// Universität Bonn
Physikalisches Institut

Noise and Performance Measurements on the BaBar Silicon Vertex Tracker Readout Prototype (AToM) Chip.

By
Armin Karcher

Abstract:

In this thesis calibration and performance measurements performed on a custom integrated circuit for readout of the BaBar Silicon Vertex Tracker (SVT) subsystem are described [1]. This prototype chip was produced using a radiation-soft process, but contains the full functionality of the chip that will be used in the BaBar experiment. It is commonly referred to as the AToM (A Time Over threshold Machine) chip [2].

The thesis describes the chip, its properties and its environment, consisting of both the SVT subsystem and the Data Acquisition (DAQ) system.

Introduction:

Modern high energy physics detectors are composed of multiple subsystems, each of which has a specialized task in determining the properties of the observed particle decay.

The design of the entire detector varies greatly with the kind of accelerator at which it is placed, and the objectives of the experiment. The most obvious distinction is the difference between fixed target experiments and collider experiments. Most modern collider detectors have a cylindrical symmetry around the beam, while fixed target detectors are built to fit the specific geometry of each experiment. When looking at collider detectors one has to differentiate between hadron and electron collider detectors. Hadron colliders can reach higher energies due to the heavier particles at the cost of higher background rates. In an electron positron collider the colliding particles annihilate into pure energy, creating a well defined initial state for the following particle production. In hadron colliders only 2 out of 6 quarks of the two hadrons involved in the collision annihilate, the other 4 quarks produce massive hadronic background.

In modern colliders the size of the beam pipe is minimized both in diameter and wall thickness. Such a beam pipe allows a better decay vertex reconstruction, because the decay particles are deflected by multiple scattering in the beampipe. In a narrower beampipe the point of deflection is closer to the vertex, therefore the error in the vertex position is smaller. A thin walled beampipe reduces the multiple scattering error even further by reducing the average deflection angle.

To take advantage of the narrow beampipe the detector closest to the beam has to have very high spatial resolution. In most modern collider experiments this detector is a silicon microstrip vertex detector (see chapter 2).

These detectors have an extremely high number of channels in a very small volume. Typical readout pitches are on the order of 100μm. This makes readout by conventional discrete electronics impossible. Some experiments designed their own custom readout integrated circuits optimized for their specific operating environment, in other experiments pre-existing designs were used. Earlier designs had a switched capacitor array to store analog voltages, so that all channels on one chip could be read out via a single analog bus [3]. Many recent designs have a fully digital readout, converting the analog data on the front end readout chips. Each new design did not only reflect the objectives of the experiment, but also the rapid advances in semiconductor technology.

The location of the detector close to the beam pipe does potentially put the readout electronics in a high radiation environment. Depending on the collider intensity this may require expensive radiation hardened processes for the fabrication of the readout chips.

The more advanced and complex the circuits are, the harder it is to verify the entire functionality using simulation models. Therefore several iterations of prototypes are built, first to prove the feasibility of the design, then to check the functionality and performance of the actual implementation. The findings on the radiation soft prototypes are then used to optimize the design and fix any problems before the first radiation hardened prototype is fabricated.

The high integration of the chips makes it impossible to measure vital parameters directly: the output noise of the preamplifier, for example, can only be inferred by measurements on the chips output, or by probing on test points (picoprobing) within the chip. Picoprobing is time consuming, and only a few channels can be measured simultaneously. Thus comprehensive readout and data processing is necessary to measure inaccessible chip parameters.

The large number of channels on each chip means that testing even a single chip produces enormous amounts of raw data. Therefore a sophisticated DAQ (Data Acquisition) system is needed even in the early testing stages. The software has to be powerful enough to make data visualization manageable, so that problems can be easily detected.

The centerpiece of the BaBar detector [4] is the Silicon Vertex Tracker (SVT), a silicon microstrip detector with the ability to reconstruct particle tracks without information from the other detector subsystems. A 128 channel full custom integrated circuit (IC) for the amplification, shaping, digitalization and readout of the detector will be described. In this thesis functionality tests, as well as noise and performance measurements performed on the SVT readout chip prototype will be described. The test stand, including both the data acquisition software and hardware, as well as the chip itself, will be discussed in detail.

1. Goal of This Thesis:

In this thesis both the environment of the SVT readout chip and the custom integrated circuit itself will be described, as far as this is necessary to understand the various performance measurements which were done during the course of this thesis. The description of the SVT detector subsystem is at sufficient depth to illuminate underlying performance requirements for the AToM chip [2].

An extensive set of measurements was performed, first testing the functionality and performance of the chip, then measuring system noise and other chip specific parameters. In this thesis the result from these measurements as well as the methods and apparatus used to make them will be described.

During the course of the thesis there will be a clear distinction between components and procedures that will be used in the BaBar experiment, and preliminary prototypes. The differences in functionality will be explained so that the reader will become familiar with both the test setup and the chip's intended environment in the experiment.

All of the custom readout electronics that were used during the chip testing will be described extensively, as this introduction is needed to understand the procedures used for the measurements and the various technical hurdles. These electronic components are prototypes of the components that will be used in the experiment. The functionality is sometimes very similar, or identical (as in case of the chip itself), and sometimes additional requirements arose between design of the prototype and completion of this thesis.

The software package that was written and used to test the chip has been written specifically for the existing data acquisition (DAQ) system with the readout chip testing in mind. It will be used in various stages of testing (e.g. subsystem test of the SVT system in early 97), but will not be incorporated in the final data-flow software for the BaBar experiment because the software framework for the BaBar DAQ system has not yet been completed.

The main objective of this thesis will be to discuss the test apparatus and the performance measurements on the AToM chip. This prototype of the SVT readout chip has the full functionality of the final chip. At most two more sets of prototypes will be manufactured before the final chip is commissioned. The measurements done for this thesis are the most comprehensive tests done to date on the full chip prototype. Individual parts of the measurements have been repeated elsewhere, and their results will be cited where applicable.

In the next chapter a summary of the operating principles of silicon vertex detectors in general will be given. The chapter will discuss spatial resolution limits, system noise components and detector speed, as well as basic readout principles.

Chapter three will concentrate on the specific features of the Silicon Vertex Tracker and the readout system. The description includes many details of the final detector that are relevant to this thesis only in terms of explaining requirements and design decisions.

In chapter four a short overview over all measurements, that were taken during the course of this thesis, will be given.

In chapter five the test bench setup that was used to obtain the measurements will be discussed in detail. Both software and hardware components in whose design, realization, or debugging I had a major responsibility will be treated with particular detail.

In chapter six the set of measurements performed during this thesis is described, the results are evaluated, plotted and discussed.

In chapter seven the results of the measurements are compared with other measurements, simulations and requirements.

2. Silicon Microstrip Detectors

To understand the design issues involved in the design of the SVT readout chip one has to be aware of the operating principles of the SVT. To give the reader a better understanding of those principles this chapter gives a general overview of the theory of Si-microstrip-detectors.

2.1. Semiconductor Detectors

Si-microstrip detectors used in high energy physics detectors are a special case of a large class of devices called semiconductor detectors. Other popular examples are the Charge Coupled Device (CCD) found in many video cameras or astronomical telescopes, or the photo diode which is used anywhere from domestic security systems to CD players. Another well-known semiconductor detector is the Ge pin diode, which is most often used in high resolution spectroscopy.

2.1.1. Semiconductors and P/N Junctions

Semiconductors are column IV elements with a relatively small band gap between the valence band and the conduction band. In the case of silicon this gap is 1.1 eV [5]. At room temperature a pure semiconductor has essentially no free charge carriers, and is therefore an insulator. By adding small amounts of impurity atoms the electrical behavior of the semiconductor can be controlled. When adding column V impurities (such as phosphorus) each impurity atom introduces an electron as essentially free charge carrier. The binding energy of those electrons to their atoms is only about 0.05 eV. On

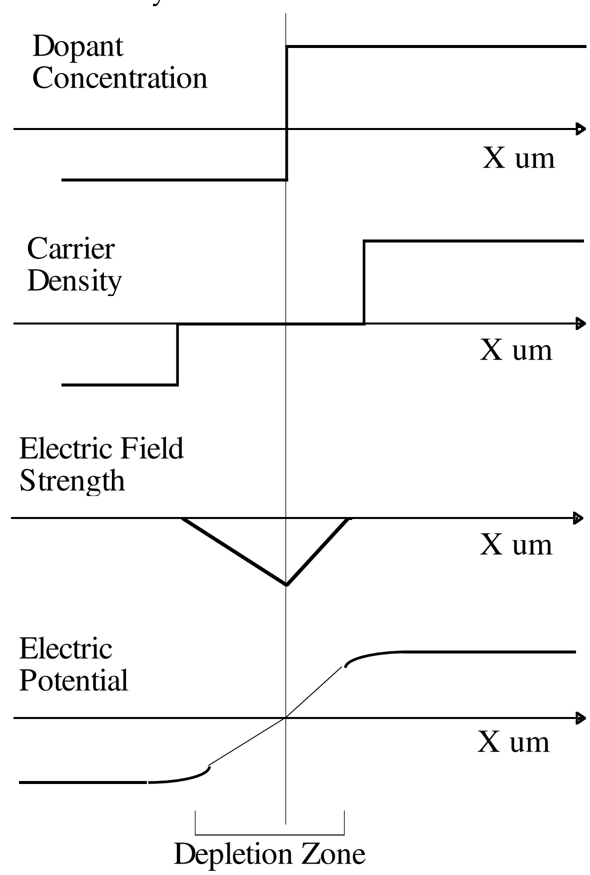


Figure 1 : Various quantities across an unbiased np junction.

the other hand column III impurities can accept additional electrons, thus creating holes as effective positive charge carriers. Usual semiconductors have dopant concentrations anywhere between $10^{13}/cm^3$ to $10^{18}/cm^3$.

When a part of the semiconductor with column V impurities (n-doped) is in direct contact with a part with column III impurities (p-doped) the contact area is called a N/P junction. The absence of electrons on one side, and holes on the other causes a diffusion current, electrons drift across the boundary and combine with holes and vice versa. This process leaves the junction area clear of free charge carriers, but slowly builds a space charge, which in turn creates an electric field. The diffusion process reaches equilibrium when the electric field is strong enough to create an equal but opposite flow of charge carriers. If the n-dopant concentration is lower than the p-dopant concentration the number of holes diffusing into the n-doped region is higher than the number of electrons diffusing into the p-doped region, thus creating an asymmetrical charge carrier free region, and an asymmetrical electric field (see Figure 1).

The area around the N/P junction which is free of charge carriers is called the depletion zone. When an external voltage is applied, positive on the p-doped side, negative on the n-side, the electric potential across the junction grows, increasing the size of the depletion zone. This is referred to as reverse biasing the junction.

The exact width of the depletion zone can be calculated using the following formula [5]:

$$\text{a)} \qquad W = \sqrt{\frac{2K_s\,\varepsilon_o}{e}\,\frac{N_A + N_D}{N_A\,N_D}\,(\phi_B + V_R)}$$

Where N_A is the acceptor (p-), N_D the donor (n-) doping concentration, K_s is the dielectric constant of the semiconductor, ε_o the permittivity of free space, ϕ_B the intrinsic potential across the junction and V_R the reverse bias voltage.

2.1.2. Semiconductor Detector Operating Principles

The active volume of any such detector is the depletion zone at a P/N junction, typically the junction between a highly doped p-layer (up to $10^{17}/cm^3$) on a much lower doped substrate (as little as $10^{13}/cm^3$). The depletion zone can be increased in size by choosing a low bulk doping concentration and reverse-biasing the detector. The voltage needed to deplete the entire detector is called the depletion voltage, it can be several kV on Ge-pin diodes, but is typically only 60V on a 300μm thick Si-microstrip detector, depending on the dopant concentration chosen (see 2.1.1 a).

Ionizing particles interact in several ways, leaving behind a trail of electron-hole pairs. The interaction responsible for almost the entire energy loss is coulomb scattering. The energy T transferred to an electron at distance b from the particle track can be approximated, considering the electron as free. This approximation is valid only for energy transfers much bigger than the ionization energy of the electron [6].

$$\text{b)} \qquad T = \frac{2e^4}{m_e c^2 b^2 \beta^2}$$

Where b is the distance of the electron from the particle track, β the usual v/c (v is the speed of the particle), m_e is the electron mass, and e the electron charge.

Integrating over the probability of such an interaction and refining the approximation for bound electrons leads to the well known Bethe-Bloch-Formula for the average energy loss [7].

$$\text{c)} \qquad -\frac{dE}{dx} = Kz^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ \ln\frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}\right]$$

Where the maximum energy transferred on a single electron is:

$$\text{d)} \qquad T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1+2\gamma m_e/M + (m_e/M)^2}$$

The definitions used in c) and d) can be found in Table 1.

Table 1 : Symbol Definitions

Symbol	**Definition**
K	$4\pi N_A r_e^2 m_e c^2$
r_e	Classical electron radius
m_e	Electron mass
z	Charge of particle in units of electron charge
A	Atomic mass of medium
Z	Atomic number of medium
β	v/c
γ	$(1-v/c)^{-1/2}$
I	Mean excitation energy
δ	Density effect correction to ionization energy loss
M	Incident particle mass

In thick samples (several mm) the energy loss distribution is a gaussian. In thin samples, such as si-microstrip detectors however, a single interaction with a large energy transfer makes up an appreciable part of the overall energy loss, the statistics of large numbers is no longer applicable. This leads to the asymmetrical Landau distribution of the energy loss probability, where the most probable energy loss is lower than the average energy loss [8]. This means that in thin samples the most probable energy loss is no longer a linear function of the sample thickness, but depends on the asymmetry of the distribution. In silicon the most probable energy loss is 280 eV/μm for a 300 μm thick detector [7].

The scattered electron dissipates the absorbed energy within a few μm of its original site by ionizing other electrons. The energy needed to create an electron hole pair in a silicon crystal is about 3.6 eV [7], therefore a minimum ionizing particle creates about 77 electrons hole pairs per μm in silicon. That corresponds to 23000 electrons in a 300 μm thick detector.

Under normal circumstances these electrons would eventually recombine with the ionization sites. In a semiconductor detector however this is avoided by the depletion of the detector. No free charge carriers are available for recombination in the active detector volume. Impurities or errors in the crystal lattice structure, however, can capture electrons or holes for several micro seconds. The so caused reduction in signal height is less important than the noise current caused by the randomly released charge carriers.

The electric field in the detector (see Figure 1) causes the electrons and holes to drift towards the anode and cathode, respectively. For E-fields < 0.2v/µm the drift velocity can be calculated as:

$$\text{e)} \qquad v_d = \mu\, \varepsilon$$

Where µ is the mobility of the charge carriers: (1350 cm^2/v sec for electrons, 480 cm^2/v sec for holes [6] in lightly doped silicon). For higher E-fields the drift velocity approaches a maximum drift velocity of about 8.5 10^6 cm/sec for electrons, 4.45 10^6 cm/sec for holes.

2.1.3. Detector Speed

The charge collection time is limited only by the drift time of the holes in the silicon. At a depletion voltage of 60 V on a 300 µm detector the field strength in the detector is 2 kV/cm. That corresponds to a drift velocity of 27 µm/ns for electrons, and 9.6 µm/ns for holes. For a 300 µm detector this corresponds to a charge collection time of 30 ns at the cathode and 11 ns at the anode. Thus a silicon vertex detector is among the fastest detectors.

2.1.4. Silicon Microstrip Detector Geometry

Si-microstrip detectors are thin (typically 300µm thick) detectors used to detect the particle position.

A typical double sided detector has a low doped n-bulk with highly doped p- and n-sides. Both n-side and p-side of these detectors are separated into long, narrow (20-50µm wide) strips (see Figure 2). The n-side strips are orthogonal to the p-side strips, forming a two dimensional readout grid. This does not provide a true two dimensional readout. Each side supplies only one dimensional information leading to ambiguities as soon as several particles are recorded at the same time.

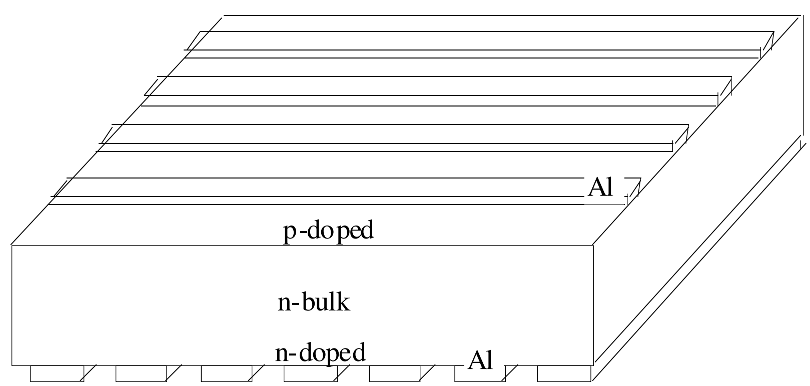


Figure 2 : Drawing of a simplified silicon microstrip detector with double sided readout

2.2. Details of Particle Detection with Si-Microstrip Detectors

A typical minimum ionizing particle ionizes about 23000 electrons when it traverses 300µm of silicon. If it traverses the detector at an angle the distance within the detector is longer, so the total deposited charge is proportionally higher, but the charge on any given channel might be considerably less. This reduces the resolution of the detector by diminishing the signal to noise ratio, and increases the demands on the dynamic range of the readout electronics.

The p-side consists of highly p-doped strips separated by the n-doped bulk. This creates insulating p-n-p junctions between the strips which allow an independent measurement of the charge collected by each strip.

The n-side is more complicated. Between the strips there is no insulating p-n junction. The high resistivity of the n-bulk silicon is not enough to separate the strips from each other, because the surface layer between the Si and the SiO_2 always contains virtually free electrons [5]. These electrons reduce the resistivity between

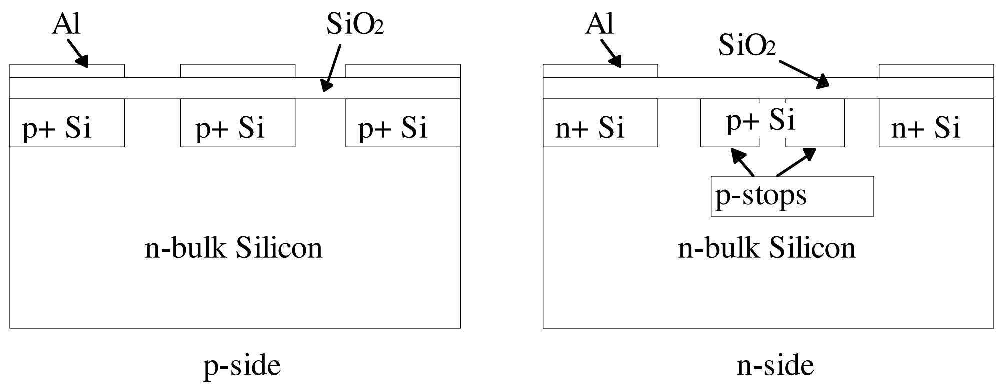


Figure 3 : Simplified drawing of the two detector sides

neighboring n-strips so far that individual readout is impossible. Therefore narrow p-doped strips (p-stops) are implanted between the n-strips creating insulating n-p-n junction between the strips. Figure 3 shows a simplified drawing of both the detector p- and n-side.

There are two different options for detector readout: The readout strips can be DC coupled to the preamplifier. In this case there is no insulating SiO_2 layer, as shown in Figure 3, between the silicon readout strips and the aluminum strips that are connected to the preamplifier input. This creates a somewhat simpler detector, increasing the yield in the detector production, but also increasing the detector noise, because the strip leakage current is integrated along with the detector signal, making it necessary to subtract the quiescent leakage current from the signal.

In the AC coupling scheme the insulating layer creates a capacitive coupling between the detector and the preamplifier. This layout is often preferred, because it has superior noise characteristics and allows simpler preamplifiers, because no current is superimposed on the signal.

In order to fully deplete the detector a bias voltage has to be applied between the two detector sides. To achieve this the readout

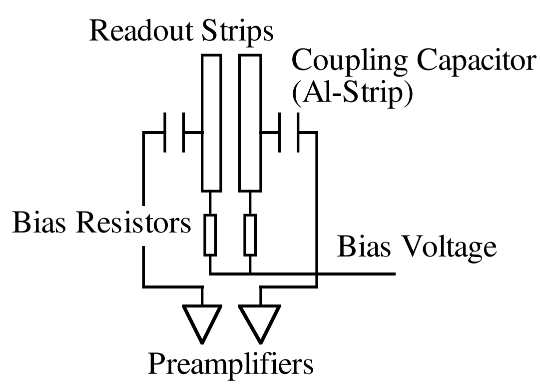


Figure 4 : Schematic of the Biasing Circuitry

strips have to be connected to a bias line, which supplies each detector side with the depletion voltage. To prevent signal loss into the bias line the readout strips have to be connected to it via a large resistor, so that the peaking time of the preamplifier is short compared to the RC constant of the system. On the other hand the resistor has to be small enough to allow for a maximum leakage current of 100 nA without significant voltage drop. Figure 4 shows the schematic of two AC coupled detector strips and their respective biasing resistors.

