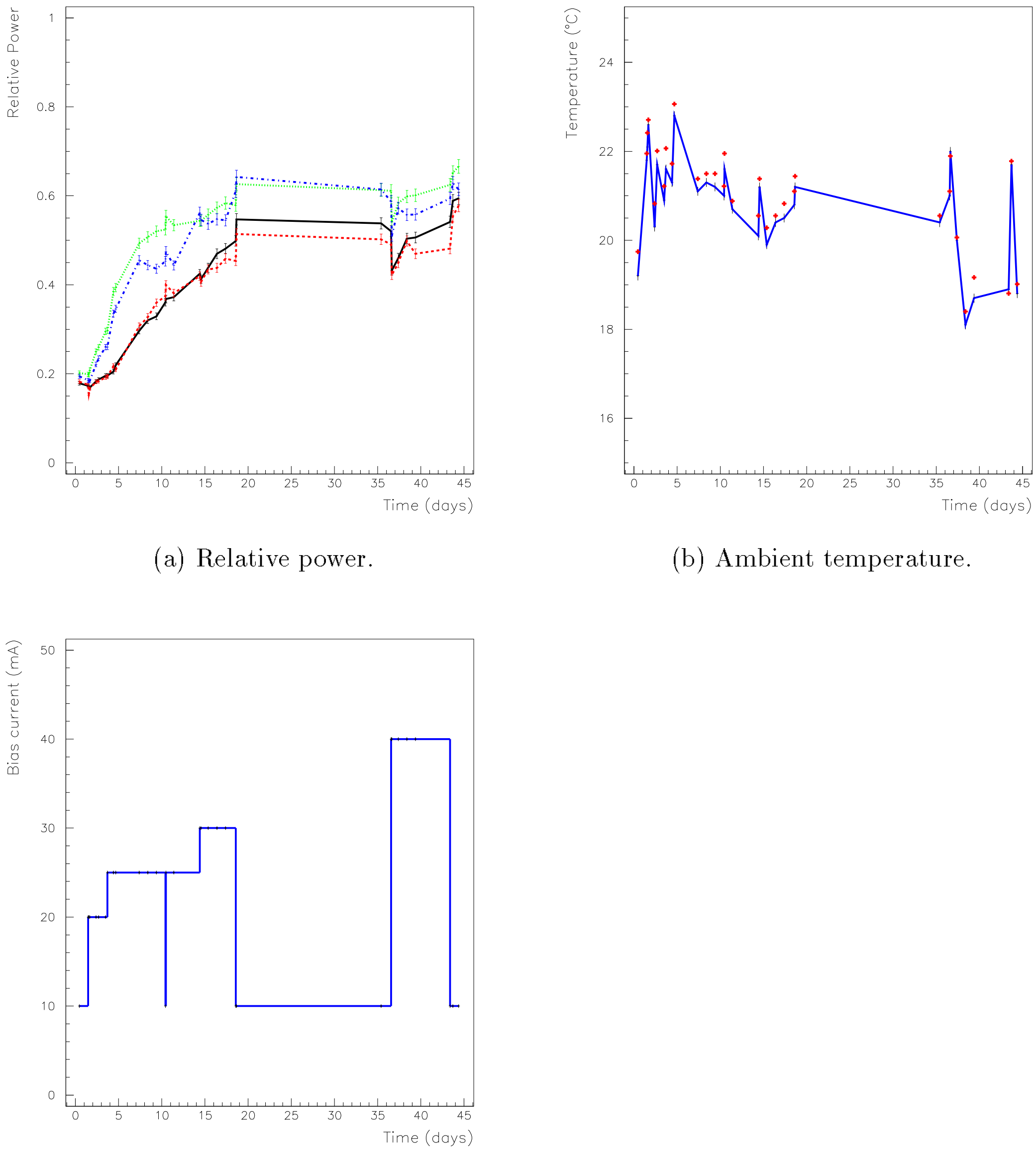
The LEDs were again run at higher currents in the same way as after the proton irradiation. The currents were limited to $40mA$ to avoid any anti-annealing that was observed previously. Figure 7.8 shows the recovery of the LEDs and the operating conditions during the annealing period.

Similar results to the post-proton irradiation are observed. All the LEDs recover to $> 57\%$ relative to the light output before the neutron irradiation.

7.3.5 Conclusions

The degradation of LEDs resulting from radiation exposure can be described by a damage constant. The damage due to different types and energies of particles can be related by the NIEL of the different particles. For the irradiations performed in the Birmingham Dynamitron and CERN PS a damage constant ratio $1 : 4.8 \pm 0.4$ is obtained. This agrees well with the expected ratio from the NIEL calculations of $1 : 4.3 \pm 0.6$.

It has also been shown that re-irradiation of devices can be described if the damage of



(a) Relative power.

(b) Ambient temperature.

(c) Bias current.

Figure 7.8: Results of post-neutron irradiation annealing.

the devices is always related to the performance of the non-irradiated device.

The total damage to the LEDs after 10 years LHC operation can be calculated from the expected fluences (see section 3.4.6) throughout the inner detector and the damage constants measured in these tests. Rearranging equation 7.12, we can obtain an expression for the relative light output, P/P_o, after a given fluence ϕ,

$$\frac{P}{P_o} = \left(\frac{1}{1+K\phi}\right)^{3/2}. \tag{7.18}$$

The expected relative light output is $\sim 3\%$ for LEDs located at the 1^{st} SCT barrel and $\sim 0.3\%$ for those located at the the outer pixel detector layer. However, this calculation neglects the beneficial effects of injection annealing.

The injection annealing of the devices results in a large recovery of the radiation induced damage. The annealing has not been studied extensively as a function of received fluence, however results have been taken for devices irradiated up to the fluences expected in the SCT. The Bern group has reported that LEDs irradiated to approximately double the fluence expected at the 1^{st} SCT barrel show recoveries to $10-30\%$ relative light output [56]. The neutron irradiation at Birmingham reached fluences of $\sim 1/3$ this and recovered to $\sim 55\%$ light output.

It has been estimated [20] that a degradation to a relative light output of only 20% would still give a link signal to noise ratio of $\sim 65:1$. This is a factor three higher than that required for data readout. With annealing the LED performance is sufficient for operation at the radius of the 1^{st} SCT barrel with a substantial safety margin. Operation for the readout of the pixel detectors seems marginal even if the optical packages are mounted at the outer pixel layer. High current injection annealing during the operation of ATLAS is incorporated into the operational schedule for the detector [20].

7.4 Irradiation of Photodiodes

Two types of Silicon PIN photodiodes have been tested. The first type of device are fabricated from bulk Silicon wafers which has dopants diffused into the semiconductor to produce the n and p-type regions. The second type of device starts out as a thin silicon wafer which then has epitaxial layers of differently doped material grown on top. The epitaxial devices can therefore be made thinner which results in a slightly lower quantum efficiency, but is expected to make the devices more radiation tolerant.

The photodiodes were biased at $-10V$ during the tests. This was sufficient to deplete fully the devices and is the expected reverse bias required for fast operation in the readout. A $100k\Omega$ series resistor allowed the current through the device to be monitored. The dark current, when no light was incident on the photodiode, and the photocurrent for incident light powers from $3-20\mu W$ in five equal steps was measured. From these measurements the response of the photodiode was determined. The first step is to make a 3^{rd} order polynomial fit to the photocurrent versus incident light power points. This can be used to calculate the incident light on the photodiode for a given LED drive current[d]. A similar fit is then made to the photocurrent vs incident light power points. The response is calculated from the slope of the fitted function for a launched optical power of $10\mu W$ into the photodiode fibre.

7.4.1 Preliminary Bulk Silicon Photodiode Irradiation

Only three Bulk Silicon photodiodes were available for the first set of radiation tests. They underwent the same cycle of irradiations as the LEDs described in the previous section. An overview of the results from these tests is given before more detailed discussion of the tests on a larger sample of photodiodes.

During the proton irradiation the photodiode response was measured, but not the dark current. The photodiodes were exposed to a total proton fluence of $0.4 \times 10^{14} cm^{-2}$. The response of the photodiodes was measured before the exposure began when the diodes were at both the beam area temperature, 30^oC, and the nominal coolbox temperature, -7^oC. In both cases the response of all three photodiodes was measured to be $\sim 0.5A/W$. The three photodiodes responded in different ways during the exposure period. The response of one photodiode dropped by only $\sim 10\%$ while another fell to $0.1A/W$, only 20% of the starting value, and the third did something in between. However, during a short stop in the irradiation, while the coolbox dewar was refilled, and following the the termination of the irradiation all three photodiodes recovered to close to their pre-irradiation values over

[d]The incident light power is measured at the launch point into the fibre patchcord running between the barracks and photodiode. The incident light on the photodiode can be expected to be systematically lower than the measured value (by $\sim 5\%$) due loss in the fibre and at connectors. However, we are more worried by changes in the response.

a period of a day. The annealing is correlated with a rise in temperature to that of the surroundings in each case. It is hard to draw quantitative conclusions on the performance of the diodes from these three different cases, however the large drop seen by one photodiode is by itself worrying.

Following the return of the devices to Birmingham the photodiodes were exposed to neutron fluences in excess of $2 \times 10^{14} cm^{-2}$. During these tests the monitoring was changed to allow the dark current to be measured in addition to the response of the diodes. The photodiodes behaved in a consistent, but disappointing, manner during these tests. All three devices showed a rapid fall in response to $< 0.1A/W$ after only a short period of irradiation. No subsequent recovery was measured. The dark currents of the diodes was measured at $\sim 15nA$ before the neutron irradiation. This shows a deterioration from the expected pre-irradiation value of $< 1nA$ due to the proton exposure. The dark current rose steadily throughout the neutron exposure to a final value of $\sim 0.8\mu A$ after the full neutron fluence. The dark current dropped and levelled out at $\sim 0.5\mu A$ over the two day period following the irradiation.

The increase in the dark current although large is well behaved and does not indicate a catastrophic failure in the device. However, the large decrease in the response of the photodiodes due to the neutron irradiation is not expected. Further tests on photodiodes have been performed and the results are discussed in more detail below.

7.4.2 Bulk Silicon Photodiode Irradiation

A yield of eight working photodiodes from the possible nine was achieved. The one diode not working was due to a bad wire bond giving no electrical connection rather than a problem with the device itself.

Neutron irradiation of these devices has been carried out using the Birmingham Dynamitron. The running schedule consisted of three ~ 8 hour long daytime irradiations followed by twelve days monitoring with the devices still in the test area.

The results of the irradiation on the photodiodes is shown in figure 7.9. The response of all the diodes drops rapidly during the first irradiation period. The plot of relative response versus the fluence highlights this rapid fall, due to a neutron fluence of only $\sim 0.25 \times 10^{14} cm^{-2}$, to relative response of only 10%. The degradation saturates at this level for the remainder of the irradiation and no post-irradiation recovery is observed. The dark current of the diodes shows a large increase over the irradiation period. The initial dark current for all the diodes was too small to measure and consistent with the $< 1nA$ value quoted by the supplier. The final value was $\sim 500nA$ immediately after the irradiation dropping to $\sim 300nA$ after several days. Equation 7.16 predicts a linear increase in the dark current with fluence which can be characterised by the damage constant K_{bulk}. The annealing between the three irradiation periods, and presumably during the irradiation, probably invalidates any fit to the data. However, if only the final value just after the finish of the irradiation is considered a damage constant $K_{bulk} \approx 3.3 \times 10^{-12} cm^2$ is obtained for a change in dark current measured in nA.

An additional neutron irradiation has been carried out in the ISIS facility [58]. $1MeV$ equivalent neutron fluences of $7.9 \pm 1.2 \times 10^{13} cm^{-2}$ were achieved over a 38 day irradiation period. The response of the diodes was unchanged. The increase in the dark currents was consistent with that expected from the damage constants.

7.4.3 Epitaxial Silicon Photodiode Irradiation

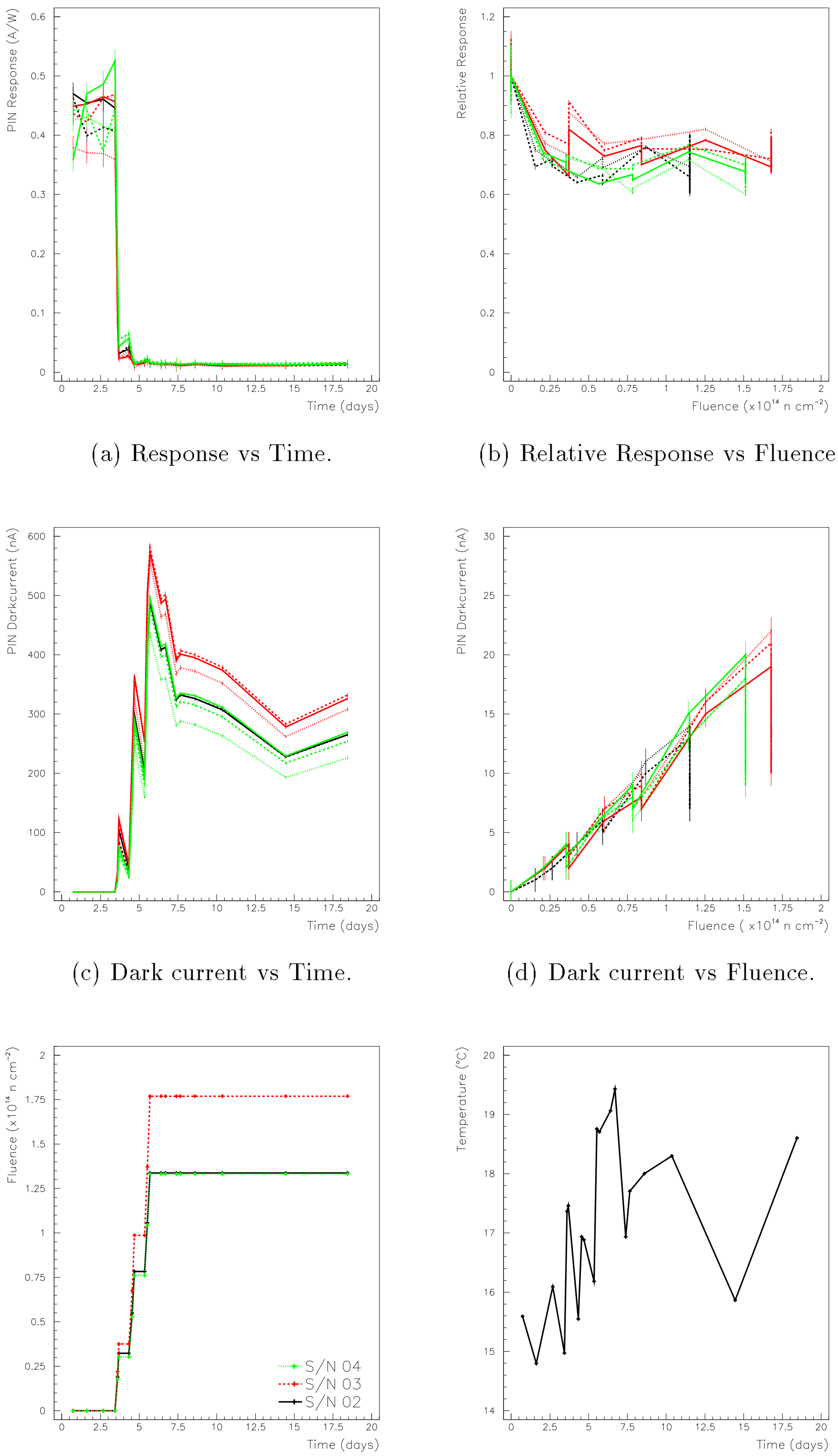
The epitaxial Silicon photodiodes were irradiated at the same time as the bulk Silicon diodes. The measurements and analysis of the data were performed in exactly the same way as previously described.

The results of the irradiation tests are shown in figure 7.10. The response of the diodes exhibits the same immediate drop and saturation as the bulk photodiodes, but with the important distinction that saturation occurs after a fall of only $\lesssim 30\%$. The dark current rises throughout the irradiation as expected with a damage constant $K_{epitaxial} \approx 0.13 \times 10^{-12} cm^2$ ignoring the annealing.

The additional neutron irradiation in the ISIS facility resulted in no change in the response of the diodes. The increase in the dark currents was again consistent with the damage constant expectation.

7.4.4 Conclusions

The dark current and response of both Bulk Silicon and Epitaxial Silicon photodiodes has been measured throughout irradiation with neutrons. Approximate damage constants have



(a) Response vs Time.

(b) Relative Response vs Fluence.

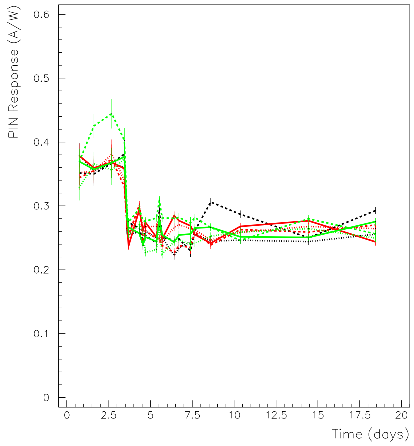
(c) Dark current vs Time.

(d) Dark current vs Fluence.