There are several different options for the resistor implementation on the n- and p- side:

1. On n- or p-sides a poly-silicon strip can be used as a resistor, connecting the bias line to the readout strips. The resistance can be calculated from the dimensions of the resistor and the specific resistance of poly silicon.
2. On the n-side the surface conductivity can be exploited to form simple channel resistors. The resistance can be calculated by the channel dimensions and the surface layer sheet conductivity, which is process dependent.
3. On the p-side punch-through biasing is often used: A p-doped bias line is run across the detector close to the end of the readout strips. The punch through effect described below[1] controls the resistivity. [5]

On both sides a guard ring surrounds the entire active area of the detector. This ring gracefully degrades the electric field at the detector edge. This avoids a high electric field at the corners of the detector, thus preventing a current flow at the detector edge.

2.2.1. Resolution Improvement Using Floating Strips

A common technique used to improve the resolution is the use of unbonded sense strips between the bonded readout strips. Sense strips and readout chips are identical, except that readout strips are connected to a preamplifier, while sense strips are left floating. This increases the resolution at constant readout strip pitch without requiring additional electronics. The two track resolution (ability to separate two tracks) is not affected.

If an unbonded channel is hit its charge is shared with its neighbors by capacitive coupling. The center of the charge distribution can then be calculated, and the beam position can be found within a fraction of the readout strip pitch.

The σ of the charge distribution of an orthogonal track is typically on the order of 10 μm, therefore the entire charge is injected in no more than 2 strips at a 50 μm strip pitch. Tracks with a low angle to the detector on the other hand deposit

---

[1] The punch through effect is a breakdown condition of an insulating pnp junction. A potential difference across a junction increases the size of the space charge region. If the n-layer is only lightly doped, as in our case, the space charge region of the pn-junction reaches the np-junction. This creates a continuous depletion zone connecting both p-regions, thus allowing a current to flow.
The voltage at which punch-through occurs can be calculated as:

$$\text{a)} \qquad V_p = \frac{e\,Q_B}{K_s \varepsilon_0}\left[W_B + \frac{Q_B}{2\,N_A}\right]$$

Where Q_B is the total number of impurities in the n-region, W_B is the width of the n-region and N_A is the dopant concentration in the sense strip.

more total charge due to the longer track in the detector, but this charge is spread out over many channels. Therefore the signal height in any given channel for an angled track can be an order of magnitude smaller than for an orthogonal track, depending on detector geometry. This means that orthogonal tracks profit from dense strip pitches, which increase the median finding precision, while angled tracks can be read with more accuracy using wider strips, because the signal of angled tracks is stronger in wider strips. The unbonded sense strips benefit both the orthogonal tracks by decreasing the strip pitch, and the angled tracks by inducing stronger signals in the readout strips. A wider sense strip pitch increases the ratio of tracks that deposit charge in only one channel, which makes it impossible to find the median, thus limiting the resolution to the sense strip pitch.

2.2.2. Detector Resolution

The theoretical spatial resolution of a silicon vertex detector is limited by two factors. For orthogonal tracks the limiting factor is given by the displacement of the center of the charge distribution by single high energy electrons. The probability of a 5 μm displacement is on the order of 10% for a 300 μm detector [6]. The transfer of a large amount of energy (>100 keV) onto a single electron has a low probability (<10% for a 300 μm detector), but those electrons then travel tens of microns until they range out, depositing a large charge relatively far away from the original particle track.

For angled tracks the high energy electrons have even more severe consequences, because a high energy electron near the end of the track can pull the center of the charge distribution over by tens of microns.

The second limiting factor is the variation in the generation of charge itself. This does not have an influence on orthogonal tracks, but the center of the charge distribution for an angled tracks will depend on the charge distribution along the track. The energy deposited in a certain length of silicon by a particle varies according to the Landau distribution [8]. The statistical fluctuation in energy loss causes an identical fluctuation in deposited charge, thus introducing an error into the position of the center of the charge distribution.

The theoretical limit in the resolution strongly depends on the angle of the track. For orthogonal tracks the theoretical limit is only a couple of microns [6], while the limit for tracks at an angle of less than 45° is about 10μm.

In practice, the resolution is a result of sense strip pitch, readout strip pitch, and detector thickness. It is also influenced by the total system noise, including the preamplifier noise, and the readout precision (step size of the ADC).

2.2.3. System Noise

The system noise is the sum of the noise contributed by the detector and the readout electronics. The detector inherent noise itself is a sum of several contributions: The first contribution is the thermal generation of electron-hole pairs at impurities or lattice defects (shot noise) [6]. The noise current can be calculated as [9]:

$$\text{a)} \qquad I_{noise} = (2\, e\, I_{dc}\, B)^{1/2}$$

where e is the electron charge, I_{dc} the dc current (in the case of the detector this is the leakage current), and B is the limiting bandwidth of the amplifier or measurement. Even after a large radiation dosage the leakage current is usually not the dominant noise source. The noise voltage can be calculated by multiplying the noise current with the bias resistor value.

$$\text{a')} \qquad V_{noise} = (2\ e\ I_{dc}\ B)^{1/2}\ R$$

The second contribution to detector generated noise is the thermal or johnson noise of the bias resistors as follows [9]:

$$\text{b)} \qquad V_{noise} = (4\ k\ T\ R\ B)^{1/2}$$

where k is Boltzmann's constant, T the temperature and R the bias resistor. To calculate the the equivalent noise charge (ENC) the noise voltage is multiplied by the input capacitance.

The readout electronics contribute to the noise mainly due to the noise of the first input transistor. Transistor noise voltage is comprised of two main components: the serial or current noise and the parallel or voltage noise. Both contributions depend strongly on the IC process, and transistor parameters [9]. For a silicon FET preamplifier the noise is calculated in [10].

$$\text{c)} \qquad ENC = \left[\frac{16}{3}\, k\, T\, \frac{0.7}{\omega_T\, T_M} \left(\left(\frac{C_D}{C_I}\right)^{1/2} + \left(\frac{C_I}{C_D}\right)^{1/2} \right)^2 C_D + e\, I_g\, \frac{14}{15}\, T_M \right]^{1/2}$$

Where ω_T is the gain-bandwidth product, T_M the peaking time, C_I the preamplifier input capacitance, C_D source (detector) capacitance, and I_g the gate leakage current. Usually the leakage current is small, so the first summand is dominating. At $C_I \ll C_D$ the inner sum can be simplified to C_D / C_I . The noise is then directly proportional to C_D. Even though all possible steps are taken to minimize the transistor noise, this is the dominating factor in the entire system noise.

2.3. The Detector Readout

2.3.1. Custom Designed Readout Integrated Circuits

Many different options for the readout of silicon microstrip detectors have been developed and used in the last decade. The design goal is always the same: to achieve the highest possible channel count in the smallest possible area, consume as little power as possible and to be resistant against radiation damage. The optimum tradeoff between all those parameters is different in every experiment. Therefore a new readout chip was designed for many of the recent experiments, each of them optimized for it's working environment.

A common architectural element in these recent designs is the compression of the data from multiple channels (128 channels in this experiment) onto a single readout bus. This can be done either analogically or digitally [3]. A switched capacitor array is often used to temporarily store analog voltages, then connect one channel after another onto an analog readout bus. The digital approach requires digitizing of the analog data on chip, then storing it, until it can be read out on a, digital (in our case serial), bus. To keep the chip small enough different approaches have been taken to simplify the ADC and the preamplifier. The design of the SVT readout chip will be explained in detail below.

2.3.2. Zero Suppression

In any detector system data from channels containing no interesting information has to be suppressed. This zero suppression is performed by differentiating between hits and background noise based on a certain criteria. In some cases these criteria can be complicated, taking into account all neighboring channels, but usually it just means that the signal in a certain channel has to rise above a threshold in order to be read out.

The large amount of data generated by thousands of channels requires zero suppression early in the data flow system to prevent inevitable bandwidth problems. In a digital system this can be done on the readout chip itself, if the threshold is set on that chip. This allows the chip to distinguish a real signal from background noise. In analog systems, or digital systems where the threshold is not set on-chip, all data has to be sent further before the zero suppression can occur.

2.3.3. Data Flow

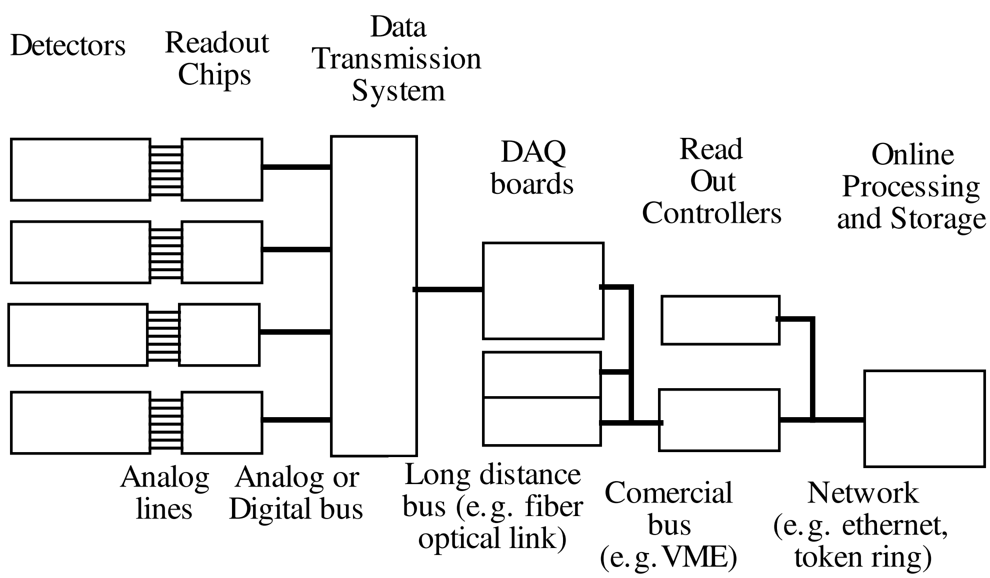


Figure 5 : Data Flow Block Diagram

The data from many (on the order of 100) detector channels is read out by each readout chip. It is then passed on to the next higher level, where data from many chips is gathered on data acquisition (DAQ) boards. These boards are as specialized as the front end chips that they are connected to. They digitize data and perform zero suppression, if this has not occurred before. They also process and reformat the digital data before passing it on to the next link in the readout chain, the read out controller.

The read out controller (ROC), typically a single-board-computer with a standard bus (such as VME) to the data acquisition boards. From this computer the data is sent to the online processing facility, usually a farm of workstations. This computing facility separates out interesting events, extracts salient parameters from the data, and sends it to temporary or permanent storage.

3. The BaBar Silicon Vertex Tracker (SVT)

3.1. Requirements for the BaBar SVT

The BaBar experiment measures CP violation by observing the time evolution of the B B-bar final state. This requires extremely precise spatial resolution, both to identify the originating vertex for the majority of the particles and to reconstruct the time at which any decay has taken place. In order to be able to reconstruct the timing from the spatial information the collider was designed to produce B B-bar pairs that are not at rest in the lab frame of reference. This is achieved by colliding a 9.00 GeV electron beam with a 3.109 GeV positron beam [4]. Any particles so produced have a strong boost along the beam axis. A cylindrical coordinate system is used to describe the detector and the interactions. The positive Z direction is the direction of the electron beam. A measurement of the Z-coordinate with an accuracy of 100 μm is sufficient to calculate the decay time with an accuracy of 100 μm/c = 0.3 ps.

The BaBar SVT system has to deal with many requirements that were not imposed on other silicon vertex detector systems. The main purpose of the SVT is to pinpoint the secondary vertices with extremely high precision. The design requirement is to determine those vertices with better than 90μm accuracy [1]. This is no easy task, because the decay takes place well within the beam pipe, introducing additional errors by multiple scattering in the beam pipe.

Another important requirement for the SVT is the ability to reconstruct entire tracks without information from the other detector subsystems. This enhances the precision for the secondary vertex position in events where some tracks have a low transverse momentum, and therefore never enter the detector subsystems that are located further away from the beam axis [1].

3.2. Layout of the SVT

3.2.1. Mechanical Design of the SVT

The B-factory collider beam energy asymmetry requires an asymmetrical detector. The angular acceptance in the forward region has to be much higher than in the backward region to achieve a homogenous acceptance in the particle's frame of reference [4]. Figure 6 shows the relationship between the polar angles in the laboratory frame of reference and the center of mass system. For photons the boost $\beta\gamma$ can be calculated as the energy in the center of mass system divided by the momentum difference.

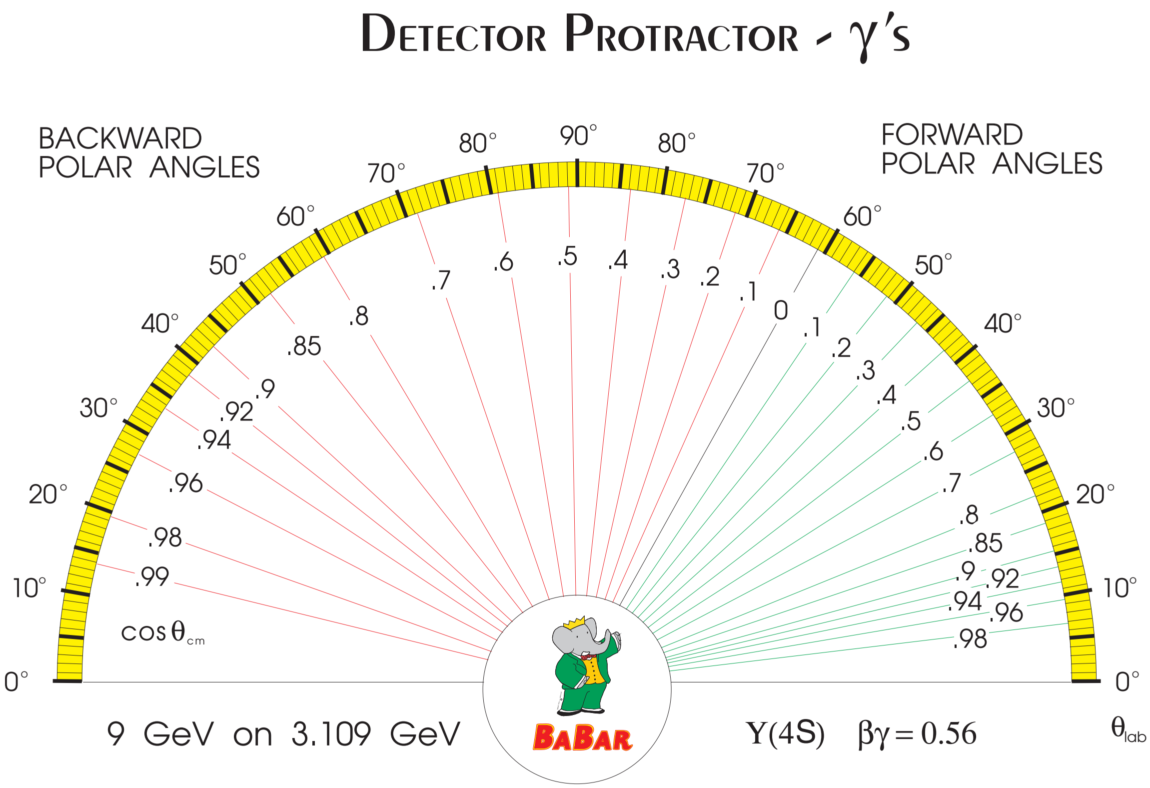


Figure 6 : Protractor showing the relationship between center-of-mass and laboratory polar angles for photons at ßγ = 0.56

The SVT is a silicon microstrip vertex detector using a total detector area of about 1 m^2. It consists of a total of 5 layers: 3 barrel shaped inner layers with 6 modules each, and 2 outer layers which are comprised of 16 and 18, respectively, bridge like detector modules. The overall mechanical design is shown in Figure 7.

The entire detector rests on two gimball rings, which are mounted on the two bending magnets closest to the interaction region (B2 magnets). The gimball rings are fixed to the detector cooling and support cones, which in turn support the detector modules. The space frame, an external carbon fiber skeleton, gives additional structural strength.

The fully assembled detector is inserted in the support tube, a carbon fiber pipe that gives mechanical support to the B2 magnets. The entire assembly is buried deep in the final detector assembly, making it virtually impossible to access during regular maintenance.

The asymmetry of the detector was one major design parameters requiring extremely large angular acceptance in the forward region. This made it necessary to minimize the size of the readout electronics in order to get adequate angular acceptance.

In the design of the position of the actual detector layers two parameters played a mayor role: The decreasing sensitivity at low angles, and the requirement to be close to the beam pipe to reduce the influence of multiple scattering on the vertex resolution. This led to three barrel shaped inner layers, to bring those detectors as close to the beam as possible. There are two bridge shaped outer layers which have the detectors in the forward and backward region tilted inwards towards the beam axis. This reduces the necessary active area at constant angular coverage, and increases the detector sensitivity in the end regions by raising the angle of impact of the incoming particles.

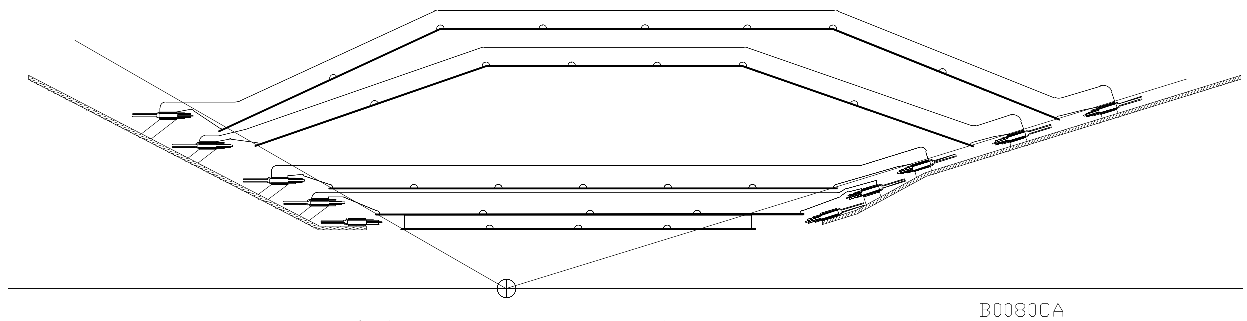


Figure 7 : Overview of the SVT Mechanical Structure. The three concentric inner layers and the two bridge shaped outer layers can be seen in the cutaway.

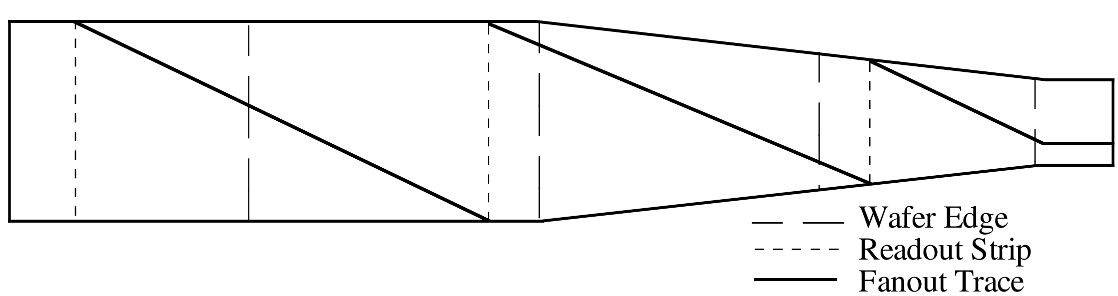
3.2.2. Sense Strip Layout

The tilted parts of the bridge layers have to be slightly wedge shaped to fit the cylindrical geometry. This means a variable strip pitch is necessary, reducing the strip pitch from 100µm in the rectangular area to 65 µm at the end of the trapezoidal section (in ϕ direction). The barrel layers have a constant sense strip pitch of 50µm.

In the inner layers the maximum readout density of the chip is utilized, in the outer layers spatial information is less critical, so a wider readout pitch is used to reduce the amount of data that needs to be processed. Thus the highest concentration of readout chips is in the 3rd layer, while the 4th layer has the highest number of overall readout channels due to its larger area.

For the inner layers the maximum strip density given by the readout chip (50µm) is used in the ϕ direction. In the Z direction only every other strip is bonded, resulting in a readout pitch of 100µm.

Even though the Z strip pitch is much lower there are still more Z readout strips than phi readout strips. The electronics however can handle only the same amount of strips on either side of the detector. Therefore ganging is used to reduce the number of Z readout channels. Ganging connects together Z strips from



different parts of the same detector module. Figure 8 illustrates times three ganging on a detector module from layer 5. Ganging introduces additional ambiguity into the detector readout. The 5 layers of the detector and its high efficiency will help to resolve those ambiguities. This can only be done during track reconstruction, where the entire data of one event is available. This would be either at the online processor farm, or the offline processing facilities.

The overall achievable resolution with the SVT detector is expected to be 10µm in phi and 12µm in z direction. Any further increase in spatial resolution would not noticeably improve the vertex resolution of the detector due to the resolution limit introduced by multiple scattering in the beam pipe and the detector itself [1].

3.2.3. Detector Biasing

The full depletion of the SVT detector wafers is achieved at 60V. On the p-side the detector is biased by punch-through biasing, described previously. This effect has been used at other detectors (e.g. Aleph) and has proven to be radiation hard. Biasing on the n-side is achieved by poly-silicon resistors. These resistors are 480µm long, connecting the n-bias line to the sense strips. They are 3µm wide, flanked on either side by the p-stops. This technique was chosen over channel resistors, because they are not radiation hard.