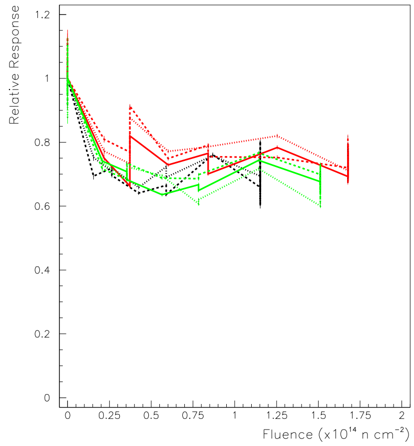
(e) Fluence vs Time.

(f) Temperature vs Time.

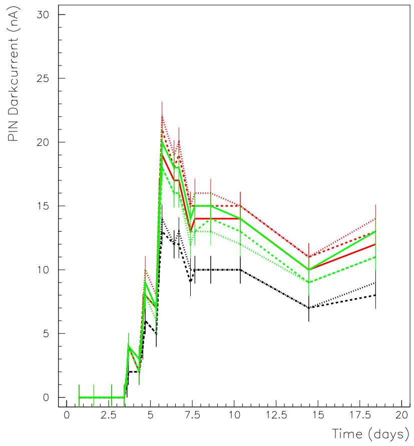
Figure 7.9: Results of Bulk Silicon photodiode neutron irradiation.



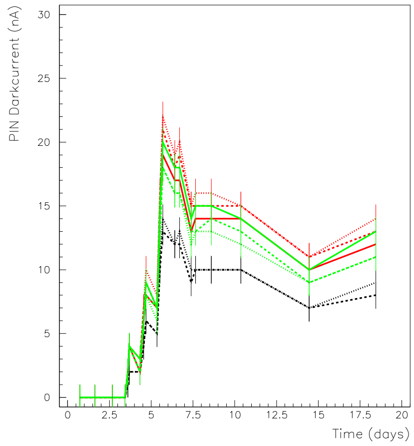
(a) Response vs Time.



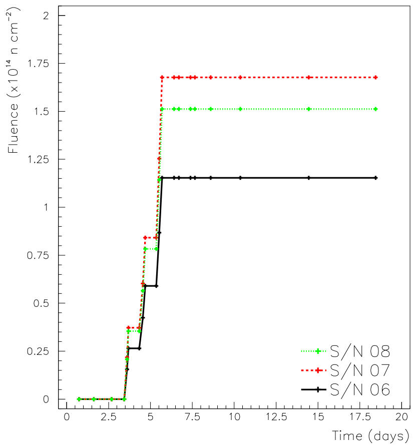
(b) Relative Response vs Fluence.



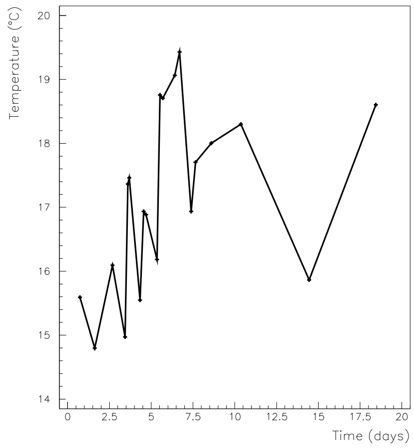
(c) Dark current vs Time.



(d) Dark current vs Fluence.



(e) Fluence vs Time.



(f) Temperature vs Time.

Figure 7.10: Results of Epitaxial Silicon photodiode neutron irradiation.

been found for the increase in dark current for the devices, the values are $3.0 \times 10^{-12}cm^2$ for the Bulk Si diodes and $0.12 \times 10^{-12}cm^2$ for Epitaxial Si diodes when scaled to $1MeV$ neutrons. After the 1^{st} SCT barrel expected $1MeV$ neutron fluence of $1.3 \times 10^{14}cm^{-2}$ the expected dark currents are $\sim 390nA$ and $\sim 16nA$ for the Bulk and Epitaxial Si devices respectively. The dark current damage in the Epitaxial Si devices is much lower due to the smaller volume of the devices by $\sim 10\times$[e].

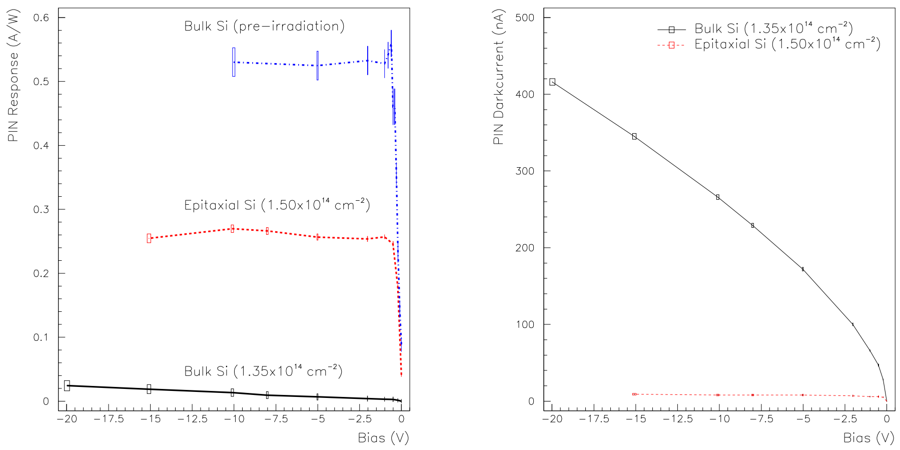
The response of both types of diode decrease after neutron fluences of only $\sim 0.25 \times 10^{14}cm^{-2}$. However, after the initial drop the diodes response is then unaffected for $1MeV$ neutron equivalent fluences up to $\sim 2.7 \times 10^{14}cm{-}2$ as expected. The initial fall in the response of the Bulk Si diodes is fatal, however, the Epitaxial diodes exhibit a fall of $\lesssim 30\%$ which is acceptable for use in the final package.

One explanation for this initial response degradation could be type inversion of the device substrate [70]. The incident radiation causes the formation of acceptor levels which result in the low-doped n-type region changing to a low-doped p-type region. The result is that the depletion region now grows from the p-type side of the diodes rather than the n-type side. The distance for absorption of 99% of $820nm$ light in Silicon is only $\sim 45\mu m$ and so it is possible that the light will not reach the depletion region in the bulk Si devices if the devices are not fully depleted.

Figure 7.11 shows the response and dark current for typical bulk and epitaxial diodes after neutron irradiation as a function of the applied bias voltage. The response of the a bulk diode before irradiation is included for comparison. The response has been calculated using the method described above. It can be observed that the response of the epitaxial diode and the bulk diode before irradiation quickly reach a plateau at a low reverse bias. This indicates that the epitaxial device is fully depleted and the bulk device is depleted at least as deep as $\sim 45\mu m$. It can also be observed that the epitaxial diode dark current approaches a plateau again indicating full depletion. However, the bulk diode after irradiation shows neither of these attributes. The response rises approximately linearly with the applied voltage and the dark current is still increasing at a bias voltage of $-20V$, the maximum specified voltages for the diodes. These results indicate that the bulk diodes are not being fully depleted after

[e] The depletion region thicknesses are $\sim 15\mu m$ for the Epitaxial diodes and $\sim 100 - 150\mu m$ for the Bulk devices. The difference in volumes will therefore be of the order a factor 10.

neutron irradiation.



(a) Response.　　(b) Dark current.

Figure 7.11: Measurement of the photodiode voltage characteristics after irradiation.

Chapter 8

Twisted Pairs

8.1 Introduction

This chapter covers work carried out evaluating the performance of the binary electrical link described in section 4.4. The link has been investigated as an alternative to the optically based binary link due to concerns about the radiation tolerances of the optical components and a projected factor three cost saving over the optical system. It utilises thin shielded twisted pair (STP) cables in an effort to minimise both the material introduced into the inner detector region and the electromagnetic radiation and pick-up of the system. A transition to cheaper unshielded twisted pair (UTP) cables is made outside the ATLAS calorimeters.

The electrical link has the added benefit of not requiring the specialised cooling contact that the LEDs require to maximise their lifetime [20, 56]. However, the electrical link does have drawbacks. It firstly introduces more material into the inner detector which is not preferred on physics grounds. Additional concerns are the possibility of both the radiation and pick-up of electromagnetic noise and the introduction of ground loops into the detector system. The SCT is designed as a groundless system to minimise the cross-coupling between the system elements. Separate power supplies are used for each module which are referenced to one another using high-value resistors. There are a number of ~DC connections between the modules and off-detector electronics such as power cables, calibration lines, control lines, bias voltage cables, and monitoring cables. An electrically based link would introduce a high-frequency connection running in close proximity which could radiate to the other cables. The shielding of the cables introduces the possibility of ground loops where different power supply grounds could be accidently connected to one another. The resulting circuit can lead to DC and AC level shifts between the driver and receiver ends of the link.