3.2.4. Radiation Length in the Detector

The radiation length of the complete SVT assembly is one of the limiting factors for the resolution of the detector components that are further away from the interaction point. The position resolution degrades with the square root of the mass between the beam and the detector itself. When considering the influence of the mass of the SVT, both the mass of the beampipe, and the mass of the support tube have to be taken into account.

The space frame, seen surrounding the detector in Figure 7, is the only structural support for the entire SVT. There is no internal support pipe. The space frame is an extremely lightweight construction, made from hollow carbon fiber rods. The detector modules themselves are part of the mechanical strength of the assembly, and are essential for a stiff detector structure.

3.3. The Readout System

The SVT readout system is a set of custom built electronics, specifically designed to meet the requirements of this detector subsystem. The first link in this chain is the custom designed SVT readout chip, the end of the chain is the DAQ board. In this section a detailed description of every element in this chain will be presented.

In some collider experiments the Trigger is derived from data collected in all subsystems. The silicon vertex detector then needs to forward a limited set of trigger data from every beam crossing. In the PEP II environment the beam crossings are virtually continuous (238 MHz). This would make such an approach very difficult. Therefore the BaBar trigger is derived from the calorimeter and drift chamber data alone. This trigger is then supplied to the SVT exactly 12.9 µs after the event occured, so no data needs to be read out until a trigger is received. This reduces the amount of data that needs to be sent from the SVT readout chips to the DAQ board by orders of magnitude.

3.3.1. The SVT Readout Chip

The chip measured for this thesis is a radiation soft prototype of the SVT readout chip. It contains the full functionality of the final chip, and is expected to behave very similarly. The chip is commonly referred to as the AToM (A Time-over-threshold Machine) chip [2].

The AToM chip converts 128 analog inputs to digital data, using a time over threshold (ToT) analog to digital converter. It buffers the data for 12.9 μs in a trigger latency buffer and stores both time and charge of the input signal in one of three buffers upon receipt of a trigger. Data is read out on a single serial data line.

The chip has two well separated sections: the analog and the digital part. In the analog part the input charge on each channel is integrated, the resulting voltage peak is amplified and shaped, using a $CR/(RC)^2$ filter. In the digital section the signal is sampled, buffered and stored. The digital section also contains the decode and control logic. Figure 10 shows a functional block diagram of the chip.

3.3.2. The SVT Design Philosophy

After installation repair work on the SVT is virtually impossible, because it is placed inside the support tube, which makes access extremely difficult. It is not expected that there will be a chance to service the detector within the first decade of the experiment. This naturally poses great demands not only on testing and quality control, but also on an inherently robust and redundant design. Every chip that will be used in the detector will have been fully tested, and is produced in a radiation hard process. However this cannot eliminate the possibility of a chip failure. The design philosophy for the SVT demands that the failure of any single component must have minimal impact on the entire subsystem. For the chip this has two very different implications. The chip itself must be designed to minimize the impact of damage to any part, and it must deal with an environment that might have defective components. For example a bad detector strip might put a current into the input of the preamplifier. The chip must prevent excessive crosstalk of this noise into the rest of the system. Each chip has been designed to receive two command and clock lines. Should external failure cause one line to become inoperative the chip will still be able to function properly by switching to the backup line.

3.3.2.1. Implications on the Chip

To make the chip itself more failure resistant most of the internal logic is duplicated once per channel, so a failure in the internal logic will effect only one channel. However some control logic, as well as all drivers and receivers have global functionality. The limited space available for wires in the support tube made it impossible to connect a dedicated command and data line to each chip. To minimize the number of lines needed to read out the detector all chips in a readout section (one side of a single detector module) can send their data on a single line. All chips receive all commands in parallel from a single command bus, but only the first chip (master chip) answers to a read event command. It sends it's data to the higher level readout electronics, then passes a token to the next chip on the line, which then starts passing data to the master, who in turn sends it on to the higher level readout electronics (see Figure 9). A chip is identified as master by bonding a dedicated pad to +5V.

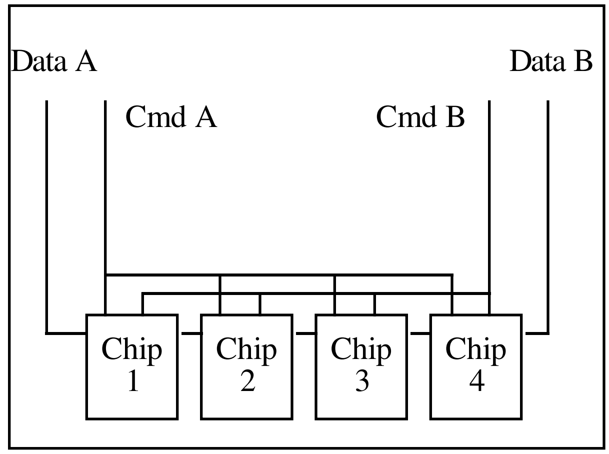


Figure 9 : Schematic of the data and command flow. Chip 1 and 4 are 'master' chips in this configuration

Still, the total failure of one chip must not affect any of the other chips in the same readout section. This is one of the reasons for the existence of the master chips, which send all data from the entire readout section to the next higher level of readout electronics. If multiple chips would share a common output bus, any one of those chips could cause a global failure by shorting this bus to power or ground.

The problem with the master slave configuration is that a single non-functional master would cause loss of all the slave's data. To prevent this, each chip can send its data either to the left or to the right neighbor in the readout section, both the outermost chips are configured as masters, and correspondingly two readout lines are connected to each readout section (see Figure 9). If any single chip should fail, only its data is lost, the other chips can be reprogrammed to successfully deliver their data.

To prevent severe consequences from a broken bond-wire or cable the command and clock lines are doubled: the chip "listens" to the default clock and command line upon power up. If the default clock line does not oscillate the chip switches to the backup command and clock lines. It can also be programmed to select the alternate command and clock lines.

3.3.2.2. The Analog Section

The analog section is fully parallel for all 128 input channels. This chapter describes one of those channels, the exact same logic is duplicated 128 times to make up the entire analog section. Figure 10 shows a diagram of the functional blocks of both the analog and digital section.

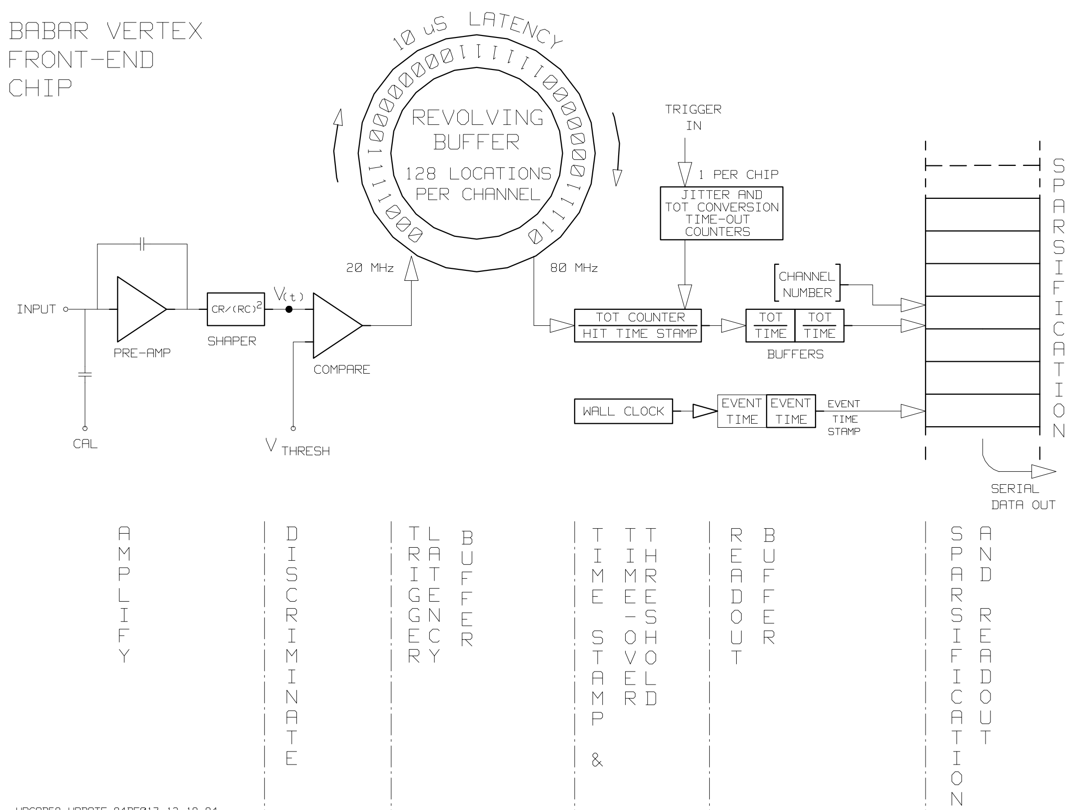


Figure 10 : Functional Diagram of the AToM Chip showing the Analog Section consisting of the Preamplifier, the Pulse Form Shaper, and the Comparator, as well as the Digital Section, consisting of the 12.9 μs Trigger Latency Buffer (pictured as revolving Buffer), the ToT and Timestamp Extraction Logic, and the Three Data Buffers.

The first stage of the chip is a charge sensitive preamplifier. This preamplifier has a continuous reset feature via a 300Ω resistor. The continuous reset is a specific requirement for this chip, because PEP II beam crossings occur at a rate of 238 MHz, essentially continuous in respect to the speed of the readout electronics. In other experiments beam crossings occurred at the kHz range, allowing a discrete reset of the integrating preamplifier after each crossing.

After being amplified the input signal is shaped via a $CR/(RC)^2$ shaper. The overall gain of the preamplifier/shaper combination is designed to be 100 mV/fC. The output of this shaper is a second order semi-gaussian pulse. Shaping is necessary to obtain a well defined charge-amplitude relationship. For a $CR/(RC)^2$ shaper the charge amplitude relationship is linear over the dynamic range, in this case 2V or 20 fC input charge [2]. A special feature of the shaper is the digitally selectable peaking time. The peaking time can be varied from 100 ns to 400 ns in increments of 100 ns. The shortest peaking time will be used in the innermost layers which have less capacitance. The higher peaking times will be used in the outer layers where ganging and long readout strips increase the capacitance. This results in an improved signal to noise ratio for the outer layers. The high background occupancy in the inner layers prohibits a longer shaping time [22].

The comparator is the only junction between the digital and the analog section. It compares the shaper output to a digitally settable threshold voltage. If the shaper output is above this voltage the comparator output goes to a logic high state, it otherwise remains at a logic low state. The threshold voltage is settable between 0 and 280 mV in 4.5 mV steps.

Right at the input of the chip is a charge injection capacitor which allows the injection of a programmable charge for channel to channel calibration purposes. An absolute calibration requires external charge injection, because the absolute size of the internal calibration capacitor is not well known. The internal calibration capacitor can be pulsed with a programmable voltage step to inject a charge into the preamplifier input. The value of the injected charge is proportional to the capacitor and the voltage step. The amplitude of the voltage step is programmable between 0 and 540 mV in steps of 8.5 mV ~ 0.5 fC [2].

3.3.2.3. The Digital Section

The digital section combines the 128 comparator outputs to deliver only one serial output. Most parts of the digital section however are still separated into 128 channels.

The comparator output value is stored in a circular buffer at a sampling rate of 1/4, 1/3 or 1/2 of the system clock speed. The system clock speed is 1/4 of the 238 MHz PEP II bunch crossing rate (59.5 MHz). Only the 15 MHz sampling rate will be used in the experiment. The other rates were included, because the speed of the system clock had not been finally decided when the chip was designed.

The circular buffer has 193 locations, supplying approximately 12.97μs trigger latency (at 15 MHz) before the data is overwritten. If no trigger is received within 12.97μs the old data is overwritten by new data. Thus the circular buffer contains a 12.97μs history of the state of the comparator output.

Upon receipt of a trigger the read pointer moves through the buffer at a speed of 59.5 MHz, starting with the oldest, not yet overwritten location, moving into the newer data. When the read pointer reaches a zero-to-one transition in the buffer it starts counting the number of successive ones in the buffer. It scans a window of up to 2μs worth of data before determining a no-hit for the channel. The window can be programmed with 5 bit accuracy from 0 to 2μs.

If a channel is hit (a zero-to one transition is found), both the position (timestamp) and time over threshold (the number of consecutive samples for which the signal is above threshold) for each channel are stored in the first of three buffers, if this buffer is free. Otherwise the data is discarded.

The method of counting the number of consecutive ones in the circular buffer gives a measure of the length of time the shaper output was above the threshold. This time over threshold (ToT) value is proportional to the input charge if the shaper is operating within its dynamic range. The pulse shape controls this dependency, an ideal gaussian pulseshape would give a logarithmic relationship between charge and ToT counts. This expands the dynamic range by compression of high amplitude signals. Therefore the chip has a high resolution especially on channels with small signals. This improves the median fitting capabilities for tracks with low signals. Typically angled tracks deposit little charge into any given channel, therefore the chip automatically improves resolution on angled tracks, which would otherwise have very poor spatial resolution.

At the same time the ToT analog to digital converter (ADC) is very simple, much simpler than any common ADC with a linear response. The design therefore combines the advantage of a smaller circuit with a better resolution. Simulations have shown that the compressed readout with just 5 bits accuracy performs almost as good as a fully analog (infinitely accurate) readout, assuming that standard median finding procedures are used[1].

To accommodate the longer ToT counts associated with the 300 and 400 ns shaping time setting a programmable 'skip-control' has been implemented. This register enables a prescaler counter which scales the ToT value by a factor 1 (no skip), 2, 3 or 4.

When the read pointer scans the circular buffer only a zero-to-one transition is registered as a start of a hit, if the comparator is always high no hit is registered. This is considered a minor drawback, because at a reasonable threshold only bad channels would have a constant high comparator value.

A 5 bit wall clock counter on the chip counts at the system clock frequency. The trigger accept command contains a 5 bit trigger identification tag. This tag can be set randomly by the module that generates the trigger accept command. The wall clock value (trigger time) and the trigger tag are stored along with the associated data in one of the three buffers.

The content of a lowest buffer is moved into the next higher buffer whenever that buffer is vacant. It is eventually moved into the third (highest) buffer. From there it is transferred off-chip in response to a read event command.

When the chip decodes a read event command, it first sends a header containing the chip address, the trigger tag and the trigger time (both of which are used to keep the readout system synchronized). It then sends the channel number, ToT value and time stamp of each channel that received a hit. This dramatically

Table 2: The SVT Readout Chip Commands

Command	Action
No Op	No action is required
Clear Readout	Reset buffer pointers, counters, occupancy flags
Sync	Reset time counter, reset divide by n counter for local clock.
L1 Trigger Accept	Read data from circular buffer, perform ToT processing.
Read Event	Send data stream from highest readout buffer.
Calibration Strobe	Inject a charge (selectable by calibration DAC) into the preamplifier input.
Write Control Register	Write data stream into control register.
Write Calibration and Channel mask registers.	Write data stream in both calibration and channel Mask Register
Read Registers	Read back value of control, calibration and channel mask registers.
Chip Reset	Reset all registers.

reduces the amount of data that needs to be sent off-chip as compared to a readout that transmits data from all channels.

The digital section also contains the command decoder, which decodes the serial bit stream on the command line. There are two types of commands: SVT local and BaBar global commands. The difference between global and local commands is that global commands can immediately be followed by a trigger, while local commands have data attached, or require a certain period of time before the next command is sent (e.g. reset). Therefore the local commands can not be sent during runtime, because they would interfere with the prompt sending of the trigger command. The first six commands in Table 2 are global commands, they can be sent at any time. The last four commands are SVT specific commands, they may be sent only during calibration and setup.

3.3.2.4. AToM Chip Registers

The control register is used to control all the above mentioned programmable features, such as the calibration and threshold digital to analog converters (DACs), along with the trigger window size and the sampling rate. A bitwise description of the AToM registers can be found in [2].

The channel mask register is used to mask of any unused or bad channels, so reconstruction software does not have to deal with invalid data. This also reduces the amount of transmitted data.

The calibration mask allows any combination of channels to receive a pulse from the internal charge injection capacitor. This will probably be less important during runtime, then it is during testing, where this feature can be used to facilitate measuring channel to channel crosstalk.

3.3.2.5. Power Consumption and Noise Considerations

Each chip has two well separated sections: the analog and the digital part. Great care has been taken to avoid crosstalk between those sections. They have a completely independent power connection, and even the ground is not connected on chip. The places of contact have been limited as far as possible. There are only three inputs to the analog section, which are connected to the digital section: the calibration and threshold DACs along with the calibration charge injection switches. The calibration charge injection switches are by far the most critical, as they are coupled to the preamplifier input, but they will not be used during data taking.

Both sections are built in MOSFET technology to conserve both space and power. Most MOSFET process parameters are very well studied for digital designs, as the vast majority of all MOSFET designs are fully digital. Yet many of the parameters that are important only for the analog designer are not well known. This makes the prototyping and exact measuring of prototypes extremely important.

To minimize both the power consumption and interference with the analog part all off-chip signals are low swing differential signals. The advantage of the differential signal is not only the excellent signal quality due to common mode noise rejection of differential receivers, but also an almost constant current requirement for the drivers. Whenever one side is turned on, the other side turns off, reducing the amount of spikes introduced on the power planes.

3.3.3. Chip Mounting

Due to the limited space available for the electronics both the readout chips and all the external components for both the z and phi readout of a detector are placed on a single hybrid. This High Density Interconnect (HDI) is fabricated from aluminum nitrate substrate to allow for a low-resistivity heatflow from the chips to the mounting points of the hybrid, which are connected to the chilled water cooling system. The hybrid has heat expansion properties similar to silicon, so the chips can be mounted directly on the HDI without chip carriers.

The kapton fanout connects the detector modules to the readout chips. One side of each trace on the fanout is bonded to a readout strip, while on the other side the fanout is glued to the front of the HDI, the fanout traces are then directly bonded to the readout chip preamplifier input pads, using a staggered, double layered bonding pattern to achieve the required bonding pitch of only 50μm (see Figure 11).

The measurements in this thesis did not include a HDI, instead a specifically designed test board, which will be described later, was used to test a single chip;

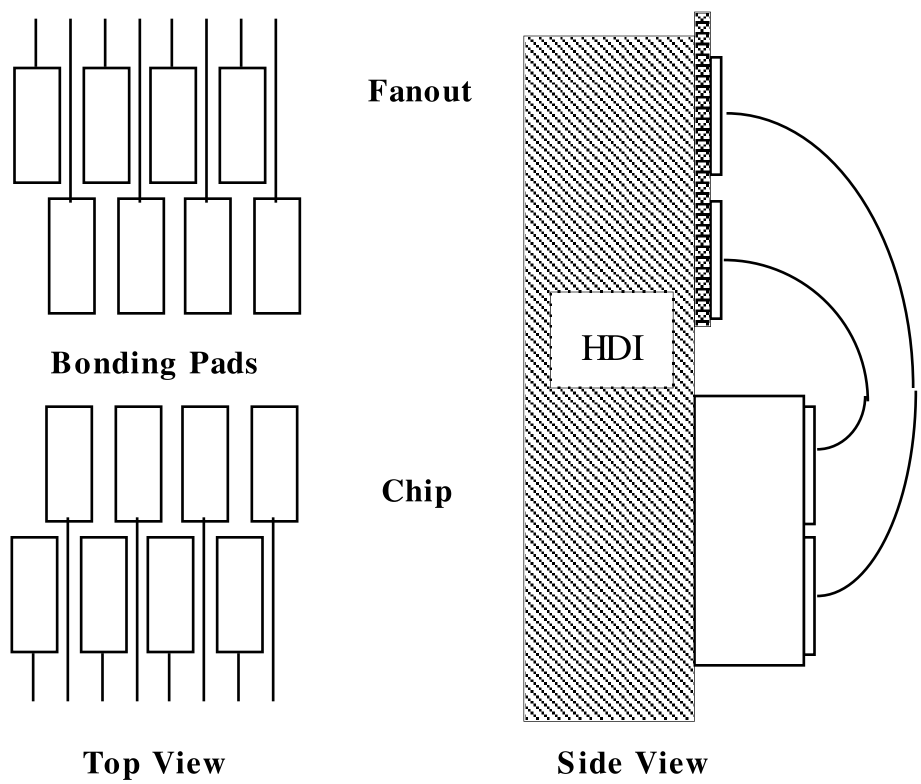


Figure 11 : Drawing of the staggered bonding pads on the AToM chip and the detector fanout. Chip height and fanout/chip gap not to Scale!

however, measurements are currently underway using an HDI substrate to test several chips simultaneously.

3.4. Data Transmission

3.4.1. The Mux-Card (Data and Command Multiplex Card)

The data and command multiplex card (Mux-card) is located outside of the detector, where more space is available. It is responsible not only for collecting data from 16 serial data links, corresponding to eight readout sections, and multiplexing all of those onto a single fiber-optic link. It also demultiplexes the command stream from the incoming fiber link into 16 command streams for the individual readout sections.

To achieve full depletion of the detector a voltage of 60 V is applied between the detector ϕ and Z side. To minimize the effect of a possible short between a Si-sense strip and the aluminum coupling layer on the detector the readout chip input reference is coupled to the depletion voltage on each side. Thus the ground planes for the electronics responsible for opposite detector sides have a voltage difference of 60 volts. The Mux card has to accept signals coming from both sides of a detector, unify them to a common signal level, and send them out on a single line, so the entire data of one detector is available in one place.

Up to 16 mux-cards physically reside in a VME crate. All settings on the mux card are programmed via the VME interface, the mux card does not decode any part of the command or data stream that is transmitted on the fiber link. Only a single chip without a detector was used in the test setup, therefore no Mux-card was needed.

3.4.2. The Read Out Module (ROM)

The Read Out Module (ROM) consists of two separate units: The Personality Module and the Data Acquisition (DAQ) board. The Personality module contains the data and command interface to the detector front end, while the DAQ board contains the memory and data processing facilities, along with the interface to the higher level readout system. Figure 12 shows a block diagram with the major functional blocks of the ROM. A more detailed block diagram can be found in the appendix.

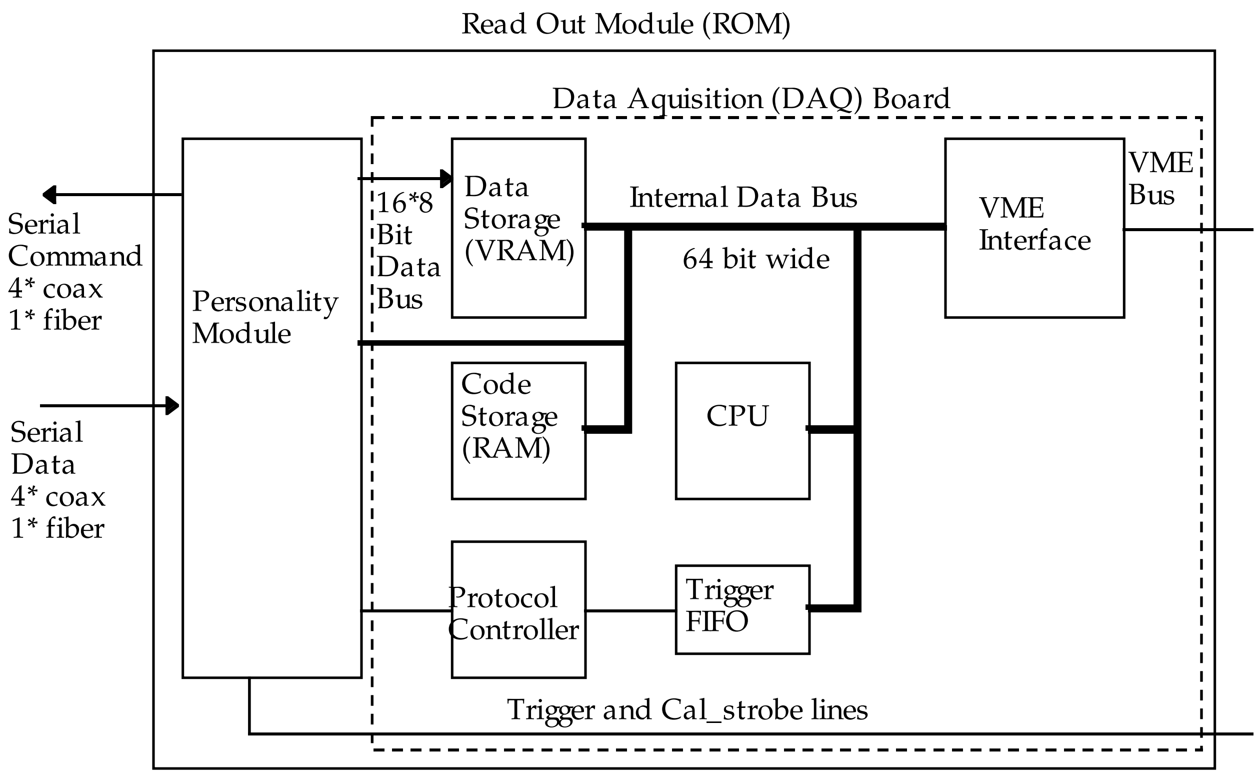


Figure 12 : The Read Out Module Block Diagram

3.4.3. The Data Acquisition Board (DAQ board)

A very detailed description of both the prototype DAQ board can be found in [11]. This description also includes some additional information on the personality module. Here the main parts of the DAQ board will be described, in particular those important to the experimental setup. A more detailed description of this complex board is inappropriate for the scope of this thesis.

3.4.3.1. Major Components of the DAQ board

This section lists the major functional blocks from Figure 12 (except the personality module) and explains each block in suficient detail to understand the basic functionality of the DAQ board.

1. **The Personality Module Data Bus**
   The connection to the Personality Module is a fully synchronous 128-bit wide bidirectional bus. The bus direction is controlled by the DAQ board, while all transfers are run by the Personality Module. Bus speed is variable, up to 30 Mwords/s.
2. **The Personality Module Command Bus**
   The Personality Module registers are accessible via a dedicated 16 bit wide bus. This bus is connected directly to the lower 16 bits of the processor data bus.

3. **Event Data Storage**
   The event data is stored in 4 Mb VRAM. VRAM stands for Video Random Access Memory, a type of memory typically used in video cards on personal computers. This memory has two sides, the RAM side, on which standard random access read or write operations can be performed. This side is connected to the DAQ board processor data bus. The serial access memory (SAM) side allows access of contiguous locations only. Read and write operations may not be mixed, a rather complex operation is necessary to switch from serial input to serial output mode. The serial memory resides in a physically separate part of the VRAM chip. Contents of the serial memory can be transferred in a refresh like cycle into the RAM side, or vice versa. In this application the VRAM is used as a simple dual ported RAM.
4. **The VME Interface**
   The board has a VME slave interface. This interface is used to send processed data off-board. A commercial Readout Controller (ROC), typically a standard high performance single board computer, is used to collect the data from all the DAQ boards and make it available for further processing. The ROC can access all features of the DAQ board, including the entire memory range, without a DAQ resident processor. This was used during debugging to prevent the necessity of writing programs for the DAQ board processor. Without any operating system this would have been extremely difficult. VME access is very slow compared to any on board functions, but during debugging and testing the performance was not an issue.
5. **The DAQ board Processor**
   The DAQ board can be equipped with either a 100 MHz (internal clock) PPC 604 processor, or with a 130 MHz PPC 603e processor. A DAQ board without processor is also fully functional as a data-taking device. We did not use the processor on the DAQ board in the test stand. Instead processing was performed in a commercial CPU board inside of the VME crate. This CPU board is the equivalent to the readout controller in the experiment. It is therefore commonly referred to as the 'Readout Controller' or 'ROC'.
6. **The Fast Control Interface**
   The DAQ board also has a custom interface to the fast control system. The fast control system will be used to send time critical commands, such as triggers, to the front end. In the prototype system the only implemented fast control functions are trigger and calibration strobe. Both of these commands have a dedicated line on the prototype.

3.4.3.2. Mechanical specifications of the DAQ board

At least 16 DAQ boards with their personality modules have to fit into a VME crate. To avoid building customized extra wide crates the prototype has been designed as a single width 6U VME card. The DAQ board physically carries the Personality Module, so the entire Read Out Module (ROM) consisting of both DAQ board and Personality Module form a unit that fits into a single slot of the VME crate.

3.4.3.3. DAQ board Design

The main purpose of the DAQ board is to processes, reformat and buffer data. The large VRAM banks buffer many events before processing has to be completed. Therefore the fluctuations in the trigger rate do not affect the system performance. A high instantaneous trigger rate for a brief period will not make throttling of the trigger necessary. Only a continuously high trigger rate will eventually fill the buffer.

The production version of the DAQ board will run the VxWorks operating system on-board. This high performance real time operating system is an ideal code development platform for embedded systems such as the DAQ board. In the test stand the code development took place on the readout controller, which also runs under the VxWorks operating system.

The design of the DAQ board is held simple by use of one large Field Programmable Gate Array (FPGA), which controls all logic functions. The advantage of the FPGA as compared to discrete logic, or a combination of simple Programmable Logic Devices (PLDs) is the high concentration of logic, allowing for relatively few other components

The FPGA is also a good choice to make the board more flexible, and to reduce development time. The FPGA can be easily reprogrammed when logic errors are found on the actual board, reducing the amount of time that is needed during simulation. Even additional registers and other additional functionality can be implemented, as long as the external connectivity remains the same.

The FPGA is programmed using VHDL, a C-like hardware description language. The modular structure of this language makes it easy to modify the design even by people unfamiliar with the entire board. Just a small fraction of the functionality needs to be understood to be able to add or modify the current functionality. This language is also invaluable during the simulation phase. Signals are much easier to trace in code, then they would be in a gate-level design description.

3.4.4. The Personality Module (PM)

The Personality Module has a dual interface to the front end electronics. It has a total of 16 data and command channels, all of which are multiplexed onto a single high speed fiber link. On the prototype four of those channels are also connected to standard coaxial connectors. This allows us to use the PM with the final Mux-card, using the fiber link, with a prototype Mux-card without fiber link, and with a primitive test bench setup with direct wiring to the chip test board. The final version of the personality module will be built without the coaxial connectors.

3.4.4.1. The Fiber Link

A fiber link has been chosen to connect the Personality Module to the Mux card. The entire data stream to and from the detector is handled on this interface. It transmits commands, setup and calibration data, and receives event data.

A fiber link was chosen because the high throughput rate allowed the use of just one fiber per 8 readout sections (16 data/command channels). This means cable cost per distance is much lower than if dozens of twisted pairs are used. Additionally the fiber link is expected to be much more robust, mostly because the number of connectors is dramatically decreased.

The 59.5 MHz clock signal is encoded on the data line, further minimizing cable cost. The transmission of the clock on the fiber link has the additional advantage of a very small clock jitter, because the high speed fiber receiver needs an extremely accurate clock to validate the ultra high speed data. When this high speed internal clock is then divided down to produce the output clock asymmetry and jitter on the output will be naturally very low.

3.4.4.2. The Coaxial Connectors

To support 4 channels of command and data a total of 10 coaxial connectors is needed. 4 command output lines and a clock output, along with 4 data input lines, and a data validating clock input. The clock output and clock input can run completely asynchronous, but will usually be run at the same clock speed. In that case the clock output can be used to drive the clock input, but care has to be taken to assure proper data/clock phasing. Data on the input lines should be valid 4 ns prior to the rising edge of the input clock.

3.4.4.3. Functionality of the PM

The main purpose of the Personality Module is to act as a serial to parallel interface with a 16 to 1 multiplexer. This that means the module converts the two fibers (one input, one output) to a 8*16=128 bit wide data bus. This ultra-wide bus keeps the bus speeds on the board at an manageable rate while permitting large data throughput. The 4 coaxial data and command lines represent the lowest 32 bit of the 128 bit wide bus.

Commands can be sent to the front end in two different ways: The current PM prototype has two dedicated lines, one for the trigger, the other for the calibration strobe command. An active low pulse on one of those lines causes the respective command to be sent to the front end system. This is an interim solution, as no trigger system is available. The final version will receive commands on a serial line, and pass them on to the front end without decoding.

Alternatively commands can be sent by writing into a specific register on the Personality Module. Another register on the PM determines how many (if any) words of data are sent down to the front end, attached to the command. This data is necessary for the setup of the readout chip. All programmable registers of the personality module, as well as their functions are listed in Table 3.

The Personality Module handles the protocol management for the data reception from the front end chips, decoding the data header (a single start bit) and the data trailer (a sequence of 32 zeroes). It checks for timeout violations, along with fiber link status errors. The PM has full interrupt capability to notify the DAQ board processor of any such violation.

Table 3 : Personality Module Registers

The Personality Module has the following programmable registers:	
Timeout	Contains the maximum number of bus clock cycles between begin and completion of a read event.
Tx count	Contains the number of bytes that will be sent to the front end.
Rx count	Contains the maximum number of bytes the personality module will store. The data stream after this point is truncated, but the read event is not considered complete until all channels are silent.
Data mask	Masks off any of the 16 input channels.
Command mask	Masks off any of the 16 output channels.[2]
Irq mask	Masks off any of the 5 interrupts that would otherwise be sent to the DAQ processor.
Irq reset	Writing a one to a bit location in this register clears the corresponding interrupt. Writing a zero has no effect.
Opto register	Gives access to the fiber link chipset internal settings.
In addition to the programmable registers there are read only registers:	
Irq register	Contains a mask indicating which of five possible conditions caused the interrupt.
EOT register	Contains a mask of all channels that caused a timeout. It is automatically updated each time a timeout condition occurs.

3.4.4.4. Personality Module Design

The design is held simple by use of one large Field Programmable Gate Array (FPGA), which controls all logic functions. The advantage of the FPGA as compared to discrete logic, or a combination of simple Programmable Logic Devices (PLDs) is the high concentration of logic, allowing for very few other components. The current design contains just a clock driver, some ECL to TTL converters and the fiber link chipset apart from the FPGA.

The FPGA is also a good choice to make the board more flexible, and to reduce development time. The FPGA can be easily reprogrammed when logic errors are found on the actual board, reducing the amount of time that is needed during simulation. Even additional registers and other additional functionality can be implemented, as long as the external connectivity remains the same. In this way the prototype personality module can actually be reprogrammed to support all the features the final version needs.

The FPGA is programmed using VHDL, a C-like hardware description language. The modular structure of this language makes it easy to modify the design even by people unfamiliar with the entire board. Just a small fraction of the functionality needs to be understood to be able to add or modify the current functionality. This language is also invaluable during the simulation phase. Signals

---

[2] On the prototype personality module there is a combined mask for all 16 I/O channels.

are much easier to trace in code, then they would be in a gate-level design description.

4. Measurement Overview

The goal of the set of measurements performed was to fully characterize the entire chip. This includes both a functionality test, and a measurement of all essential chip parameters.

To better understand the single measurements, and the order in which they were done a short summary of all the measurements, their goals, and the problems involved in the experimental setup and the data interpretation will be given.

4.1. Calibration Measurements

In the first measurement the offset, step size and linearity of the threshold digital-to-analog converter (DAC) was measured. All three parameters can be extracted from a simple measurement of the threshold DAC voltage for each DAC setting. A linear regression of those data pairs gives the offset (intercept) and step size (slope). The linearity can be calculated from the quadratic sum of the errors. The threshold voltage is essential to convert later measurements that have results in DAC counts into units of mV.

The second measurement is very similar to the first. The calibration DAC offset, stepsize and linearity was measured by the same process as above. After both the DAC voltages are known the only other unknown on the chip that prevents us from converting chip units to SI units is the preamplifier and shaper gain. This can be directly measured by injecting a small, known charge into the preamplifier input. The output of the shaper is not accessible, so a simple analog measurement of the voltage at the shaper output is impossible. Therefore the on chip comparators were used to compare the shaper output to the threshold voltage, while a threshold voltage scan over the entire threshold voltage range was performed. At high threshold values the comparator never fires, while it is constantly on at low values. The 50% or 'turn-on' point of this curve is the point at which the threshold is just as high as the shaper output.

The turn-on point however has to be well defined. In an ideal system with no noise the comparator would fire every time the input signal is above threshold, and never when the signal is below threshold. The result would be a perfect step function for the firing probability of the comparator. Under non-idealized circumstances however there is a thermal noise. The noise charge simply superimposes on the signal charge, so it can both cause a below threshold signal to fire the comparator, or cause an above threshold signal to not fire the comparator.

In order to find a function that can be fitted to the data assumptions about the noise spectrum have to be made. The noise amplitude spectrum is assumed to be gaussian, which correspondingly gives an error function for the turn-on probability function. Comparison to the data shows that an error function is a good fit. The program used to fit the error function to the data can not use a simple χ^2 fit, because a χ^2 fit assumes that for each data point the measured values are distributed around the true value according to a gaussian distribution. The 0% and 100% probability limits, however, are hard cut-offs, meaning that values below 0% or above 100% can never be measured. This deforms the symmetric gaussian to an asymmetrical binomial distribution for values close to either cut-off. Therefore a maximum likelihood fit is used to fit the error function to the data set. A maximum likelihood fit maximizes the probability that a measurement of the fitted function would result in the actual data that was measured. The overall probability for a measurement resulting in the actual data is the product of the probabilities for measurements of each data point to result in the actual data. This product is very hard to differentiate. Therefore the fitting algorithm finds the zero of the differentiated logarithm using a standard Newton iteration. This point gives the maximum probability, because the logarithm is a monotonic function. The errors are estimated by the smallest error rectangle that fully contains the error ellipse.

The threshold DAC has a negative slope, which can be very confusing during data interpretation. Therefore the fit program inverts the slope by inverting the DAC setting number. Thus a low fit value now corresponds to a low threshold, a fit value corresponds to a high threshold.

After finally deriving the turn-on point in DAC settings it is easy to convert to an absolute value using the above mentioned measurements. Thus the gain of the input section can be derived in mV/fC.

4.2. Performance Measurements

As the first in a series of measurements that use the chip itself as a measuring instrument the capacitance of the internal injection capacitors is determined. By comparing the shift of the turn-on point per injected charge (measured above) to the shift per internal calibration DAC step, one can calculate the capacitance.

The next important measurement is the pedestal variation of the individual channels. The pedestal is defined as the turn-on point at no injected charge. Usually the pedestal is not directly measured by measuring the turn-on point at zero charge, but by measuring the turn-on point at several known charges, and then extrapolating to zero. This type of measurement makes it possible to accurately measure the pedestal even for channels that have a below zero pedestal. These channels would never fire without injected charge, thus preventing a pedestal measurement.

Two values are of interest: the RMS variation of the pedestal and the maximum peak to peak pedestal variation. It is desirable to have a low pedestal variation, because there is just one threshold setting for the entire chip, so a large variation would cause inefficiency for low charge hits in some channels, while other channels would be subject to noise hits.

Another very important measurement is the inherent system noise. The noise is defined as the RMS of the noise amplitude spectrum (assuming again that the spectrum is a gaussian). This parameter can be extracted from a threshold scan as the RMS of the fitted error function. The preamplifier noise is linear with the input capacitance of the system (see section 2.2.3 c). The total system noise is dominated by the preamplifier noise, so the system noise is roughly linearly dependent on the input capacitance. This means that the additional detector capacitance will be the dominating factor. The test bench system did not allow an easy change of the input capacitance, so the noise was measured for only two different configurations: no external capacitors, and 15 pF capacitors to both neighbors. The exact dependency on capacitance has been measured in Pavia [12].

Three more measurements are necessary to define all the essential chip parameters: The first one is the measurement of the transfer function of the time over threshold analog to digital converter (ADC). The exact form of the transfer function does not have a closed analytic equation. As mentioned above the transfer function has to extend the dynamic range by compressing large signals. This leads to a roughly logarithmic response. The exact shape of the function is not critical to chip operation, as long as it is monotonic.

The next measurement examines the shift in timestamp due to a variable peak height. The timewalk so introduced could be reduced using analog techniques, such as constant fraction discrimination, but the space available on the chip does not permit this. The true timestamp value can be estimated later on using both the measured timestamp and the ToT value.

At last the channel to channel influence was measured. There are two contributions, the crosstalk, which is largely caused by capacitive coupling between neighboring channels, and the change in the calibration charge injection voltage when the injection circuitry is loaded by simultaneously injecting charge into many channels.

5. The Test Bench Setup

The Test apparatus can be divided into several parts: The test stand software, UNIX workstation, DAQ board, Personality Module, test board, and the setup, including wiring, clocks and power supplies.

5.1. The Software

The test stand software is based on a very user friendly and flexible VME access package called Oasis, which was developed at the University of New Mexico [13]. The Oasis software package is split into three very distinct parts: The front end C code, running on the Readout Controller (ROC) system, the back end C code, running on a Sun workstation, and the graphical user interface, programmed in the Tcl/Tk scripting language [14]. The functional design of the Oasis code is ilustrated in Figure 13. The user enters commands or starts measurements by interacting with the graphical user interface (GUI). The interface script calls C routines running on the workstation to access the DAQ board. The C routines on the back end call proxy C routines running on the front end via a Remote Procedure Call (RPC) (see 5.1.2). The C routines on the front end use the hardware access routines supplied by the VxWorks operating system to access DAQ board features.

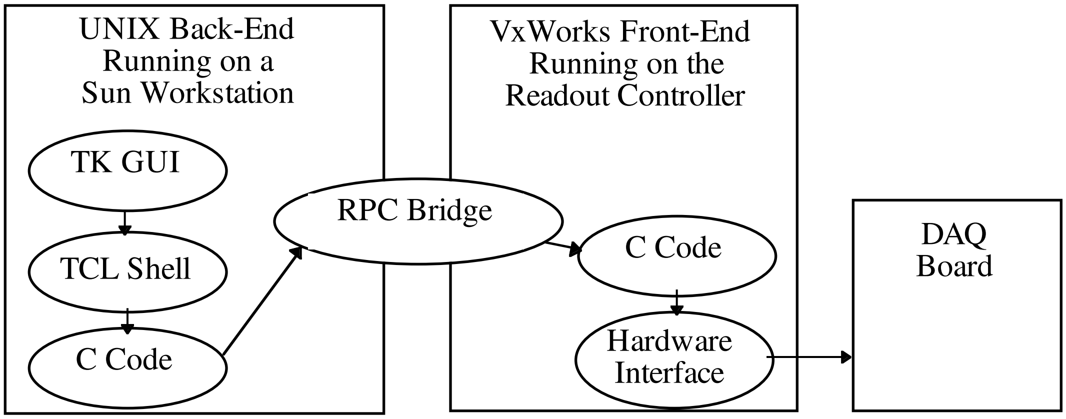


Figure 13 : Functional Diagram of the OASIS Software

5.1.1. The Foundation -- The Oasis Software Package

After some deliberation we decided to adapt existing code to our hardware platform. This would give us a more flexible, modular and user friendly program with no more effort than a primitive new program would have cost. The Oasis code was chosen as a template for our system because it is flexible by design, so adapting to our hardware was comparatively easy. It supports a graphical user interface, so it is easy to learn how to use. This is particularly important for a test system which is used by many different international groups, where some of the groups use the software just for a short time to test their hardware during an assembly process.