Multiple links have been tested in close proximity to test for noise pick-up. Different shield connections have been investigated and ground loops simulated by the injection of DC and AC signals between the two ends of the link. The links have also been tested with the cables running close to the front-end electronics of a Silicon microstrip module again to evaluate any noise effects.

The work has been carried out as a collaboration between the SCT links group and the TRT links group. The author assisted in the preliminary tests of the cables.

8.2 Test Setup

A custom test board[(a)], shown in figure 8.1, was used to make bit error rate (BER) tests on the STP cables. The board is split into two independent circuits: a driver side, on the left in the photograph; and a receiver side, on the right. The two halves of the board are powered from independent power supplies.

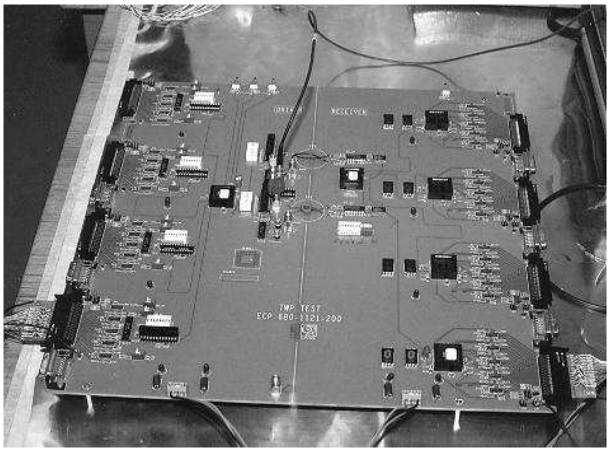


Figure 8.1: Photograph of the BER test board [37].

The board is designed to test bundles made of 27 STP cables connectorised as shown in figure 8.2. Each bundle contains 20 pairs for the transmission of data, 4 dummy pairs for

[(a)] The board was designed by Peter Lichard of the TRT links group at CERN.

the study of cross-talk, and 3 spare pairs. At each end of the bundle the shields of all the pairs are connected together.

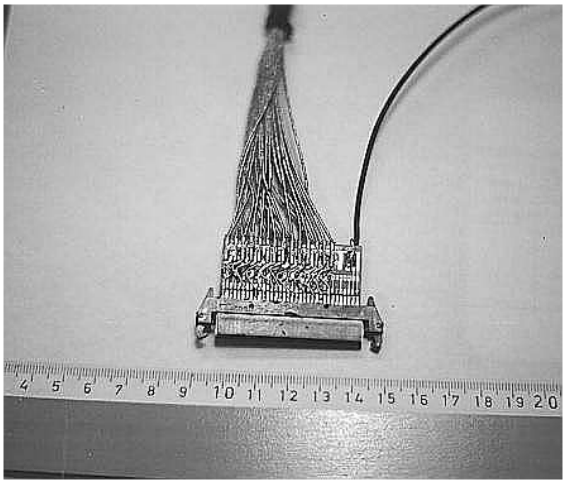


Figure 8.2: Photograph of a 27-pair connectorised AWG36 cable [37].

The driver side of the board can be set to generate either periodic or random TTL data at rates up to $80Mbit/s$. The TTL data are sent to both TTL-LVDS driver chips, which generate identical signals for the 20 data pairs, and to a comparator on the receiver side of the board via a strip line. LVDS-TTL receiver chips are used at the receiver side of the board to convert the LVDS STP data signals to TTL signals which are sent to the comparator. Delays are used to synchronise the data signals and adjust the timing of the comparator strobe. Separate counters are kept for errors on single pairs and simultaneous errors on two or more pairs. At data rates of $40Mbit/s$ BER tests to the required level (10^{-9}) could be performed in a few minutes.

The termination resistors at the driver and receiver sides allow the board to be configured to send signals with the full LVDS swing or with roughly one half of the LVDS swing.

The board allows the cable shields to be terminated to the driver side ground and/or the receiver side ground. The split nature of the board allows both DC and AC signals to be injected between the grounds of the two sides of the board to simulate ground loop scenarios.

8.2.1 System Performance

The first tests were to study the transmission performance of the system. It was evaluated in two ways:

- Eyeplots were made of the received signals, before the LVDS-TTL receiver chips, as shown in figure 8.3. To make an eyeplot the oscilloscope was set to infinite persistence so that multiple signal data are plotted one on top of the other. The eyeplot therefore essentially shows all the possible transitions between low and high data states, eg. 001 or 101. This is important since the 01 transition may be affected by the bit preceding the transition. The opening of the eye is a good indication of the transmission quality. It may become more closed vertically due to a reduction in the signal to noise ratio of the link or due to a lengthening of the rise/fall-time of the transmission so that the 0 and 1 levels are not reached within a one bit time window. Laterally the eye can close due to time jitter of the signal.

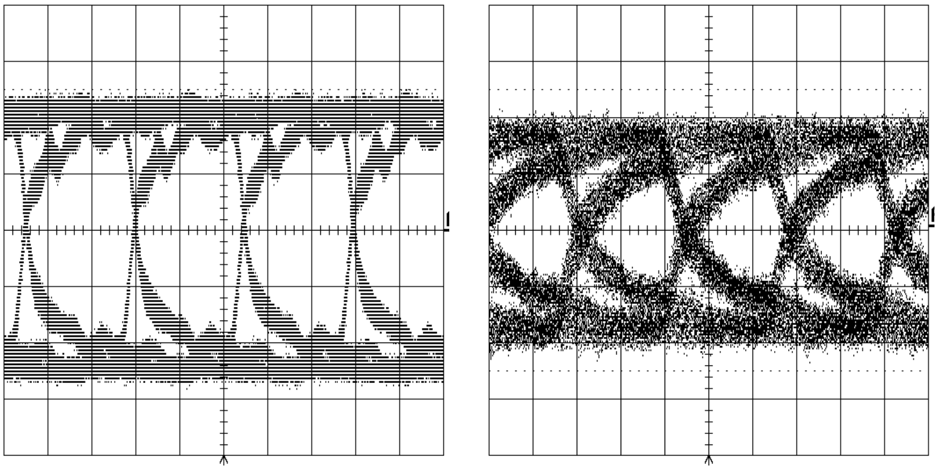
- The timing window of the comparator strobe for which no errors were produced was measured. This is essentially a measure of the opening in the eyeplot.

The transmission performance was measured for different system configurations. Four arrangements for shield connections of the STP cable were studied:

- no shield connection,
- shield connected at driver side,
- shield connected at receiver side,
- shield connected at both sides.

Additionally AC and DC signals of up to 2V were injected between the ground levels of the driver and receiver sides of the board.

Eyeplots for 10 and 20m of AWG36 STP cable are shown in figure 8.3 for data rates of $40 Mbit/s$. It can be seen that the opening of the eyeplot decreases significantly for the longer cable. This is due to the combined effects of a reduced signal amplitude and a decrease in the rise and fall times of the signals due to the larger capacitive load. However, the link



(a) 10m cable. (b) 20m cable.

Figure 8.3: Persistence traces for random data transmission at $40Mbit/s$ over AWG36 STP cable. The x-axis is set at $10ns/div$ and the y-axis $50mV/div$. Each plot corresponds to approximately 2500 sweeps.

still performs adequately with 20m of cable. Little or no difference in the link performance was found due to the shield connections or injection of ground loop signals. Further tests have been carried out by the SCT and TRT links groups [37] with different cable lengths and data transmission rates. The unshielded AWG28 cable for data transmission outside the detector has also been studied. The results are summarised in table 8.2.1. Only those results designated as bad would result in the link being below the performance specification for data readout.

For the module tests the STP cable was run $\sim 4mm$ below the silicon detector and hybrid. The results are discussed in detail in [37]. To summarise it was found that a small increase ($\sim 4.5\%$) in the noise of the module was introduced by the link. However, this noise was found to be due to radiation from the twisted pair shields rather than from the pairs themselves, the shields having picked up signals from the test board. The consequences for a module in the SCT are therefore hard to quantify since any pick up in the cable shields would be from different sources from those in the laboratory tests.

Table 8.1: Summary of twisted pair performance tests [37].