The original Oasis code was designed to access any kind of registers or memory, including first-in first-out memories (FIFOs). This made the code attractive to us, as all features on our board were simpler than the ones foreseen by the Oasis code designers. We did not need to add any code in the VME access section. On the other hand it did not include any actual measuring routines, so there was no reasonable way to examine and process the data.

The software was adapted to our specific hardware platform, some of the memory access code was edited to better suit our purposes and additional, more complex functions were implemented, that combine a set of standard VME memory and register operations to communicate with the SVT readout chip through the DAQ board.

We added automated testing and measurement capabilities that were written specifically for the SVT readout chip in conjunction with the Read Out Module (ROM). These routines would perform a series of writes to a set of registers to set up the ROM for communication with the front end chips. Similar routines for other subsystems can easily be written using our templates.

5.1.2. The Data Transmission Structure

The Oasis code connects a Tcl/Tk based graphical user interface running on a standard UNIX workstation with the ROC running VxWorks in the VME crate (see Figure 13). This connection is done via remote procedure calls (RPC), a standard software protocol that is typically used for command and status messages rather than actual data transmission. Nevertheless it is used in the current version for all data transfer between front and back-end.

Data transmission from back-end to front-end looks like a standard procedure call, with the only difference that no pointers can be passed as arguments, because there is no shared memory space between the front-end and back-end processes. The front-end processes can return only one variable upon completion. The version of the RPC protocol used allows the passing of any size of structure between the front and back-end. Therefore all data that is returned to a certain calling routine has to be grouped in a single structure, and the entire structure has to be transferred via RPC. So the only drawback remains that the maximum transfer size has to be determined at compile time by defining the structure size.

5.1.3. Software Design Philosophy

5.1.3.1. Portability

The code was designed to be easily portable between multiple back-ends and front-ends. Portability depends mostly on availability of all the necessary libraries and the operating system dependent RPC preprocessor which is responsible for automatic code generation for the RPC bridge. Portability onto a new platform is very straightforward; it should be possible to transfer the entire code to a new platform in a week if all necessary libraries are present. Currently the code runs on Sun, SGI and DEC/Alpha computers.

5.1.3.2. Flexibility

A second point during code design was optimal flexibility. The Oasis code can be adapted to any custom VME slave board with very little work. As soon as the VME board has been fully defined within Oasis all features can be used to access the board.

This flexibility is largely possible due to the use of definable access methods for any register. The standard implemented access methods can deal with all normal registers, memory, or FIFOs. Additional access methods can be defined for extremely unusual devices.

5.1.3.2.1.VME Access Implementation

A standard read command in Oasis does not only look up the register address in the board definition, but also its defined access method. It then uses this address and method to access the register or memory. A typical access method for a register masks out all non relevant bits, and shifts the data appropriately. A typical access method for a FIFO will read one or more words from the FIFO, then format these properly into a data structure that can be returned to the calling routine. The FIFO full and FIFO empty bits are checked before each access, preventing reading from an empty FIFO, or writing into a full one. The access method for memory allows reading a block of contiguous data. The access function first checks if the entire block lies within the memory boundaries, so overwriting of memory boundaries is impossible. The board definition allows virtual memory boundaries, so code memory can be effectively shielded from data memory. The general idea of Oasis is to include all functionality that might be needed, so no patching is necessary once the system is completed.

The program was designed to be both user friendly and easily maintainable. The user friendly interface was designed as a fully graphical interface using the Tcl/Tk scripting language. Register access fields can be added during runtime using a menu that contains all registers that are defined in the board definition file. So by simply writing the board definition file the new board is already included in the graphical user interface.

To make the software easily maintainable two directives have been implemented. First the program is partitioned into easy to understand modules, so any modification requires knowledge of a minimal part of the system. Secondly a high degree of error reporting and debugging information is accessible to the user. The program has a global debugging variable which causes, if set, comprehensive debugging information to be printed on the front end console.

5.1.4. Levels of Software Functionality

The user interface supports several levels of VME access. The user can choose direct, low level routines, which cause a simple, and easily debuggable VME accesses. Alternatively he can use higher level routines which provide a more intuitive interface to the board functions.

5.1.4.1. The Board Debugging Tools

The lowest level are the board debugging tools, such as the 'VME' and 'Test Stand' windows (see Figure 14). The 'VME' window allows low-level VME access to any VME location. The user enters the VME address, no address translation is performed. This window supports D32/A32 access only, no complex access methods are involved. Only a single word is read or written.

The 'Test Stand' window gives access to all memories and registers on the board. The registers, memories or FIFOs are listed by name, not address, so no knowledge of the board is required to use this window. The number of registers and the order in which they are shown can be customized using the window's menu.

5.1.4.2. SVT Specific Features

The more specialized commands are accessible through three chip testing windows (see Figure 15): the 'Send Command' window is the most primitive chip specific window. It is used to monitor the chip's reaction on single commands. E.g. sending a 'read register' command gives the opportunity to check the entire signal path with an oscilloscope. The command can be monitored as it is sent out on the serial command line, and the chip's response can be verified, making it possible to pinpoint data or command transmission problems. The window supports the sending of data to set control or mask registers on the readout chip, but this is awkward, because the user has to encode the command register bits by hand. This window is used more to verify and debug the entire setup, rather than to test the chips. Still it remains a fast way to find obviously bad chips.

The 'Register Test' window gives easier read and write access to all chip registers. It can be used to manually check a suspicious bit. The single read or write

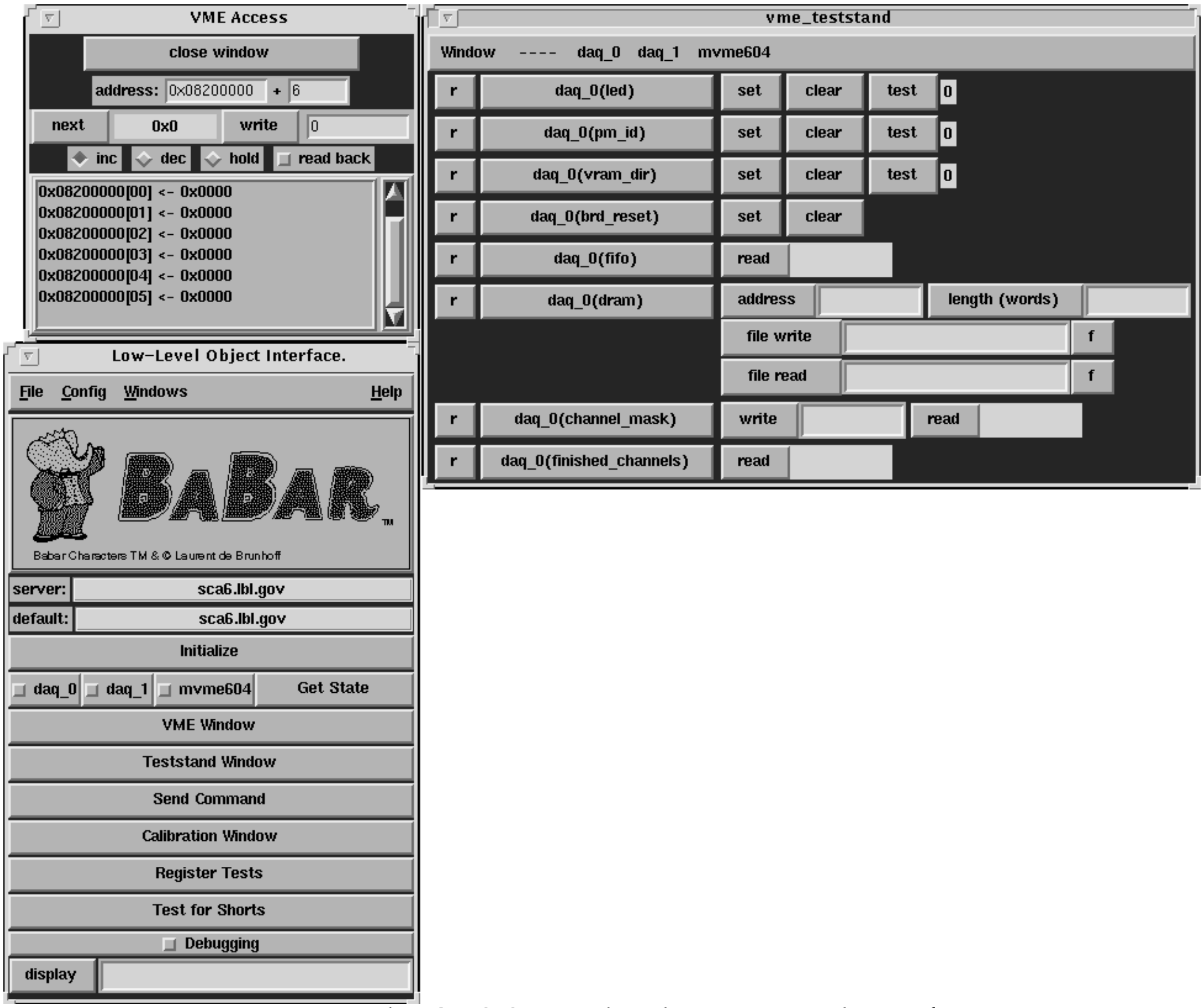


Figure 14 : The OASIS Board Debugging Tool Interfaces

function makes it possible to analyze the data stream with an oscilloscope. There is also a test that walks 1's and 0's through all register positions. This is the fastest way of checking the entire register set.

The 'Calibration' window supplies a succession of cal_strobe, trigger, read event (see section 6.2.8 for details) and reads the event data from VRAM. It properly formats all the data, and interfaces to a histogramming package. The chip setup options are very flexible:

- The calibration DAC can be varied.
- The threshold DAC can be varied.
- Both can be varied simultaneously, giving a two dimensional histogram.
- Any readout channel can be masked off individually.
- The charge injection can be masked of on any channel.
- The data taking run can be suspended at any point in time.

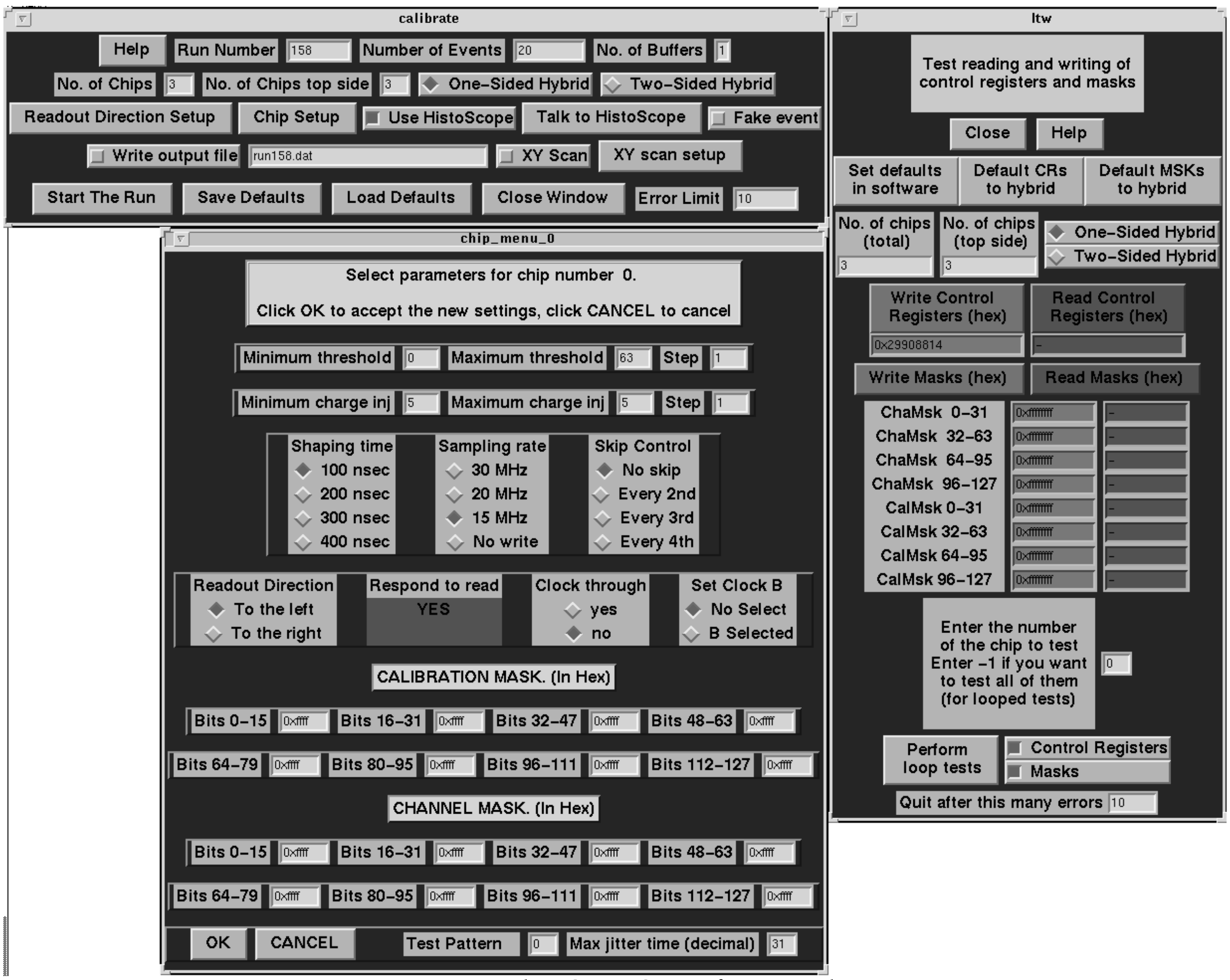


Figure 15 : The SVT Specific Windows

5.1.5. Data Visualization -- The Histoscope Package

Histoscope is a complete histogramming package with emphasis on user friendliness and intuitive handling. It was developed at Fermilab, and is available as freeware for most UNIX platforms.

Data can be stored in data structures (ntuples) or histograms. Histoscope features simple or 2D histograms, both of which can be defined in the program or be filled from an existing ntuple during runtime. The ntuples allow one to introduce and refine cuts performed on the data set before displaying as an histogram. Scaling is automatic, but manual zooming is possible. The orientation of the 3D plots can be controlled and rotated with the mouse. This makes it much easier to understand the structure of complicated plots. The entire user interface is graphic oriented, and bears a resemblance to the Oasis user interface, making it easy to use both programs in concert.

5.2. The Hardware

The test stand hardware consists of all the equipment necessary to read out the AToM chip: the UNIX back end workstation, the readout controler, the readout module consisting of DAQ board and personality module, the test board and external suport logic. For this test only a single chip without a detector connected to

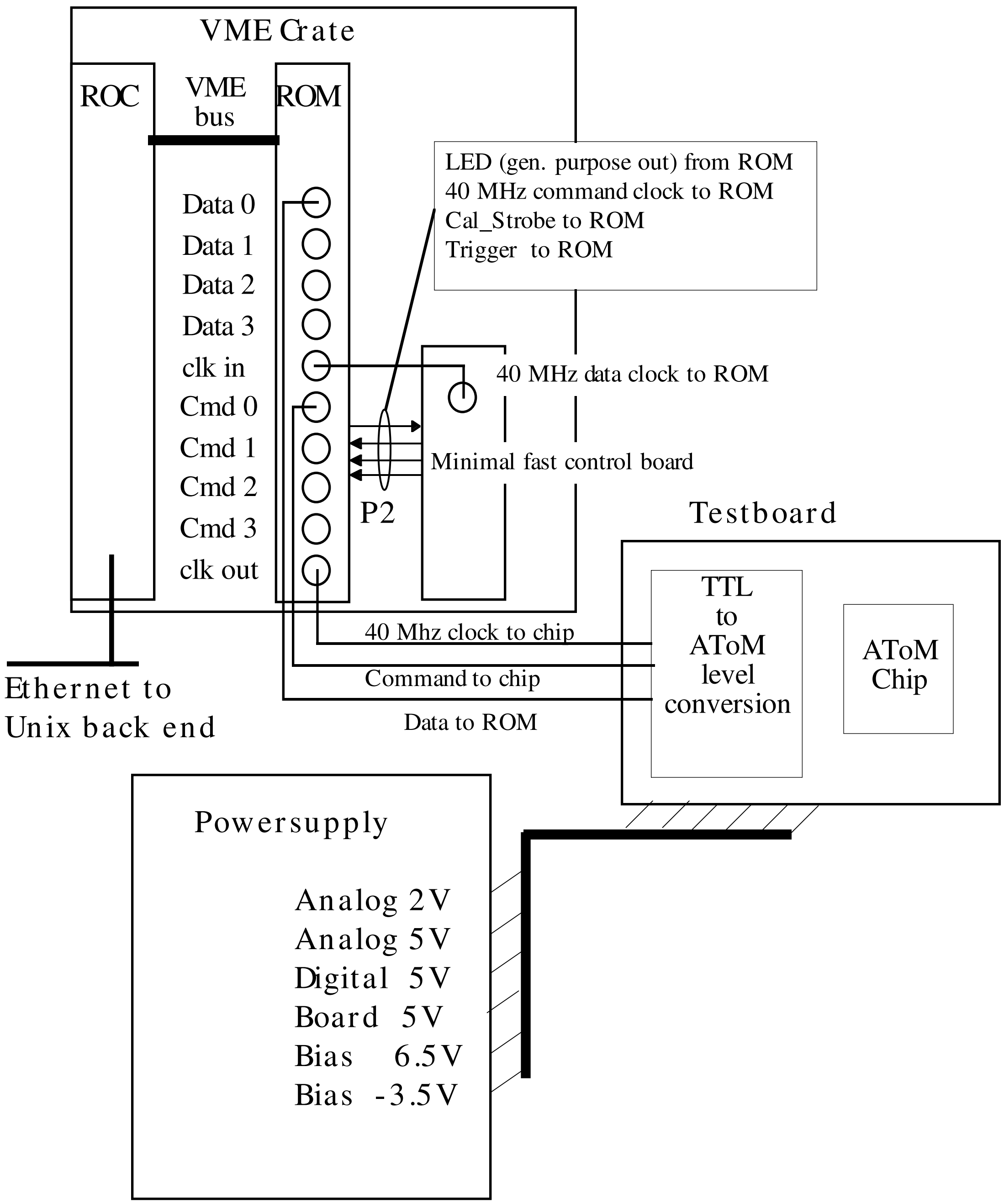


Figure 16 : The Test Bench

it will be tested. To date no detector has been bonded to the chip, first tests with a detector are expected in early 1997.

The current chip prototype has multiple problems above clock speeds of 45 MHz. All of the known problems have been traced to design errors, and will be corrected in the next chip version. We decided to run all characterization measurements on the prototype at 40 MHz. This is adequate to characterize all the chip parameters.

5.2.1. The Test Board

The test board was designed at LBL for testing the functionality of the SVT readout chip both on a very low level using a pattern generator as input device and an oscilloscope to analyze the data, as well as on a system level, using the ROM both to send and receive data

The test board held a single SVT chip, which had only 10 of its 128 input channels bonded to external inputs. This allows for a comprehensive test of the chip's functionality, without overly complicating the board, or increasing the data size beyond manageable levels. However, the chip to chip communication can not be fully tested with any board that contains just one chip. A second board designed at UCSC was used to test the functionality with several chips.

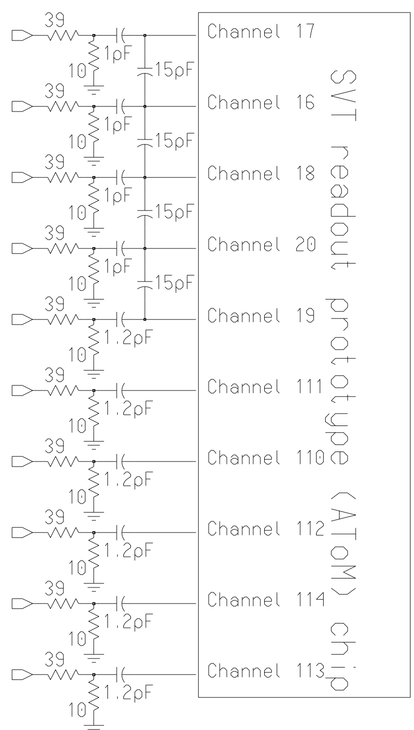


Figure 17 : Schematic of the Bonded Chip Input Channels

The ten bonded input channels are ordered in two groups of five consecutive channels (16-20 and 110-113). Due to the interleaved bonding pads the order of the input channels is scrambled (see Figure 17). On one of the groups, 15 pF strip to strip capacitors were installed to simulate the detector capacitance. On all lines either 1 pF or 1.2 pF capacitors for external charge injection were connected. Only four precise (10% tolerance) 1.0±0.1 pF capacitors, and five high precision 1.2±0.12 pF capacitors were available. A sixth 1.2 pF capacitor with a higher (20%) tolerance had to be used on one of the lines due to the shortage of high precision capacitors.

5.2.1.1. External Charge Injection

The internal calibration capacitance on each channel has a precision of 30% with respect to its design value of 0.05 pF. In a single chip production run, and especially on a single chip, the channel to channel variation is much smaller. We expect a variation on the order of 5% over all channels on a single chip.

Therefore the internal charge injection capacitor is not well suited to give absolute calibration values for the preamplifier gain. However, once an absolute calibration is established on one channel it can be used to calibrate all channels with relative accuracy. To get an absolute calibration the external capacitors were used to inject a well known charge into the preamplifier input.

The high input sensitivity of the preamplifier requires a charge injection in the 10^{-15} Coulomb range. With the external capacitors of ~ 1 pF an injection voltage between 0.2 mV and 4 mV is needed. This small voltage step is achieved by using a 1/100 resistive divider and a standard pulse generator.

Larger capacitors would not only have reduced the voltage level even further, they would also couple external noise in with a higher coupling factor. The 1 pF capacitors seem to be a good compromise between a large, precise capacitor, which reduces measurement precision by introducing additional noise, and a small capacitor which reduces precision due to its unknown capacitance.

5.2.1.2. Level Conversion

The main purpose of the test board is to convert the custom low swing bidirectional chip levels to standard TTL levels and vice versa. Additionally there is some power filtering done on board. The chip is sensitive to power supply ripple, especially on the analog +2V line. The analog +5V line is less critical, and the digital +5V is non critical.

The board allows either the default or backup clock and command lines to be selected (see section 3.3.2.1). Both default and backup clock and command lines can be wired together, to allow the programmable chip register to determine which line is used. When the lines are not wired together the clock and command lines that are driven are selectable by a jumper pair. Only the default clock and command input were used after checking the functionality of the other set of inputs. This reduces the number of changes to the setup to a minimum during the data taking period.

5.2.2. The Test Bench Wiring

The connection to the personality module is made with standard 50Ω coaxial cables. Differences in cable length are used to adjust clock/command phasing. The cables are terminated with 100Ω terminators on the board, because the driver on the personality module is not strong enough to drive a proper 50Ω termination. The 100Ω termination gives adequate reflection attenuation.

The clock/command phasing should be adjusted so that command changes slightly after the falling edge of clock when those signals arrive at the test board. An error in the command polarity on the board requires cross jumpers on the command a/b select lines. If the clock is also cross jumpered command phase has to be adjusted accordingly.

Clock phasing for the receive data clk on the personality module should be such that data is valid about 4 ns prior to the rising edge of clock. This is not a very strict requirement, the personality module seems to receive data properly over a fairly wide phase range (~ 8 ns).

5.2.3. The Personality Module

I was strongly involved in the design of the personality module. The hardware was designed elsewhere, but only a conceptual design existed for the logic inside the field programmable gate array (FPGA). I was responsible for the entire development, optimization, testing and debugging of this logic. This required extensive simulation and actual hardware testing.

The Personality Module that was used in the test bench had no fiber link. The chip test board was connected using the coaxial connectors. Only one channel was used, the other channels were tested just superficially. During further tests with the HDI (see section 3.3.3) all four electrical channels will be used (two channels per side).

5.2.4. The DAQ Board

The DAQ-board used in the tests bench is the DAQ board prototype, designed at LBL. I was responsible for testing and debugging of the essential functions during setup of the teststand.

During debugging a persistent bug in the VRAM controller became apparent. This bug allowed only 512 words to be written to the front-end in a single transfer, the read function, however, was not affected. The entire VRAM can be filled with a single transfer. This limitation has no effect on the ability to communicate with the SVT readout chip, because the longest command that needs to be sent to the readout chip is the write mask command, consisting of 32 words.

All three varieties of boards were tested: without processor, with a PowerPC 603, and with a PPC 604. All of them performed equally well. In the test bench a board with a PPC 604 was used. However, the operating system had not yet been ported to the DAQ board, so we did not use the on board processor, but executed all our programs on the ROC.

5.2.5. The Read Out Controller

The Read Out controller (ROC) is a Motorolla 1604 VME single board computer. It has an internal 33 MHz data bus and a 99 MHz PPC 604 Processor. The memory is configured as a single 8 Mbyte block of non-interleaved DRAM [15].

The software running on the ROC is the VxWorks operating system (version 5.2b) [16] and the Oasis front end code (as described in section 5.1), including the RPC server to connect to the UNIX workstation with the user interface. The read out controller is mounted in a desk-top VME-crate. It is connected to the UNIX back end workstation via a standard ethernet. The additional serial port with a terminal was used to configure the system and to monitor the automatic booting.

As a UNIX back end a Sun workstation running the Solaris operating system was used.

5.2.6. The Minimal Fast Control System

The SVT readout chip looks for a pulse in a comparatively small trigger window (max. 2µs). This window is 18 µs (at 40 MHz chip clock frequency) after the pulse arrived at the input. This timing of signals was accomplished by way of a software controllable and user accessible bit on the DAQ board hooked up to the trigger input of a pulse generator. Upon toggling of the LED bit this pulse generator would generate a 100 ns wide pulse to the Cal_strobe input of the PM.

Exactly 18µs (at 40 MHz) later a second pulse is generated by the pulse generator, this time to the trigger input of the PM. Thus the PM sends those two commands with exactly 18µs spacing. The delay can be easily varied to check the accuracy of the time stamp counter on the readout chip.

A second pulse generator is set up as system clock for the DAQ board. It has two 40 MHz outputs, one connected to the P2 backplane input of the DAQ board. This is a non terminated TTL input to allow bussing of several DAQ boards on the same clock line. Some experimenting showed that the on board clock signal (after the clock driver) looks best when there is no termination on the cable, but a 40% duty cycle is selected.

The other clock output is 11 ns delayed, and then used to latch the data coming from the test board. The entire clock distribution seems to be very stable to minor variations in duty cycle, delay, frequency or termination, except for the clock to command phasing on the test board.

Later on I built a wire-wrap board taking over the functionality of this minimal fast control system. This board is plugged on the rear P2 connector, so it shares all P2 pins with the DAQ board. It has an additional coaxial output, which is connected to the clock input on the DAQ board front panel with a 10 ns cable to latch the data from the test-board. The 18µs delay is implemented as a 10 bit counter with two comparators to send the trigger and cal_strobe pulses to the DAQ board. The board will be referred to as the 'clock board'. This setup is less flexible, but more compact and user friendly, so we expect to use it in the upcoming system test. The schematic of this board can be found in the appendix.

5.3. Test Procedures

The most critical point in the entire test procedure was the fact that nothing could be adequately tested without the DAQ board. It would have been much simpler to determine the cause of any malfunction if each part of the system could have been tested alone.

The DAQ board was tested alone, first for VME I/O capabilities, memory and register operation. Then the personality interface was tested using a pattern generator and a logic analyzer. As a last test the processor functionality on the DAQ board was verified.

The Personality Module was then installed. Testing of the personality module had very specific problems: the basic functionality tests, such as register tests and I/O functionality were simple. It turned out, however, that the serial data and command I/O was very hard to test at full speed, since the only available pattern generator was rated at only 40 MHz. It turned out that testing up to 50 MHz with simple test vectors was possible, but full-speed testing was only possible using two ROMs connected in a loop-back configuration.

The last step was the hookup of the actual readout chip and the debugging of the test board. Only then could we debug the full software functionality, as well as some previously undetected problems with the personality module.

6. The Measurements

6.1. Tuning of the Test Bench Setup

To get the entire test bench setup running the command/clock phasing had to be carefully adjusted. This was achieved by inserting a 6 ns cable delay in the DAQ board clock out, which moved the falling edge of clock just ahead of the change of the command line. The clock to command phasing seems to be the only truly critical setting on the board. At high speeds the I/O signal current levels have to be adjusted more carefully, but at speeds of 40 MHz the board was very robust toward changes in operating voltage, on board signal levels, and even clock symmetry. Higher speeds were not examined in great detail, as earlier tests and simulations had shown design errors preventing correct chip function at those speeds.

The AToM chip does not have a clock output which indicates when the data on the chip output is valid. Therefore a return clock line has to be supplied to the personality module. The clock frequency for the data output is the same as for the command input (the main chip clock frequency), so the same clock frequency as for the personality module clock out can be used. The phasing has to be matched with the phasing of the data coming from the readout chip at the personality module input. This value depends on the total cable length in the command and data pathway. Current cable length is approximately 6 ns (3 ns in the command path, 3 ns in the data path). It also depends on the delay on the test board and the chip itself. The time jitter in the data phasing, presumably caused by the readout chip, was less than 1 ns, which is more precise than our determination of the optimum clock/data phasing. In the pulse generator setup a 11 ns delay was used between system clock in on the DAQ board and data clock in on the personality module. This setting proved to be non-critical, therefore a readily available delay cable of 10 ns was used once the clock board was installed.

6.2. Digital Performance

The chip performs fully at 40 MHz, very few errors are detected in either the register test or the data taking runs using the above mentioned settings. There are occasional errors which appeared only during data readout. These errors occur in the form of no returned data from the chip. This could be caused either by a missed read-event or a missed trigger. These errors occur at a rate of less than 1 in 100000, which made it difficult to diagnose. The two most likely parameters responsible for the problem is the command/clock phasing and the command and clock signal swing at the chip input.

Two design errors on the AToM chip were detected during the test of the chip functionality. The first error was obvious only when looking at the data stream with an oscilloscope. When event data was read back from the chip a short (~1 ns) spike appeared immediately after the last data bit. The cause for this glitch has been found using the SPICE simulation of the digital part of the AToM chip.

The other error that was found was more subtle. It was known that the duration of the calibration charge injection pulse was given by the skip rate register setting instead of the shaping time register setting. This causes a calibration pulse length of only 1 μs if the skip rate is set to 'no skip' (see section 3.3.2.3). This in turn causes the truncation of the charge pulse at the preamplifier input after 1μs, so ToT values above 1.1 μs are not observed, regardless of injected charge. The unknown problem with this known bug was that the falling edge injects a large amount of noise in the next lower channel. Fortunately this occurs 1μs after the positive edge, so the trigger window can be used to mask off those unwanted pulses.

6.3. Calibration measurements

All measurements in this and all following chapters were done at the default chip settings, unless otherwise noted: 100 ns shaping time, no skip, 1μs trigger jitter window.

6.3.1. The Threshold Voltage

The threshold voltage is common for all channels on the chip. A probe connection to readily accessible probe points on the test board for this voltage simplifies calibration of this critical part of the chip. The measurement of the threshold DAC voltage can be done by simply connecting a high impedance voltmeter to both the threshold voltage and ground, and the threshold reference

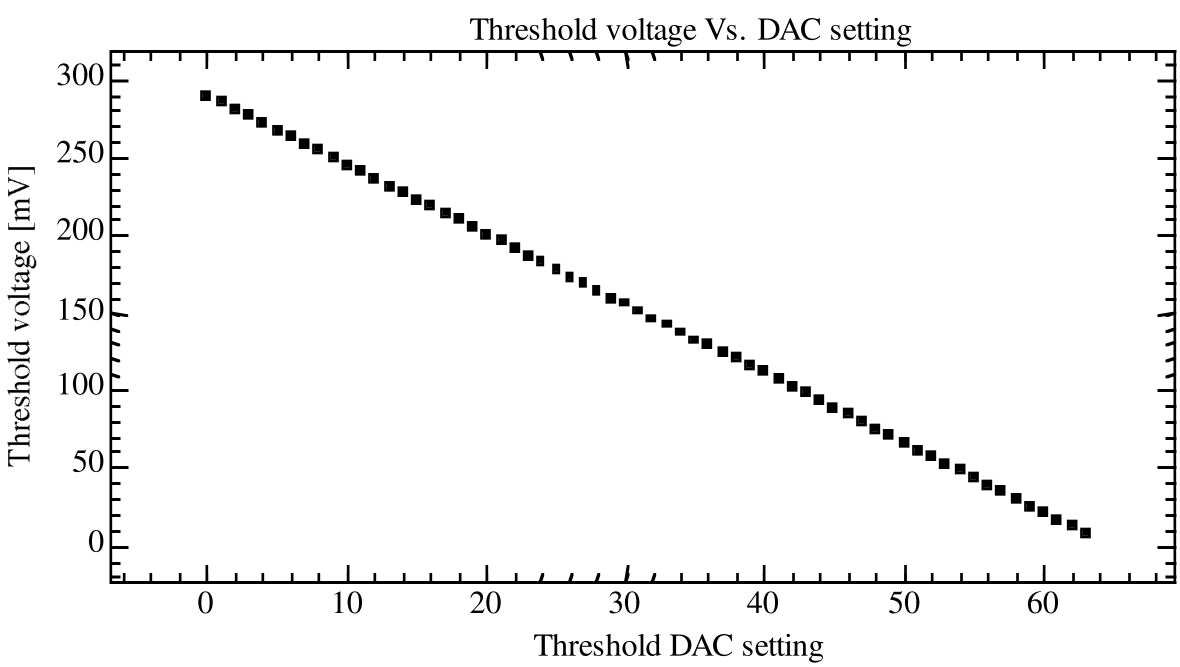


Figure 18 : Threshold Voltage vs. DAC Setting

and ground. The difference between the voltmeter readings is the actual threshold voltage. This measurement is better than a direct differential measurement, because the change in the reference voltage can be observed along with the change in the threshold voltage, revealing possible causes for nonlinearity due to an unstable reference voltage.

The reference, however, turned out to be stable to 0.2 mV at 303.8 mV. This variation is not significant, because it is less than the instrumentation error of 1‰.

A linear regression fit to the data shown in Figure 18 gives us the following two parameters:

The offset is: **290.28±0.05** mV

The slope is: **-4.48±0.001** mV/step

The Errors are calculated as the standard deviation of the linear regression. The differential measurement cancels any systematic error in the voltmeter.

The result lies within 2 sigma of the expected results, as given in [17]. Their measurement on 30 different chips led to an average threshold offset of 279.16±9.63 mV, and an threshold slope of 4.29±0.12 mV/step. This shows that the threshold DAC on the tested chip has relatively high offset and slope values, but the values are within the range observed in a random sample of 30 other chips.

6.3.2. The Calibration Charge injection Voltage

The voltage that is used to inject the calibration charge is not as readily accessible. It has to be probed with a microprobe on the chip itself. The probe pads for the calibration DAC voltage and its reference are very close together, making it impossible to probe both pads simultaneously. The calibration DAC voltages were measured at every setting in one run, measuring the reference directly before and

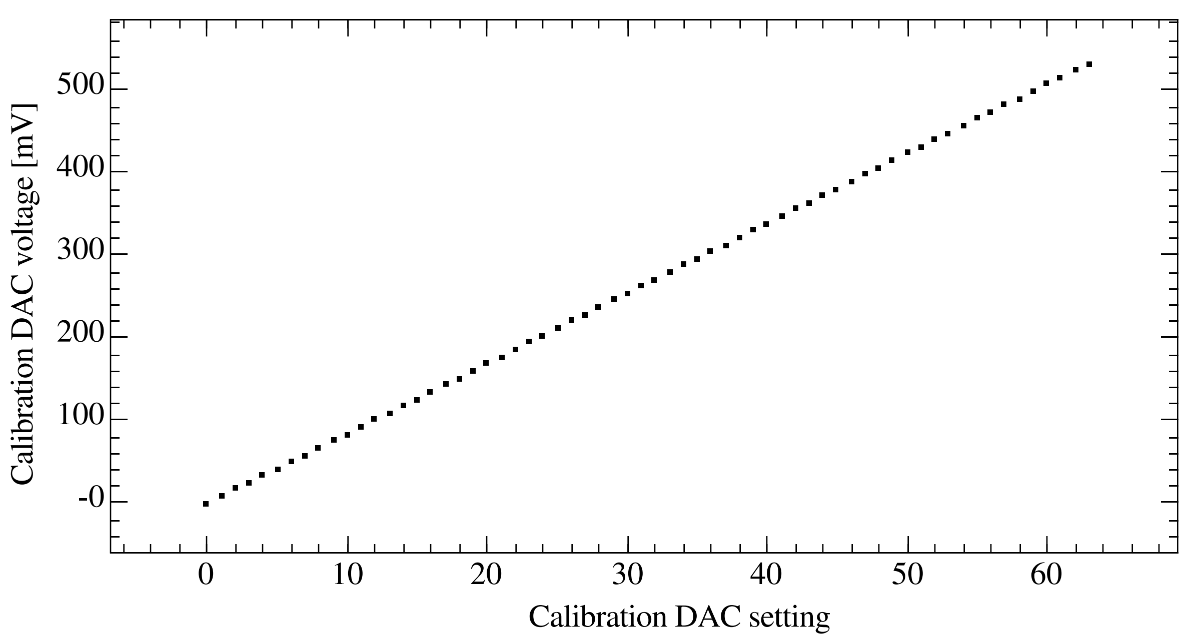


Figure 19 : Calibration Voltage vs. Calibration DAC Setting

after the series of calibration DAC measurements. The difference of the measured values was only 0.3 mV at 2.4717 V. This variation is not significant, because it is less than the instrumentation error of 1‰.

The reference voltage was read at every fourth calibration DAC setting, to see if there was any variation coupled to the calibration DAC voltage. But again there was no measurable variation. The measurement has an accuracy of 3 mV due to the 1‰ instrumentation error, therefore it can be concluded that the variation is lower than 3 mV. The extremely low fluctuation in the measurements suggest that the reference voltage is stable to a fraction of this upper limit.

The offset and slope value are calculated using a linear regression fit to the data shown in Figure 19. The offset so calculated is **1.2±0.2** mV. This does not correspond to the offset of 517.84±33.34 mV as given in [17]. The polarity select pin on the board in [17] was pulled high, which makes the comparison invalid.

The slope is **8.49±0.01** mV/step. This is less than 1 σ above the average slope measured on 30 chips [17]. The average value for the slope among those 30 chips is 8.217±0.17 mV/step, where ± denotes the 1 σ width of the distribution of measured values.

6.3.3. Summary of the DAC Measurement Results

Both the threshold voltage DAC and the calibration charge injection DAC show a very good linearity. None of the measured values was more than 1 σ away from the linear regression value. This means the measurement was not sensitive enough to detect any nonlinearity.

6.3.4. External Charge Injection Measurement Procedure

External charge can only be injected in one of the ten bonded input channels at a time. The other 118 channels can only be tested with internal charge injection, due to the limitations of the test board.

To perform a charge injection measurement the pulse generator output supplying the cal_strobe pulse to the DAQ board is disconnected from the DAQ board. The pulse length is increased to 5 μs. It is then connected to an attenuator to reduce the output voltage to less than 20 mV. The attenuator output is connected to one of the ten external charge injection inputs.

At each charge injection voltage a threshold scan is performed, injecting charge 200 times into the chip input. The ratio of hits that are recorded by the chip to the number of tries (200) defines the hit ratio. An error function can be fitted to the hit ratio (see 4.1). The fitting program returns a value for the turn-on point and the width of the error function, as well as the respective 1 σ errors.

The data from the SVT readout chip can not be fitted to regular error functions. At very low thresholds the comparator will continuously fire if the inevitable threshold offset brings the actual threshold below zero. The chip does only recognize a zero to one transition as hit, a constant one will be reported as no hit. This means that at very low thresholds there will be inefficiencies, some channels might only report a hit if there is a negative noise contribution. The fitting algorithm has been modified to disregard this drop in firing probability at low thresholds.

6.3.4.1. Measurements

The following measurements were performed:

- The pulse generator was set to output voltages of 0.8V to 1.4 V in steps of 0.2V. At the output of the attenuator the voltage was measured using an oscilloscope. This limits the total error in the injection voltage, caused by attenuator and pulse generator to 2%. For each voltage level and each of the 10 channels a threshold scan is performed. In order to reduce the statistical error the charge injection circuit is fired 200 times at each threshold value. An error function is fitted to each of the threshold scan curves to extract the turn-on point.
- The above measurements are repeated for channel 112 (the middle channel of the other group) with voltages of 0.1V to 1.4V in steps of 0.1V.

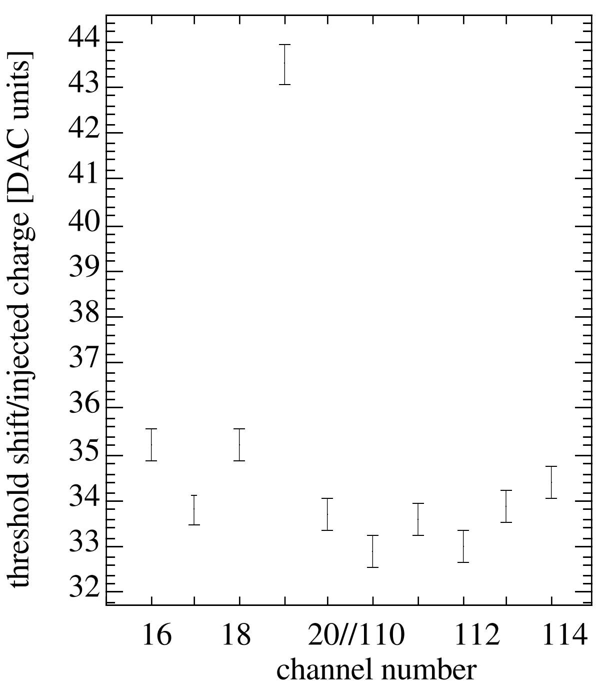


Figure 20 : Channel to Channel Variation in the Input Sensitivity of the Bonded Input Channels

For each channel the turn-on points at the 4 different charge injection voltages were used to calculate a linear regression. The slope of the linear regression is plotted in Figure 20. The slope is equivalent to the input sensitivity of each channel, because it marks the dependence of the shift in the threshold turn-on point due to change in injected charge. Where the charge can be calculated as the injection voltage times the injection capacitance.