Cable type	Length (m)	Signal size	Data rate ($Mbit/s$)	Performance
AWG36 STP	10	LVDS	40	Excellent
AWG36 STP	10	LVDS/2	40	Excellent
AWG36 STP	10	LVDS	80	Excellent
AWG36 STP	10	LVDS/2	80	Excellent
AWG36 STP	20	LVDS	40	Good
AWG36 STP	20	LVDS/2	40	Good
AWG36 STP	20	LVDS	80	Marginal
AWG36 STP	20	LVDS/2	80	Marginal
AWG28 UTP	20	LVDS	40	Excellent
AWG28 UTP	20	LVDS/2	40	Excellent
AWG28 UTP	30	LVDS	40	Good
AWG28 UTP	30	LVDS/2	40	Good
AWG28 UTP	50	LVDS	40	Marginal
AWG28 UTP	50	LVDS/2	40	Marginal
AWG28 UTP	80	LVDS	40	Bad
AWG28 UTP	80	LVDS/2	40	Bad

8.3 Conclusions

The laboratory tests on the transmission performance of the electrical link indicate that it would perform to the SCT specification. However, there is always a concern that the noise introduced to other systems by the link will scale with the size of the system. The cable tests show that the STP cable could be used for the data transmission within the detector volume at rates up to $80Mbit/s$. A repeater would then be used at the transition between the STP and UTP cables. The UTP cable shows acceptable performance at $40Mbit/s$ for lengths up to $50m$. A further repeater would therefore be required to span the $80m$ distance between the detector and the barracks.

Chapter 9

Conclusions

This chapter summarises the work carried out in this thesis in terms of the requirements for the ATLAS SCT data link as discussed in section 3.4. It also presents work carried out by different authors which is relevant to the choice of the link. Each link is presented in turn. The reason for the baseline choice of link is then discussed followed by future work and avenues for the development of this link.

9.1 Analogue Optical Link

The performance of the analogue optical readout option that is described in chapters 4 and 6 is now reviewed.

9.1.1 Data Transmission

The data transmission performance of the link is discussed in detail in chapter 6. The performance target is a dynamic range of > 250 with a non-linearity $\lesssim 2\%$ for the analogue transmission of data. A signal to noise ratio of $> 20 : 1$ is required to transmit the binary clock, trigger and control data to the modules. From the results taken with the present test setup, the analogue performance target would be expected to be met in the next iteration of modulator packages. These chips would use the old chip design in the new low mass package. However, further work is required on the packaging technique to ensure reproducible results and minimise the spread in characteristics between packages.

A major concern in the system is the noise generated due to reflections. They can increase the noise by the conversion of laser phase noise to amplitude noise (PAN) and by increasing the natural noise of the laser (RIR). The PAN should be kept to acceptable levels by the use of low-reflection connectors. However, the RIR may still cause problems in a final system that does not include an isolator for the laser. More research would be required to show that a system could be made to work within the cost constraints of the SCT.

The use of the modulators as photodiodes for the reception of the clock, trigger and control data has been investigated [30]. The modulators show good performance for this application.

9.1.2 Magnetic Properties

The packaging of the modulators is non-magnetic to allow it to be used in the SCT environment. Furthermore the performance of the device has been measured and modelled in magnetic fields up to $8T$ and shows only a slight change in the modulator characteristics [71].

9.1.3 Material

A calculation of the contribution to the material in the Inner Detector up to the last TRT hit has been made for the two binary link options (see figures 9.2 and 9.3) [37]. The cable routing for each case is the same, as shown in figure 9.1, therefore the distribution of material with η will be the same, but, scaled due to the different link materials and number of links. For the analogue optical link an eight fibre ribbon cable link would run to each module. The material contribution would be less than three times that of the binary option using ribbon fibre (figure 9.2). A peak contribution would therefore be expected of $\sim 1\%$ of a radiation length at $\eta = 2.2$. This is acceptable for use in the SCT.

9.1.4 Power Budget

The modulators are operated as essentially passive devices. The current flow when reverse biased is low. They therefore contribute little to the overall SCT power budget.

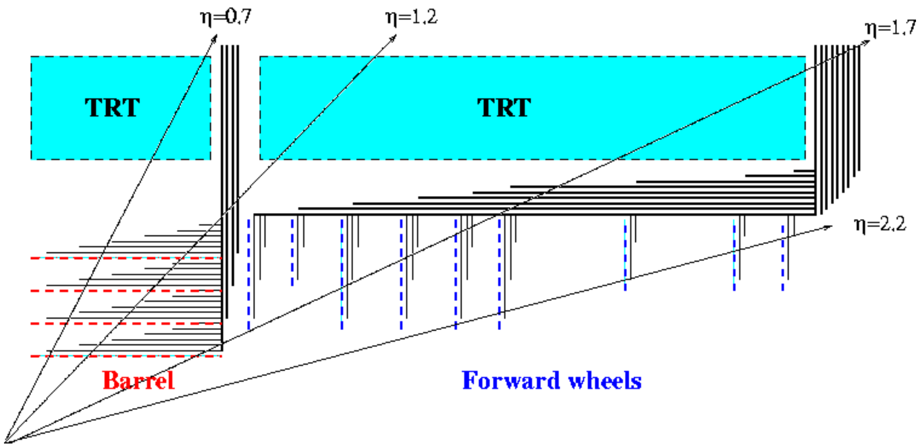


Figure 9.1: The routing of the SCT data and clock, trigger and control cables used for the calculation of the material budget [37].

9.1.5 Reliability

Lifetime tests on the modulators have not been performed and would be required before they could be used in the SCT. However, the structure of the modulators is similar to that of photodiodes and so a similar lifetime may be expected [30]. The median lifetime of GaAlAs photodiodes at room temperature is estimated to be in excess of $10^9 hours$. A MTTF for each modulator channel of $\gtrsim 10^7 hours$ would result in less then 0.2% loss of data.

9.1.6 Radiation Tolerance

Radiation tolerance measurements of the modulators are described in chapter 6. A performance drop of $< 50\%$ would be expected during the ATLAS lifetime of the devices. This performance margin would be required at the start of operation. It has not been included in the previous signal to noise ratio requirement for the link. A requirement for such a performance margin at the start of operation would make the analogue link an unworkable option.

The radiation tolerance of single-mode fibres has been studied [30]. Pure silica core fibres exhibit acceptable characteristics for use in the Inner Detector region of the ATLAS detector.

9.1.7 Thermal Properties

The packaging technology requires further study to reach maturity. At present the devices suffer from temperature dependencies due to inhomogeneities introduced during packaging. The modulators themselves have a temperature dependence due to the shifting of the Fabry-Perot resonance with temperature, however this will not cause problems in the $\pm 2.5^o C$ temperature range of operation in the SCT.

9.1.8 Cost

The main drawback with the passive modulator design is the high cost of the system. In effect an extra component is introduced in addition to the normal transmitter and receiver elements. Furthermore the noise considerations of the link require high specification low-reflection connectors throughout the system. The system cost far exceeds the SCT budget by a factor three.

9.2 Binary Optical Link

The performance of the binary optical readout option that has been described in chapters 4 and 7 is now reviewed.

9.2.1 Data Transmission

The transmission requirements for a binary link are much less harsh than those for an analogue link. The binary optical link has been shown to have a huge performance margin before exposure to radiation for the transmission of the front-end chip data [20]. The degradation due to radiation will be discussed below. The transmission of the clock, trigger and control data has not been fully investigated. However, no problems are envisaged since the performance is largely controlled by the optical power of the transmitter, which will sit in a non-radiation environment and therefore not deteriorate.

9.2.2 Magnetic Properties

The optical package is non-magnetic.

9.2.3 Material

The material contribution to the Inner Detector up to the last TRT hit is shown in figure 9.2. The most likely option for fibre protection is the use of fibre cables. This results in a material contribution peak of 0.35% of a radiation length at $\eta = 2.2$. This is well below the limit set for use in the SCT.

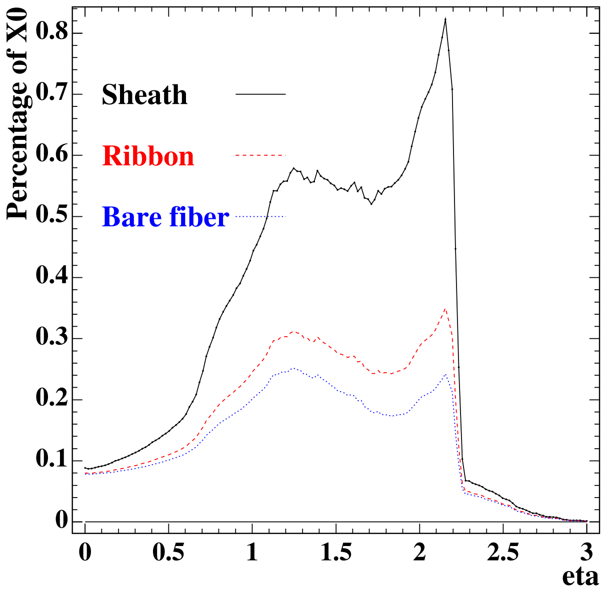


Figure 9.2: Binary optical link system material contribution for different fibre protection options [37].