The channel to channel variation at the same injected charge for channels with the same nominal capacitance is much less than the 10% given by the capacitor tolerance. This can be expected considering that all of the capacitors with the same nominal value come from the same batch. The variation within one batch is usually much smaller than the labeled tolerance. This does not help us to increase the accuracy of the chip calibration, because the average value of the entire batch of capacitors is not the nominal value.

Channel 19 has a roughly 25% higher sensitivity compared to channels 16,17,18 and 20. This is caused by its 1.2 pF injection capacitor. At the same voltage rise this capacitor injects nominally 20% more charge. The additional 5% are well within the range expected for capacitor variation.

Channels 16 to 20 have a relatively higher sensitivity than channels 110 to 114, when the lower capacitance is taken into account. This is caused by the increase in input sensitivity due to the added strip to strip capacitance. This effect was expected and has been seen in analog front end simulations [18].

To calibrate the internal charge injection the sensitivity of one channel has to be measured especially precisely, to compare the slope at known injected charge to the slope at unknown (internal) injected charge. Channel 112 was chosen for this purpose.

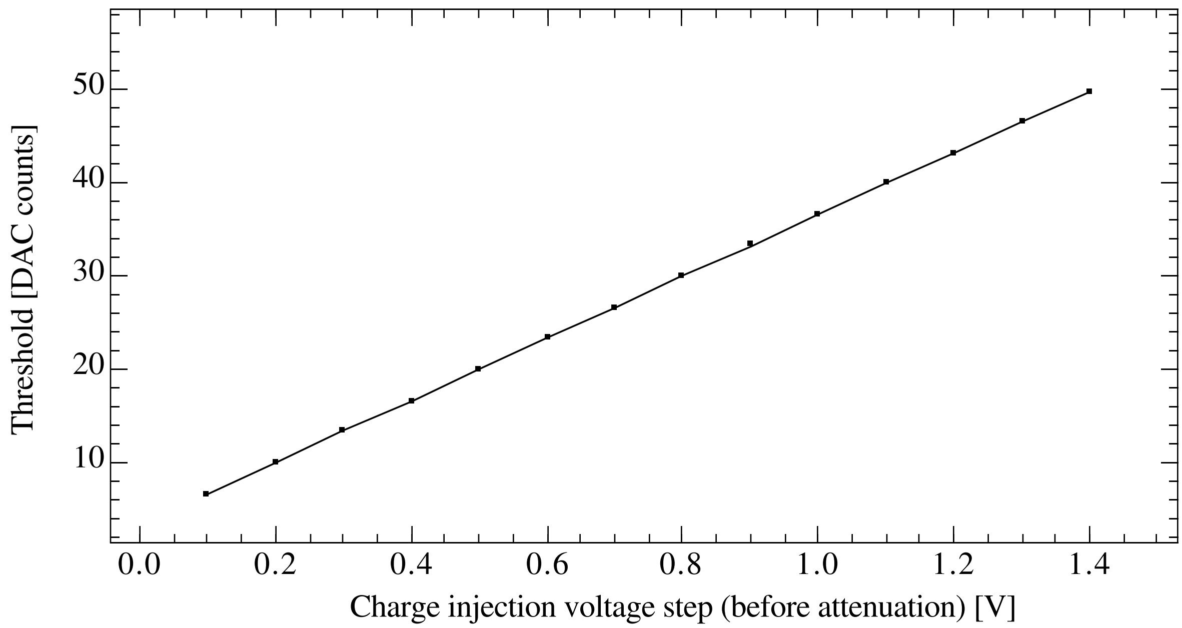


Figure 21 : Injection Voltage vs. Threshold Turn-on Point for Channel 112

The linear regression on the data plotted in Figure 21 gives the dependency of the turn-on voltage to the injection voltage:

a) Turn-on [DAC counts] = 3.32±0.05 counts +33.28±0.6 counts/V * U_i [V]

The errors given here are the standard deviation errors from the linear regression. The error in the slope is consistent with an expected error of 2% in the injection voltage.

The injected charge in the preamplifier can be calculated from the Voltage step, the attenuation and the injection capacitance. Thus we have:

b) Q_i [fC] =U_i [mV] *(0.002±2e-5) *C_i [pF]

Where Q_i is the injected charge, U_i the injection voltage (before attenuation) and C_i the injection capacitor. U_i is known to within 2%, so the 10% error in C_i is dominant.

Combining a) and b) we get:

c) Threshold [DAC counts] = 3.32±0.05 counts +13.9±1.4 counts/fC *Q_i[fC]

The threshold DAC transfer function was:

d) U_t [mV] = 290.3±0.5mV -4.482±0.005mV/count * T [counts]

With the reversed Threshold from the fitting program this corresponds to:

e) U_t [mV] = 7.93±0.5mV +4.482±0.005mV/count * T [counts]

Combining c) and e) we get:

f) U_t [mV] = 22.81±0.55mV +62±7mV/fC * Q_i[fC]

This is the gain of the preamplifier/shaper at the fastest shaping time (100 ns). Gain for the longer shaping times is about 50% higher [19]. The gain was measured only at the fastest shaping time, because this shaping time will be used during all further measurements. Measurements at slower peaking times are influenced by the bug in the calibration charge injection pulse duration (see 6.2).

6.4. Noise and Pedestal Variation

After calibrating the entire system the noise and pedestal variation can be measured. This is done in a single measurement by performing pedestal scans on all channels at variable injected charge.

6.4.1. Capacitance of the Charge Injection Capacitors

The crosstalk measurements show that the charge injection circuitry injects a considerably lower charge when all channels are turned on (see 6.7.2). To measure the calibration charge injection capacitance charge was injected only into channel 112. At each threshold level 200 charge injections were performed to obtain threshold scans with low statistical errors.

Table 4

DAC setting	Threshold Turn-on Point [DAC units]
0	6.4±0.8
1	14.0±0.6
2	21.6±0.7
3	29.5±0.7
4	37.1±0.7
5	44.9±0.9
6	53.0±0.8
7	61.1±0.8

Again channel 112 is measured, to allow a direct calibration of this channel with the data obtained above. Figure 22 shows the threshold turn-on point dependence on the calibration charge injection. The error bars represent the errors given by the fit program. The plot shows that the linear regression falls well within those errors, which proves that the internal charge injection has a very good linearity.

The linear regression gives:

a) Threshold [cts] = 6.15±0.15 counts + 7.80±0.04 counts/count *Cal [cts]

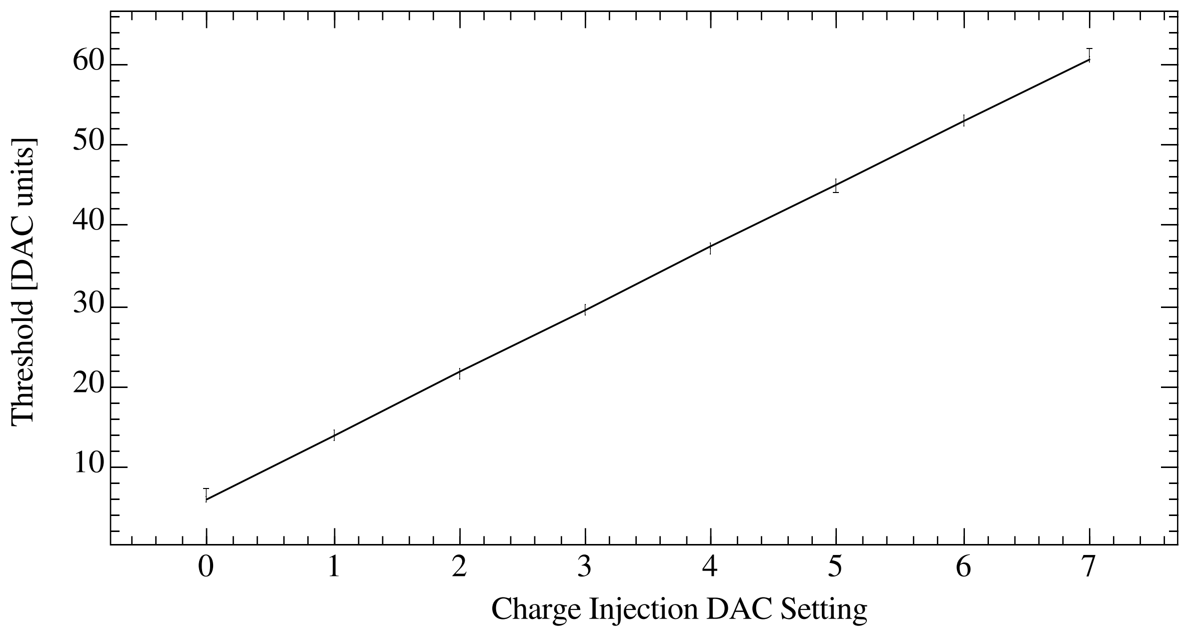


Figure 22 : Internal injected Charge vs. Threshold Turn-on Point

And the calibration voltage (see 6.3.2):

b) U_c = 1.23±0.08 mV + 8.465±0.002 mV/count * Cal [counts]

Combining 1) and 2) we get:

c) Threshold [counts] = 5.02 counts + 0.92 counts/mV *Uc [mV]

In 5.2.1.1 c) the slope is **13.9±1.4 counts/fC**.

Combining the slopes from 5.2.1.1 c) and c) we get:

d) Cc =0.071±0.008 pF

6.4.2. Simultaneous Measurements on all Channels

To obtain data for all channels the following set of measurements was performed on all channels: all channels are turned on, the masks on the charge injectors are set to inject charge in all channels. At each calibration DAC setting the calibration circuit is fired 200 times to give good statistics. This is then repeated for all other thresholds. For each channel the hit ratio versus threshold and injected charge can be plotted (Figure 23 shows channel 64). The plot clearly shows a narrow turn-on curve, indicating a low noise. The turn-on point shifts towards higher

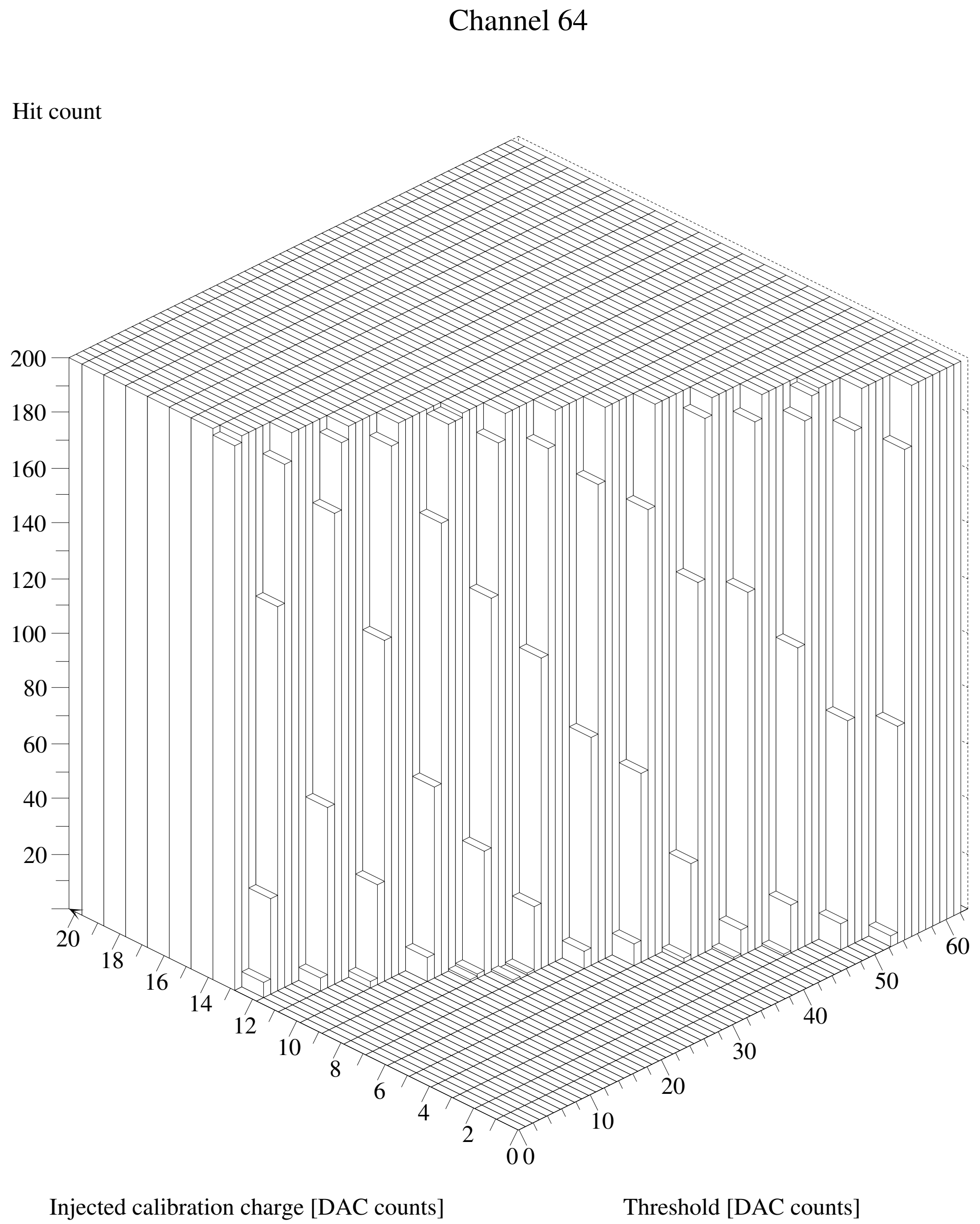


Figure 23 : This 2-D Histogram represents the hit count vs. threshold and injected charge for channel 64. The histogram was taken using the Histoscope package.

Another interesting plot is the hit ratio versus channel and threshold at a fixed injected charge. Figure 25 shows this plot at an injected charge of 5 calibration DAC units. The 5 channels with external capacitors are clearly visible due to their higher noise (wider turn-on curve). The channel to channel variation of the turn-on point at identical injected charge can also be seen on this plot. The third possible plot is the hit ratio versus channel and injected charge at fixed threshold. This is very similar to Figure 24, and does not give any additional information. Therefore it was not included here.

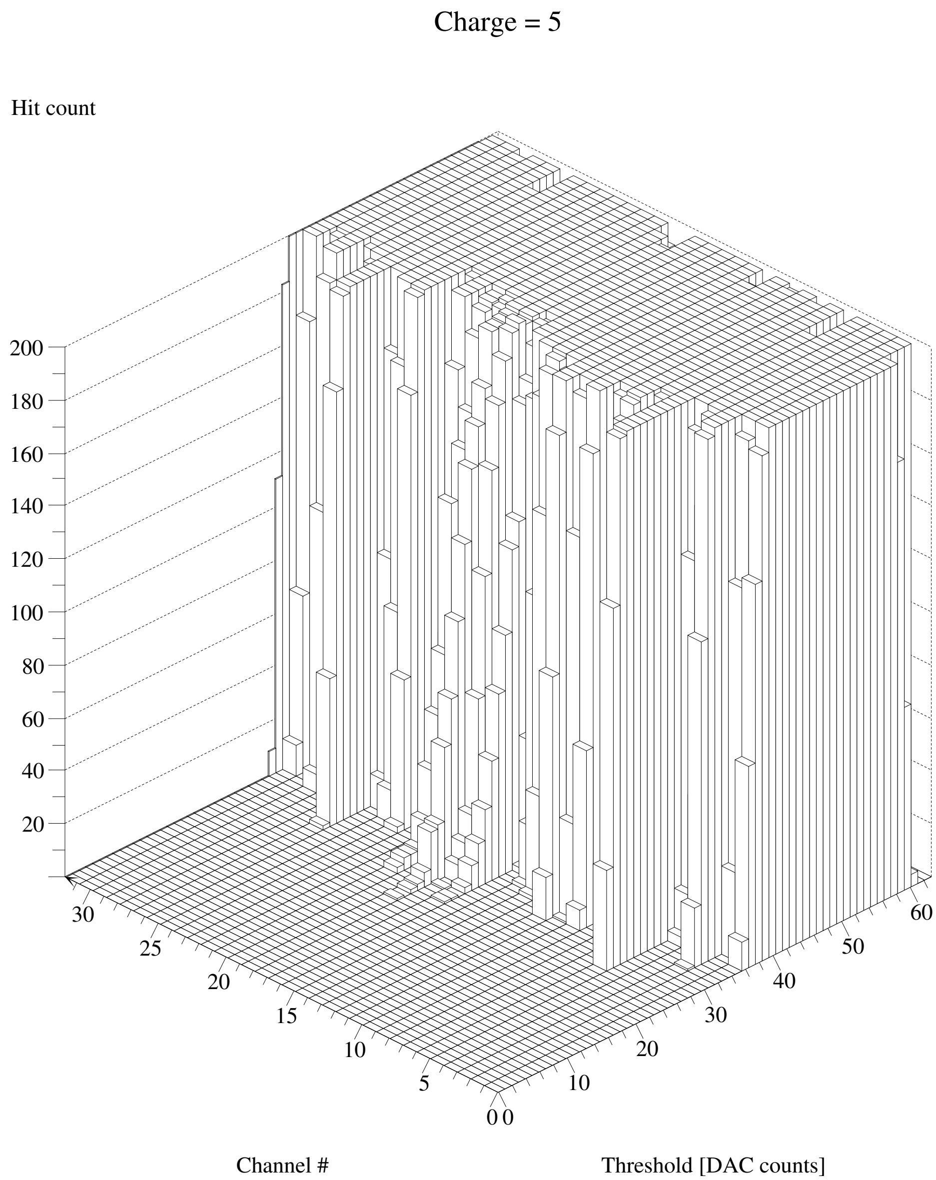


Figure 24 : This 2-D histogram shows both pedestal and noise variation at an injected charge of 5 DAC units. The histogram shows only 32 channels, because the full 128 channels makes the picture too confusing.

6.4.3. Results of the Noise and Pedestal Measurements

The data set from the measurements described above is much too big to include here. The two plots above should give some understanding of the complexity of the data, while allowing to estimate both noise and pedestal variation.

Further data processing is necessary to gain a full understanding of the measurement results.

The error function fitting program is used again to determine the turn-on point and the width of the error function along with the respective errors. Three results can be extracted from the fitted data:

- The channel to channel pedestal variation.
- The channel to channel variation of the sensitivity to injected charge (gain).
- The noise for each channel and its distribution over the chip.

For each channel a linear regression on the turn-on point versus injected charge was calculated to determine the intercept and slope. The drop in injected charge due to the simultaneous firing of all channels (see 6.7.2) changes the slope, but does not effect the intercept.

The intercept of the linear regression defines the pedestal. This is more precise than measuring the turn-on point at zero injected charge. The linear regression also allows negative intercepts, while a pure measurement without injected charge would give no value for channels with a negative threshold offset.

The slope of the linear regression is a measure for the gain and calibration capacitance. The absolute value is not of interest here, but the channel to channel variation has to be determined.

6.4.3.1. The Pedestal Variation

Figure 25 shows pedestals for all channels plotted as a histogram. The pedestal is converted into units of mV at the comparator input according to:

a) Pedestal [mV] = Pedestal [threshold DAC counts] * (4.48±0.001) mV/count

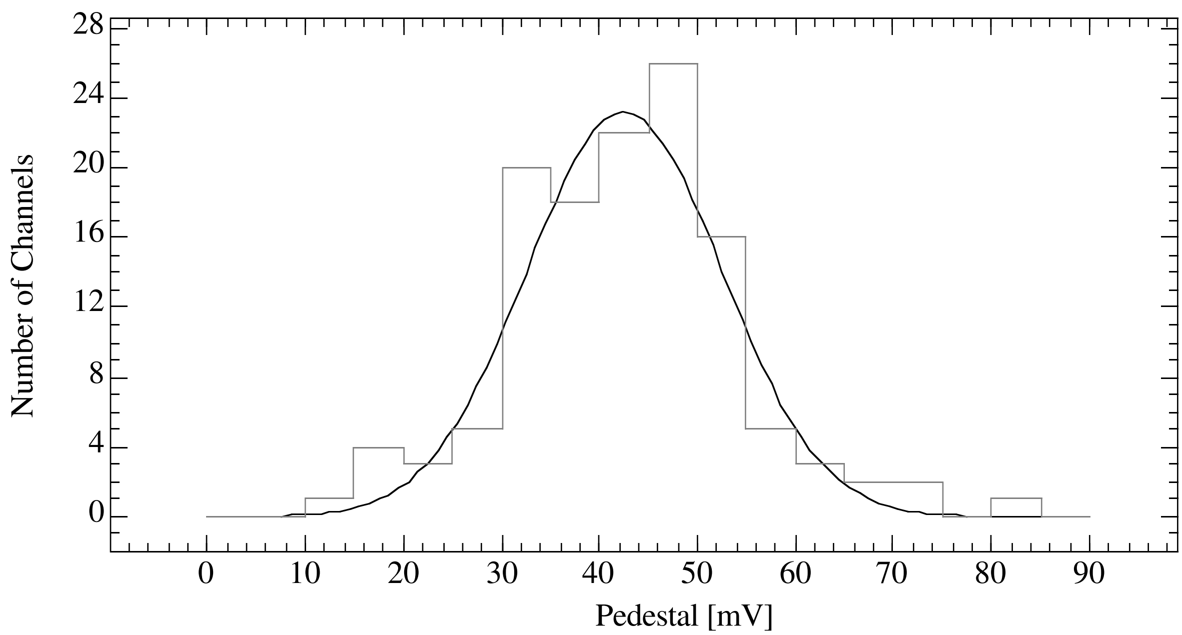


Figure 25 : Histogram of the pedestal variation with a gaussian fit

With just 128 channels as the number of entries the histogram has fairly low statistics, still a gaussian can be fitted to the histogram. The gaussian fit matches the histogram with a χ^2 of 11.9 with 14 non-zero histogram entries used for the fit and 3 degrees of freedom for a gaussian this results in an overall goodness of the fit of 11.9/(14-3) = 1.082.