9.2.4 Power Budget

The power consumption of the optical binary link comes from two sources. Each LED and associated driver circuitry contributes a power $\sim 20mW$[a] during normal running. During an annealing cycle at $50mA$ forward bias the power consumption would be $\sim 200mW$ per LED. The second source consumption is the DORIC receiver chip. Its power consumption is expected to be $\sim 200mW$.

During normal operation the optical system will therefore have a power consumption of $\sim 0.16mW$ per SCT channel. The total power consumption for the binary based SCT

[a] This is based on the assumptions of: a) a 50% link occupancy, b) the LEDs being off when no data is transmitted, and c) an equal number of 0 and 1 bits in the data.

utilising the optical readout option has been calculated to be $3.0-3.2mW$ per SCT channel [20]. The optical readout contributes only a small part.

During an annealing cycle the power consumption of the optical readout would increase to $\sim 0.4mW$ per SCT channel. The total power consumption of the SCT would therefore still fall within the power budget of $3.8mW$ per SCT channel.

9.2.5 Reliability

Comprehensive lifetime tests have been performed on the LEDs that are to be used in the optical package [55, 56]. The tests have been performed on irradiated diodes due to a worry that defects caused by radiation may increase the failure rate of the LEDs. The lifetime tests are accelerated by running the LEDs in a continuous current mode[(b)] and at an increased temperature[(c)] than that of the SCT. The tests show that a failure rate $\sim 5\%$ could be expected over the lifetime of the SCT. The redundancy scheme described in chapter 4 is therefore mandatory.

As mentioned in the analogue optical link review section above, the lifetime of photodiodes is not expected to be a problem. However, large sample tests of photodiodes are planned at Birmingham utilising the Bern group scanning machine [55, 56] during the next year.

9.2.6 Radiation Tolerance

The radiation tolerance of the LEDs and photodiodes used in the optical package is discussed in chapter 7. In addition the radiation tolerance of the LEDs has been studied for large samples of devices by the Bern group [56].

The LEDs show a substantial degradation due to exposure to both neutron and proton radiation. Damage constants have been found for the degradation. The ratio between the constants for proton and neutron damage compare well with that predicted by assuming a NIEL scaling between the radiation types. The damage constants have been used to calculate

---

[(b)] An acceleration factor of ~ 12 relative to LHC is estimated based on the assumptions that: a)LHC will run for 100 days a year, b) the average link occupancy is about 50%, c) the LEDs are off when no data is transmitted, and d) the 0 and 1 bits are balanced.

[(c)] The dependence of the mean lifetime of an LED on operating temperature and current is given by the Arrhenius equation. A conservative acceleration factor of 10 is used in these tests.

the expected link performance at the end of the SCT lifetime. The performance is marginal. However, it is found that the link has a healthy safety margin of operation if the beneficial effects of injection annealing are included in the estimate. High current annealing cycles will therefore be required throughout the operational lifetime of the links.

Two types of photodiodes have been tested for use in the package. The radiation tolerance of the bulk Silicon diodes is poor, however the epitaxially grown diodes have shown a good tolerance to neutron irradiation. The damage constant calculated for the radiation induced increase in dark current predicts a $\lesssim 16nA$[d] current after the expected integrated fluence at the 1^{st} SCT barrel. An initial $\lesssim 30\%$ drop in the response of the diodes was observed. The diode response were subsequently unaffected by $1MeV$ equivalent neutron fluences up to $\sim 2.7 \times 10^{14} cm^{-2}$. Assuming a response of $0.2A/W$ at the end of ATLAS operation and an input optical power of $10\mu W$ at the diodes (easily obtainable), the link would have a signal to noise ratio of $\sim 125 : 1$. This appears to no pose problems for the transmission of the clock, trigger and control data. However, irradiation tests with protons (which are in progress) and measurements of the rise and fall-times of the diodes are required before we can be fully confident that the diodes are suitable for use in the SCT and Pixel Detector.

The radiation tolerance of multi-mode fibres has been tested [72]. It is again found that pure silica core fibres show an acceptable behaviour.

9.2.7 Thermal Properties

Measurements show that the LEDs have a well behaved output power temperature response. An increase is power of $1.0 \pm 0.2\%/^oC$ has been measured [56]. Observations made during the proton irradiation of the LEDs (see chapter 7) agree with this value. For binary operation a temperature change of $\pm 2.5^oC$ will not affect the link performance.

The thermal properties of the photodiodes have not been measured. A decrease in temperature is expected to result in a slight drop in the response of the diodes due to a drop in the absorption coefficient of Silicon at $820nm$. However, for the temperature change in question this effect would be small. The decrease in the dark current due to the reduction

[d] The prediction is made from measurements at $\sim +17^oC$. The dark current at -10^oC would be smaller.

in temperature will counteract this change. The temperature characteristics of the diodes requires further examination and will be necessary during the lifetime test.

9.2.8 Cost

The currently estimated cost of the complete binary optical link including the front-end LDC and DORIC, the back-end receiver module and the fibres is 435CHF. This is 20% higher than the SCT budget.

9.3 Binary Electrical Link

The performance of the binary electrical readout option that is described in chapters 4 and 8 is now reviewed.

9.3.1 Data Transmission

The binary electrical link has been tested on a small scale involving bundles of 27 cables as described in chapter 8. Tests have been performed to simulate the expected noise contributions to the link. The electromagnetic radiation to other cables has also been measured as well as that to and from an SCT module. These tests indicate that the system would function acceptably in the SCT. However, there is a worry that noise contributions could escalate when the system is expanded to that required for readout of the full SCT. The system could therefore only really be tested fully after the construction of the detector.

9.3.2 Magnetic Properties

The front-end chips are designed to drive and receive LVDS signals. However, drivers and receivers may be required for the STP cables on the hybrid to match the impedances of the cables and hybrid transmission lines. These chips could be made with similar technology to the front-end chips and so would not be expected to cause problems.

9.3.3 Material

The material contribution to the Inner Detector up to the last TRT hit is shown in figure 9.3. Three cabling options are shown:

- A link consisting of four AWG36 STP cables. Two are used for the front-end data transmission from the module, one for the clock transmission to the module, and one for the trigger and control signal transmission to the module. This option has been described in chapter 4;
- A link consisting of three AWG36 STP cables. Two are used for the front-end data transmission from the module. However, an encoding scheme similar to that for the binary optical link would be used to send the clock, control and trigger signals down a single cable;
- A link consisting of three AWG38 STP cables. The AWG38 STP cables are similar to, but smaller than, the AWG36 STP cables. The transmission scheme would use three cables. The AWG38 STP cables have not been tested fully for their data transmission qualities.

The contribution for all the cabling options is much higher than that for the optical options. Reduction of the cabling material by either removing the requirement for one of the links or using smaller cables is required for the electrical link to be an option for SCT readout.

9.3.4 Power Budget

The electrical link would only contribute to the power budget if front-end driver and receiver chips were required, or data encoding was used facilitating the use of a DORIC-like decoder chip. In either case the contribution would be similar to that of the binary optical link and therefore acceptable for use in the SCT.

9.3.5 Reliability

In general integrated circuits have MTTF $\sim 10^9 hours$. They would therefore be expected to last the lifetime of the SCT.

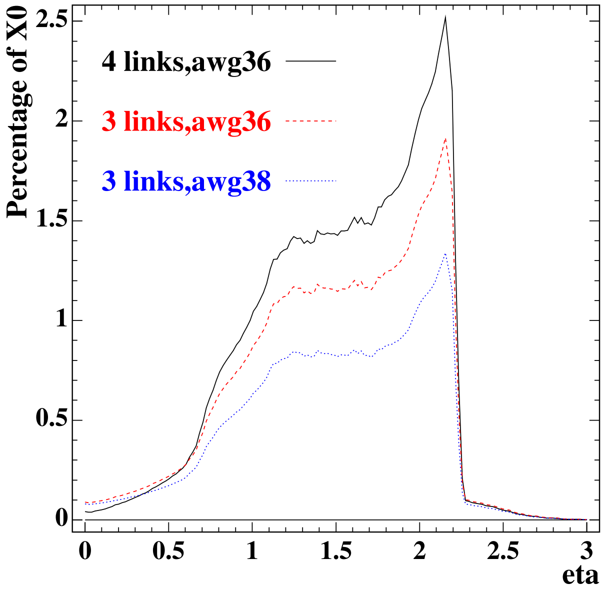


Figure 9.3: Binary electrical link system material contribution for different STP options [37].

9.3.6 Radiation Tolerance

A radiation hard repeater will be required at the patch panel between the STP and UTP cables. However, the radiation tolerance required here is much lower than in the Inner Detector. If radiation hard drivers and receivers were required on the hybrid similar designs to the LVDS driver and receiver circuits in front-end chips could be used.

9.3.7 Thermal Properties

The properties of the cables do not change greatly with temperature.

9.3.8 Cost

The binary electrical link offers a substantial cost saving over the binary optical link. The cost is estimated to be < 143CHF per module for the AWG36 STP four cable option. The cost saving of the cable would balance that for the encoders and decoders if a three cable link was used.

9.4 The Past

Various milestones have been encountered throughout the time-scale of this thesis. The milestones have been set to focus the work of the community onto the options that show the greatest promise for success. A choice between the binary and analogue readout options was made by the ATLAS collaboration in March 1996. The choice was made to pursue the binary option due to a large cost disparity between the two solutions and the more mature technological solutions of the binary readout, in particular the front-end chips. The analogue link contributed highly to the cost of the analogue readout option.

The choice of the baseline binary readout technology has been made quite recently in May 1997. The optical system was chosen. This decision was largely motivated by the material and noise advantages offered by this system. The predicted material in the SCT has been rising steadily as the module and support structure designs have been better investigated. The SCT is now very close to the limit of detrimentally affecting the rest of the ATLAS detector. For this reason any saving of material in the Inner Detector is greatly valued. The worries of noise in the electrical link were a second motivation. Although the electrical system offers a substantial cost saving the advantages of the optical system were deemed worth the investment.

9.5 The Future

Although the binary optical link has been chosen as the baseline choice there are still numerous tests required before it offers a complete solution. Much of the future work will focus on the clock, trigger and control link, in particular the testing of the photodiode.

However, there is an alternative transmitter technology that has been investigated for high radiation environment applications and shows considerable promise [56,72,73]. Vertical surface emitting laser diodes (VCSELs) will therefore be investigated as a replacement for LEDs in the optical package. Laser diodes offer the advantage of high radiation hardness. The reduction of the non-radiative minority carrier lifetime (see section 7.1.1) has the effect of increasing the threshold current required for the devices to lase, but has little effect on the output power which is controlled by stimulated emission. VCSELs can be fabricated in arrays, unlike end emitting lasers, and therefore offer a distinct cost saving. The main problems with VCSELs to date has been the coupling of the devices into fibres, causing modal noise problems. Early indications show that this may be solved by the butt-coupling of the fibre that is used in the present optical package.

Appendix A

Radiation Testing

A.1 Neutron irradiation

The University of Birmingham Dynamitron was used for the neutron tests. Neutrons are extracted via the reaction ^{9}Be(d,n)^{10}B, that is, the stripping of deuterons on a thick Beryllium target. The Dynamitron is a variable energy potential drop machine capable of terminal potentials up to $3MV$, and is powered by a radio-frequency oscillator operating at about $130kHz$. For the irradiation tests, a deuteron beam energy of $2.6MeV$ and a current of $200\mu A$ was used. This gives a neutron energy centred at around 2MeV [74] with a spread of approximately $2MeV$, and a yield of $\sim 10^{11} sr^{-1} s^{-1}$.

The operator shift rota at the Dynamitron results in a daily cycle of $\sim 8hrs$ beam in the daytime followed by $\sim 16hrs$ with no beam overnight. Fluences upto $\sim 10^{14} cm^{-2} day^{-1}$ can be achieved at small distances.

Activation foils were used to measure the total neutron fluences. Measuring the gamma activity of the foils after irradiation allows the flux to be determined. For the neutron tests, indium foils were cut to sizes corresponding to the components to be irradiated and positioned nearby. The reaction ^{115}In(n,n$'$)^{115m}In, with its low threshold of $0.339MeV$ and its product with a 4.486 hour half-life and 0.459 gamma rays per meta-stable decay was suitable for these tests. After irradiation, the foils were placed in a high precision Germanium detector (cooled to $77K$) and the gamma activity measured. Typically errors of $\sim 10\%$ were achieved in the activity measurement which is limited by the time of the measurement. This is the dominant error on the fluence measurement.

A.2 Proton irradiation

The proton irradiation studies used the CERN Proton Synchrotron (PS), Geneva. The proton energy is 24GeV and fluences of $10^{14} cm^{-2} day^{-1}$ are easily obtainable. (Typically there are bunches of 10^{11} protons every 14 seconds; the spread is dependent on the beam optics but an area 1cm $\times$ 1cm is achievable). In these irradiation tests the devices under test were mounted on an x-y stage (designed for the irradiation of silicon detectors) which scans an area of $6cm \times 6cm$ transverse to the beam direction. The PS is operational 24 hours a day. Consequently, expected fluences are of the order of $10^{13} cm^{-2} day^{-1}$.

For the proton tests, aluminium foils were used for an initial calibration run. The fluences obtained, via the reaction ^{27}Al(p,3pn)^{24}Na (producing a $1.368 MeV$ gamma with a $\sim$ 15 hour half-life), were used to calibrate a Secondary Emission Counter. The activity of the aluminium foils was measured in the same way as for the indium foils. The errors on the activity measurements were of comparable size.

Glossary

There follows a glossary of the abbreviations used in this thesis:

AC Alternating Current

ADC Analogue to Digital Converter

ALICE A Large Ion Collider Experiment

APC Angled Physical Contact connector

APV5 Analogue Pipeline Version 5

ATLAS A Large Toroidal Apparatus

BER Bit Error Rate

CERN European Laboratory for Particle Physics

CMS Compact Muon Solenoid

CSC Cathode Strip Chamber

CW Continuous Wave

DAC Digital to Analogue Converter

DAQ Data Acquisition

DC Direct Current

DFB Distributed Feedback

DORIC Digital Optical Receiver Integrated Circuit

FBT Fused Biconical Taper

FC Fibre Connector

FP Fabry-Perot

FWHM Full Width at Half Maximum

GPIB General Purpose Interface Bus

GUT Grand Unified Theory

LAr Liquid Argon

LDC LED Driver Circuit

LED Light Emitting Diode

LEP Large Electron Positron Collider

LHC Large Hadron Collider

LVDS Low Voltage Differential Signals

MDT Monitored Drift Tube

MIP Minimum Ionising Particle

MOCVD Metal Organic Chemical Vapour Deposition

MQW Multi Quantum Well

MSSM Minimum Supersymmetric Standard Model

MT Multi-tip

MTTF Mean Time To Failure

MUX Multiplex

NIEL Non-Ionising Energy Loss

NRF Noise Reduction Factor

PAN Phase to Amplitude Noise

PC Physical Contact connector

PKA Primary Knock-on Atom

PS Proton Synchrotron

QCD Quantum Chromo-Dynamics

QED Quantum Electro-Dynamics

QGP Quark Gluon Plasma

RF Radio Frequency

RoI Region of Interest

RIN Relative Intensity Noise

RIR Reflection Induced RIN

RMS Root Mean Squared

ROC Readout Controller

RPC Resistive Plate Chamber

SCT Semiconductor Central Tracker

SM Standard Model

S/N Serial Number

SNR Signal to Noise Ratio

SNR$_{1V}$ Signal to Noise Ratio for 1V input signal

SPS Super Proton Synchrotron

ST Straight Tip connector

STP Shielded Twisted Pair

SUSY Super Symmetry

SWP Spatial Walkoff Polariser

TGC Thin Gap Chamber

TRT Transition Radiation Tracker

TTL Transistor Transistor Logic

UTP Unshielded Twisted Pair

VCSEL Vertical Cavity Surface Emitting Laser Diode

References

[1] F. Halzen and A.D. Martin, *Quarks and Leptons : An Introductory Course in Modern Particle Physics*, John Wiley & Sons, 1984.

[2] P.A.M.Dirac, *Proc. Roy. Soc* **A117**, 610 (1928), and also *Principles of Quantum Mechanics*, Oxford Press, 1947.

[3] R.P. Feynman, *Quantum Electrodynamics*, Benjamin, 1961.

[4] D.H.Perkins, *Introduction to High Energy Physics*, Addison-Wesley Publishing Company, Inc., 1987.

[5] W.Marciano and H.Pagels, *Phys. Reports* **36C**, 137 (1978).

[6] S.L.Glashow, *Nucl. Phys.* **22**, 579 (1961).

[7] S.Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967).

[8] A.Salam, *Elementary Particle Theory* (edited by N.Svartholm) Almqvist and Wikell, Stockholm, 1968.

[9] M.Kobayashi and K.Maskawa, *Prog. Theor. Phys.* **49**, 282 (1972).

[10] P.W.Higgs, *Phys. Rev. Lett.* **13**, 508 (1964), and *Phys. Rev.* **145**, 1156 (1964).

[11] M.'Quirós, "Bounds on the Higgs mass in the Standard Model and Minimal Supersymmetric Standard Model", *CERN Report* **CERN-TH.7507/94** (1994).

[12] H.Georgi and S.L.Glashow, *Phys. Rev. Lett.* **32**, 487 (1974).

[13] C.W.Fabjan, "LHC: Physics, Machine, Experiments", *CERN Report* **CERN-PPE/95-25** (1995).

[14] "Design Study of the Large Hadron Collider (LHC)", *CERN Report* **CERN/91-03** (1991).

[15] J.H.Christenson, J.Cronin, V.Fitch and R.Turlay, *Phys. Rev. Lett.* **13**, 138 (1964).

[16] W.W.Armstrong *et al.*, "ATLAS : Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN", *CERN Report* **CERN/LHCC/94-43** (1994).

[17] G.L.Bayatian *et al.*, "CMS : Technical Proposal", *CERN Report* **CERN/LHCC/94-38** (1994).

[18] ALICE Collaboration, "ALICE : Technical Proposal for A Large Ion Collider Experiment at the CERN LHC", *CERN Report* **CERN/LHCC/95-71** (1995).

[19] LHC-B Collaboration, "LHC-B : Letter of Intent for a Dedicated LHC Collider Beauty Experiment for Precision Measurements of CP-Violation", *CERN Report* **CERN/LHCC 95-5** (1995).

[20] ATLAS Inner Detector Community, "ATLAS Inner Detector Technical Design Report", *CERN Report* **CERN/LHCC/97-16** (1997).

[21] ATLAS Muon Collaboration, "ATLAS Muon Spectrometer Technical Design Report", *CERN Report* **CERN/LHCC/97-22** (1997).

[22] S.Gadomski *et al.*, "The Deconvolution Method of Fast Pulse Shaping at Hadron Colliders", *CERN Report* **CERN-PPE/92-24** (1992).

[23] G.P.Summers *et al.*, *IEEE Trans. Nucl. Sci.* **NS-34**, 1134 (1987)

[24] T.F.Luera *et al.*, *IEEE Trans. Nucl. Sci.* **NS-34**, 1557 (1987)

[25] C.J.Dale *et al.*, *IEEE Trans. Nucl. Sci.* **NS-35**, 1208 (1988)

[26] G.P.Summers *et al.*, *IEEE Trans. Nucl. Sci.* **NS-35**, 1221 (1988)

[27] G.P.Summers *et al.*, *IEEE Trans. Nucl. Sci.* **NS-40**, 1372 (1993)

[28] A.L.Barry *et al.*, *IEEE Trans. Nucl. Sci.* **NS-42**, 2104 (1993).

[29] A.Chilingarov, J.S.Meyer and T.Sloan, "Radiation Damage due to NIEL in GaAs Particle Detectors", *Atlas Internal Note* **INDET-NO-134** (1996).

[30] RD 23 Collaboration, "Optoelectronic Analogue Signal Transfer for LHC Detectors", *CERN Report* **CERN/LHCC 95-61** (1995), and "Modulator based lightwave links for analogue signal transfer at LHC", *Proceedings of the First Workshop on Electronics for the LHC Experiments, Lisbon, 1995*, *CERN Report* **CERN/LHCC/95-56** (1995).

[31] C.Weisbuch and B.Vinter, *Quantum Semiconductor Structures'*, Academic Press, 1991.

[32] G.I.Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*, Halsted Press, 1988.

[33] S.Schmitt-Rink, D.S.Chelma and D.A.B.Miller, Advances in Physics **38**, 89 (1989).

[34] J.Thompson *et al.*, *Journal of Crystal Growth* **107**, 860 (1991).

[35] N.Green *et al.*, "Optoelectronic Detector Readout for Large Hadron Collider", *MDS Report* **SLM8709** (1992).

[36] C.B.Brooks *et al.*, "Optical Links for the ATLAS SCT", *Proceedings of the First Workshop on Electronics for the LHC Experiments, Lisbon, 1995*, *CERN Report* **CERN/LHCC/95-56** (1995).

[37] R.Kowalewski *et al.*, "Evaluation of twisted pair data transmission for the SCT", *ATLAS Internal Note* **INDET-NO-172** (1997).

[38] A.Yariv, *Optical Electronics, 4th edition*, Saunders College Publishing, 1991.

[39] S.M.Sze, *Physics of Semiconductor Devices, 2nd edition*, John Wiley & Sons, 1981.

[40] J.M.Senior, *Optical Fiber Communications, Principles and Practice, 2nd edition*, Prentice Hall International, 1992.

[41] H.C.Lefevre, *Elect. Lett.* **16**, 778 (1980).

[42] T.Matsumoto, *Trans. IECE Japan* **62**, 516 (1979).

[43] K.Chang and W.Sorin, *Opt. Lett* **15**, 449 (1990).

[44] K.Sharaishi, *Elect. Lett.* **27**, 302 (1991).

[45] K.Sato, *IEEE Journal of Quantum Electronics* **QE-19**, 1380 (1983).

[46] M.M.Choy *et al.*, *Elect. Lett.* **23**, 1151 (1987).

[47] R.W.Tkach and A.R.Chraplyvy, *Journal of Lightwave Technology* **LT-4**, 1711 (1986).

[48] P.K.Pepeljugoski and K.Y.Lau, *Journal of Lightwave Technology* **10**, 957 (1992).

[49] T.Okoshi, K.Kikuchi and A.Nakayama, *Elect. Lett.* **16**, 630 (1980).

[50] I.Dawson *et al.*, *Nucl. Inst. and Meth. in Phys. Res.* **A387**, 369 (1997).

[51] *private communication with* F.Vasey, ECP Division, CERN.

[52] I.Dawson *et al.*, "MQW-Modulator Based Optical Links for Atlas Tracking Detectors", *Atlas Internal Note* **INDET-NO-119** (1995).

[53] J.Beringer *et al.*, "ATLAS Silicon Strip Beam Test Results", *Proceedings of the Second International Symposium on the Development and Application of Semiconductor Tracking Detectors, Hiroshima, 1995.*

[54] M.French *et al.*, "APV5RH: A 128 channel radiation hard pipeline chip for LHC Tracker applications", *Proceedings of the First Workshop on Electronics for the LHC Experiments, Lisbon, 1995*, *CERN Report* **CERN/LHCC/95-56** (1995).

[55] J.Beringer *et al.*, "A Life Time Test of Neutron Irradiated Light Emitting Diodes", *ATLAS Internal Note* **INDET-NO-105** (1995).

[56] J.Beringer *et al.*, "Radiation Hardness and Lifetime Studies of LEDs and VCSELs for the Optical Readout of the ATLAS SCT", *ATLAS Internal Note* **INDET-NO-183** (1997).

[57] A.H.Kalma and W.H.Hardwick, *IEEE Trans. Nucl. Sci.* **NS-25**, 1483 (1978).

[58] M.Edwards and D.R.Perry, "The Radiation Hardness Test Facility", *Rutherford Appleton Laboratory Report* **RAL-90-065**.

[59] J.Lindhard *et al.*, "Range concepts and heavy ion ranges", *Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* **33**, 2 (1963).

[60] J.Lindhard *et al.*, "Integral Equations Governing Radiation Effects", *Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* **33**, 3 (1963).

[61] B.H.Rose and C.E.Barnes, *J. Appl. Phys.* **53**, 1772 (1982).

[62] A.L.Barry, R.Maxseiner and D.Bräunig, *IEEE Trans. Nucl. Sci.* **NS-37**, 1726 (1990).

[63] G.C.Messenger, *IEEE Trans. Nucl. Sci.* **39**, 468 (1992).

[64] H.W.Kraner, *Nucl. Inst. and Meth.* **225**, 615 (1984).

[65] A.L.Barry, R.Wojcik and A.L.MacDiarmid, *IEEE Trans. Nucl. Sci.* **NS-36**, 2400 (1989).

[66] L.W.Aukerman *et al.*, *J. Appl. Phys.* **34**, 3590 (1963).

[67] D.V.Lang and L.C.Kimerling, *Phys. Rev. Lett.* **33**, 489 (1974).

[68] P.J.Drevinsky *et al.*, *IEEE Trans. Nucl. Sci.* **NS-41**, 1913 (1994).

[69] D.V.Lang, *J. Appl. Phys.* **45**, 3023 (1974).

[70] N.Tamura *et al.*, *Nucl. Inst. and Meth. in Phys. Res.* **A342**, 131 (1994).

[71] K.A.Webster, "Investigation of the Use of Optical Modulators for Analogue Data Readout from Particle Physics Detectors", *PhD Thesis*, University of Birmingham 1995.

[72] B.Dinkespiler *et al.*, "Neutron Irradiation of Optical Link Components", *Atlas Internal Note* **LARG-NO-002** (1994).

[73] B.Dinkespiler *et al.*, "Analogue Optical Links for the Liquid Argon Calorimeters", *Proceedings of the First Workshop on Electronics for the LHC Experiments, Lisbon, 1995*, *CERN Report* **CERN/LHCC/95-56** (1995).

[74] J.W.Meadows, *Nucl. Inst. and Meth. in Phys. Res.* **A324**, 239 (1993).