The histogram shows a RMS channel to channel pedestal variation of **11.6±0.1 mV**. The maximum variation is ±**64 mV** peak to peak. Figure 26 shows the distribution of the pedestal values across the chip. The plot shows that there is no systematic drop or rise of pedestal values across the chip.

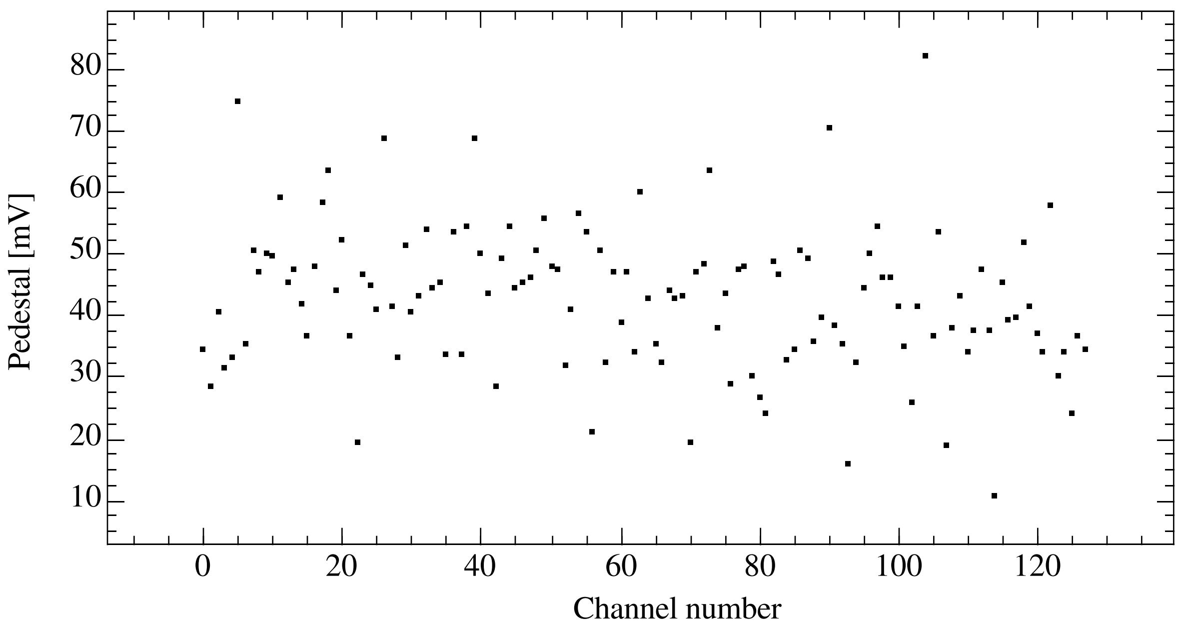


Figure 26 : Distribution of pedestals across the chip

The mean pedestal value is **43±2 mV**. This is caused by the internal charge injection mechanism. There is not only the calibration DAC offset, but also some additional contribution. There is a small amount of charge injected just by activating the internal charge injection switch. Once the internal injection is disabled or masked off the average pedestal is roughly at zero. This can be verified by a threshold scan while charge injection is disabled.

Figure 27 shows the turn-on points of all channels obtained with a threshold scan without any charge injection. No error function can be fitted to channels which have a pedestal lower than 1 threshold DAC unit, because too large a part of the function is cut off. Therefore pedestals below one are marked as zeros on the plot. The plot clearly illustrates that the true pedestals are closer to zero than indicated by the measurements using charge injection. This plot can not be directly used to calculate the true average pedestal position, because no values for the channels with below zero pedestals are available. A possibility to estimate the average pedestal is to calculate the average shift for all above-zero channels, then shift the mean value measured above by this correction factor. This assumes that at a calibration DAC

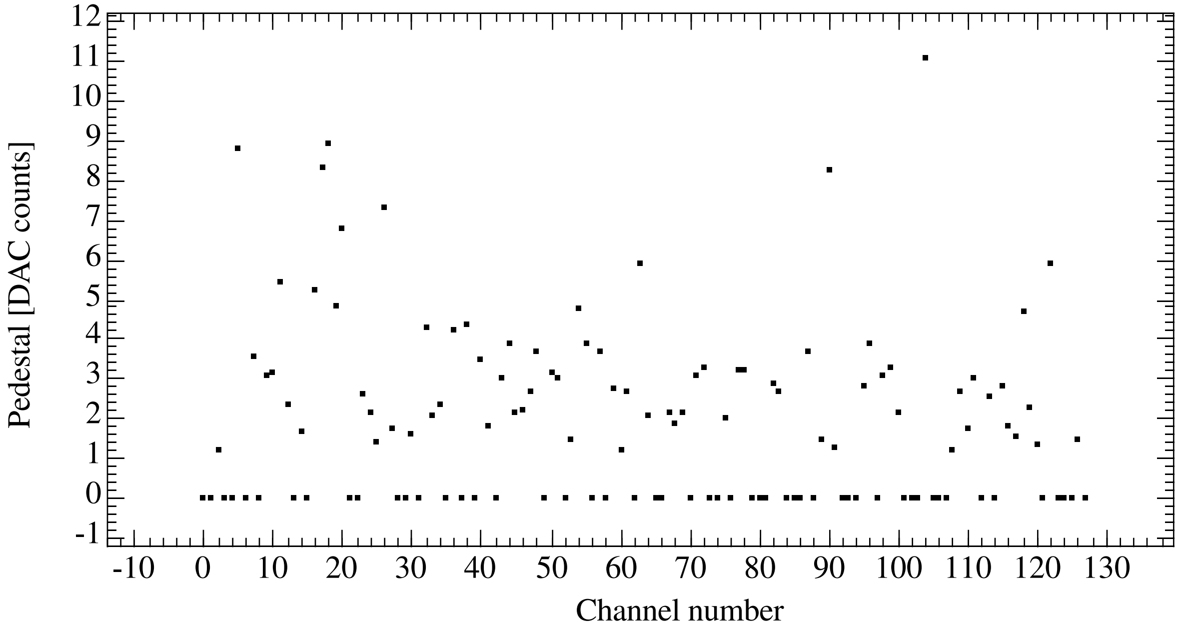


Figure 27 : A threshold scan without any charge injection shows the true position of the positive pedestals.

setting of zero the same ammount of charge is injected into all channels. Such a correction is possible only if this offset charge is almost constant. The average shift is **40.8±0.2 mV**, therefore the corrected average threshold is **2±2 mV**.

6.4.3.2. The Variation of the Sensitivity

The second result that was extracted from the measurements above is the channel to channel sensitivity variation. The RMS channel to channel variation of the slope of the linear regression, which is equivalent to the input sensitivity , is only 4%. This variation can be caused by either gain or charge injection capacitor variance. The gain variance is expected to be small (<1%), while the variation of 4% is within the range expected for the calibration capacitor variance (5%) alone.

Figure 28 shows a channel to channel comparison of the slope of the linear regression. The channels in the middle of the chip have higher sensitivity values than the side channels. The statistical error at just 128 channels is too large, so the tendency is not conclusive. Measurements on other chips will be needed to verify if the sensitivity is homogenous.

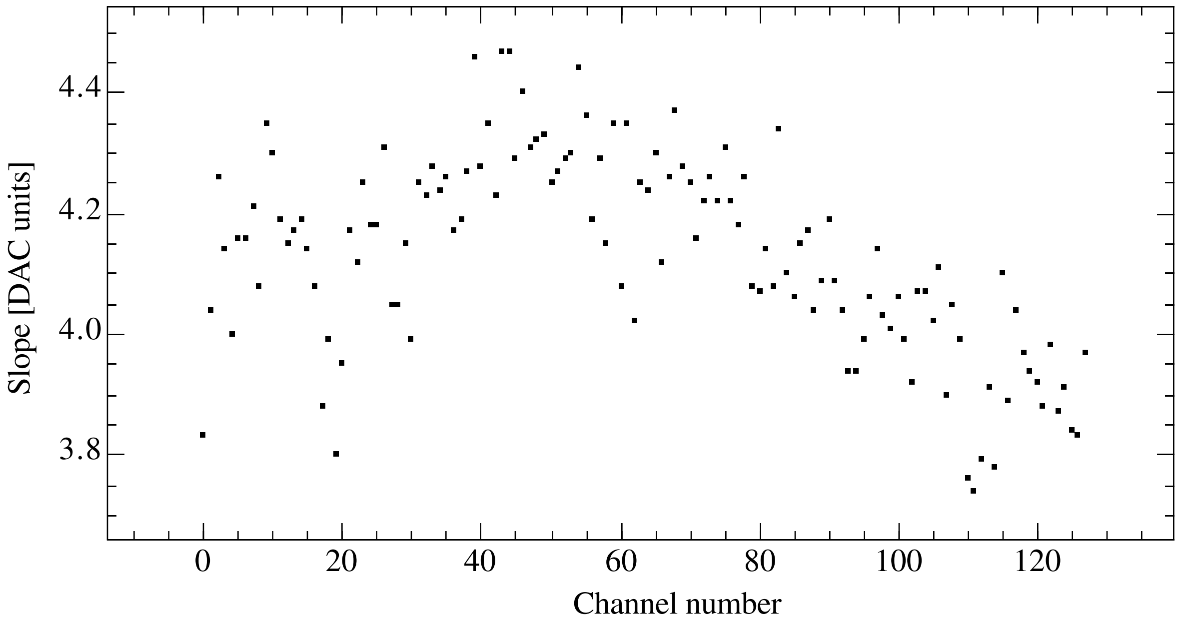


Figure 28 : The Distribution of the Sensitivity of the Channels to the Calibration Charge Injection Circuitry across the chip.

6.4.3.3. System Noise

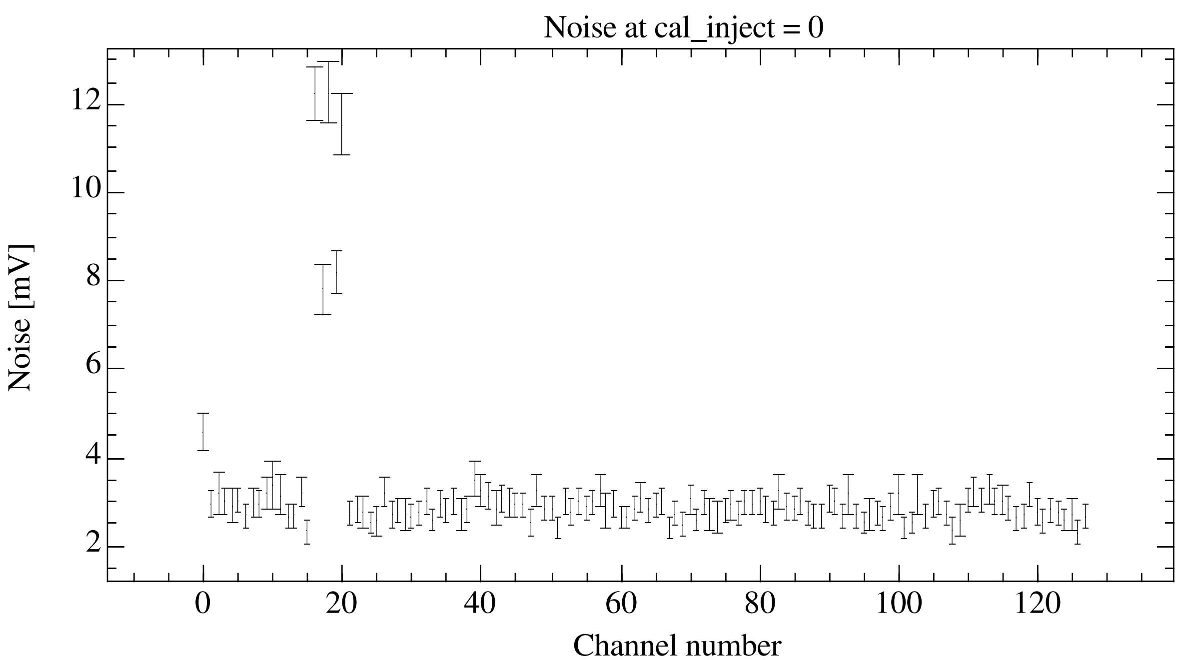


Figure 29 : Noise vs. Channel Number. Error bars correspond to the errors given by the fit program.

The last result that can be extracted from the set of measurements above is the system noise. As mentioned above the noise can be extracted as the width of the error function (see 4.1). The noise on channels without external input capacitors is extremely low, while the noise on the channels with channel to channel capacitors is entirely dominated by the noise component of the external input capacitor (see 2.2.3).

Figure 29 shows the distribution of channel noise across the chip. The measurement was taken using the lowest possible charge injection setting (calibration DAC = 0) to reduce the influence of crosstalk noise. The measurement can not be taken without charge injection, because the turn-on point of each channel has to be shifted above zero to successfully fit an error function. Channels 16 to 20 have the channel to channel capacitors installed, but due to the layout of the test board the outermost of those five channels on the test board (the ones with just one neighbor) are channels 17 and 19. The sequence of channels on the test board is: 17, 16, 18, 20, 19. The graph clearly shows the three channels with capacitors on both sides having the highest channel noise, while the noise of the two outer channels with capacitors to just one neighboring channel is considerably lower.

The noise is expected to be linearly dependent on the input capacitance, therefore this measurement gives a fair measurement for the intercept and a rough estimate for the slope.

As mentioned above, the shaping time has been set to 100 ns. Longer shaping times reduce the preamplifier/shaper bandwidth, resulting in an overall lower noise. Therefore the shaping time of 100 ns is well suited to find the maximum noise contribution. At calibration charge injection DAC setting of zero (this is not zero injected charge) the average noise on nonbonded channels is **2.86±0.03 mV** or **288±3 electrons equivalent noise charge (ENC)**, while the average noise on the three channels with 30 pF total channel to channel capacitance is 12±1 mV or 1200±100 electrons ENC at the preamplifier input . On the two channels with 15 pF to just one neighbor the average noise is 8.0±1 mV or 800±100 electrons ENC.

That corresponds to a slope of 31±3 electrons/pF. The chip noise should be close to the value of 288 electrons ENC, but the error is very hard to determine, because the test board and power supply influence is unknown. The estimated error is 10%.

The error for the bonded channels is naturally much higher, not only due to the unknown capacitance, but also due to possible external noise coupling. This measurement therefore does not allow quantitative conclusions about the noise values for the final system, but it is expected that those values will be similar or better than the values achieved on the test setup.

The width of the error function at higher injected charges is slightly wider. This does not indicate, however, that the noise is increasing with higher injected charge. The broadening of the error function is caused by the variation in injected charge. The charge varies by less than 1% or 1/10th of the least significant bit, but that is still enough to make a noticable impact on the width of the error function. The increase in the width of the error function between calibration DAC setting = 0 and calibration DAC setting = 10 is 13%: It rises from 288±3 electrons ENC to 327±3 electrons ENC.

Exact measurements of the noise to input capacitance relationship were done in Pavia [12]. The results of both measurements show good agreement in the slope: They measured 29 electrons/pF, which is within the measurement error range of the 31 electrons/pF measured in our setup. The chip noise of the unbonded channels, however, shows substantial disagreement. We measured only 288±3 electrons ENC, while the result obtained in [12] was 341.5 electrons ENC. This is due to the fundamentally different method used: they used a scan of injected charge at fixed threshold, while we used a threshold scan at fixed injected charge. This makes a direct comparison of the results invalid.

6.5. The Analog to Digital Conversion

The time over threshold (ToT) A/D conversion is non linear by design. It is intended to be sensitive for small signals, and less sensitive at large signals, therefore performing a compression of the dynamic range. The transfer function should be roughly logarithmic, but the exact shape is not critical as long as the charge to ToT value relationship is well known. To measure this relationship the ToT value was read 1000 times at each injected charge, and the mean ToT value was plotted against the injected charge for four different threshold settings: 0,20,40 and 60 DAC counts. Channel 63 was used for this measurement.

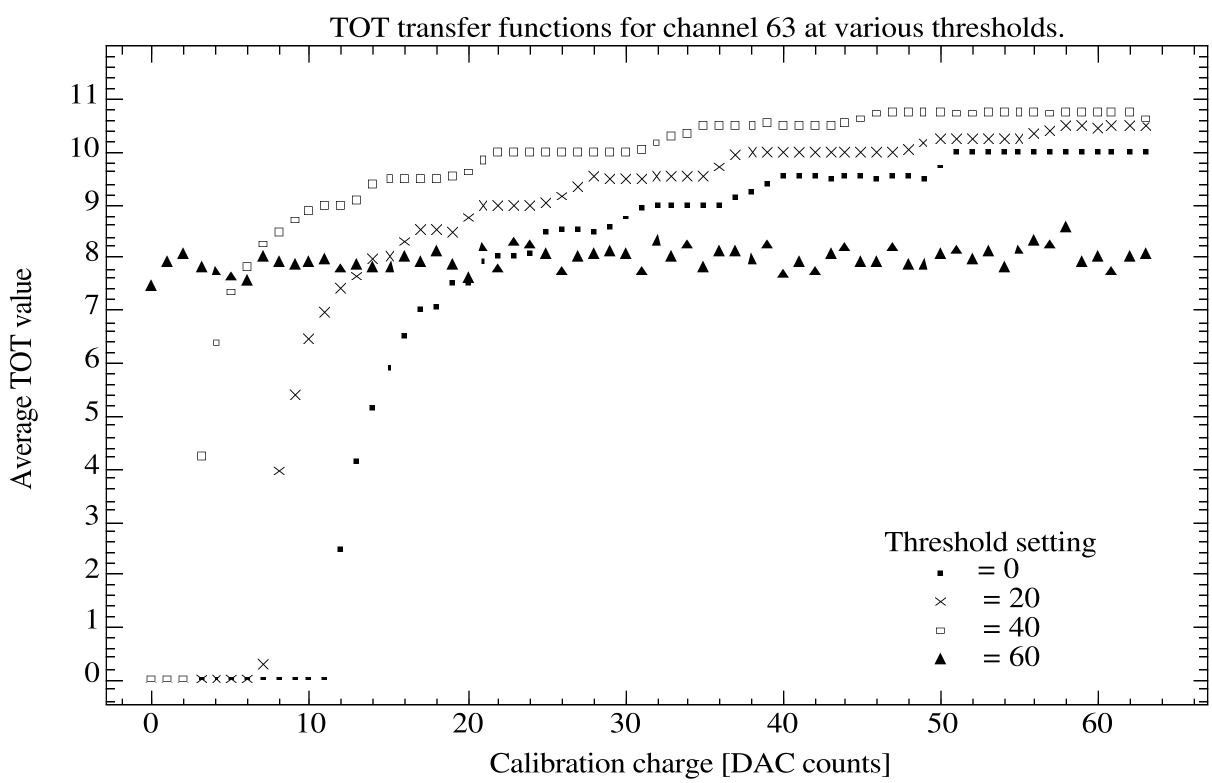


Figure 30: The ToT transfer Function at different threshold settings

Figure 30 shows that the transfer functions behaves as expected: with increasing threshold (decreasing threshold DAC setting) the curves are shifted towards higher charges to achieve the same ToT value. Otherwise the form of the curves is very similar. The graph at threshold DAC = 60 is an exception: The channel is clearly in the noise limit, so even without injecting a charge the channel receives hits. The ToT value does not depend on the injected charge, because at high injected charges, where the ToT value of the injected charge would be higher than the ToT value that is recorded, the chip records the ToT value of the first hit, which is usually a noise hit, and the 'true' hit is disregarded.

The next somewhat peculiar thing is the stepsize of the functions. The functions are not very smooth, instead there are pronounced steps, with the step size being about 1/2 of a ToT count in most cases and 1/4 of a ToT count in three cases: The last step for threshold = 40 and the last two steps for threshold = 20. This can be explained by the fact, that the injections are done by the internal charge injection circuitry, therefore they are synchronized to the chip clock. The chip clock runs at 4 times the ToT sampling speed. Thus at any injected charge the random phasing of the cal_inject pulse to the sampling clock determines if the ToT records a one count higher or lower value. For a certain pulse height there are only four discrete phasing possibilities, some of them causing the higher, some the lower count. This ratio causes the 1/4 step sizes. It has not become clear why the 1/2 step size is so much more pronounced than the 1/4 step. If the charge would be injected at a truly random time there should be no noticeable steps in the graph.

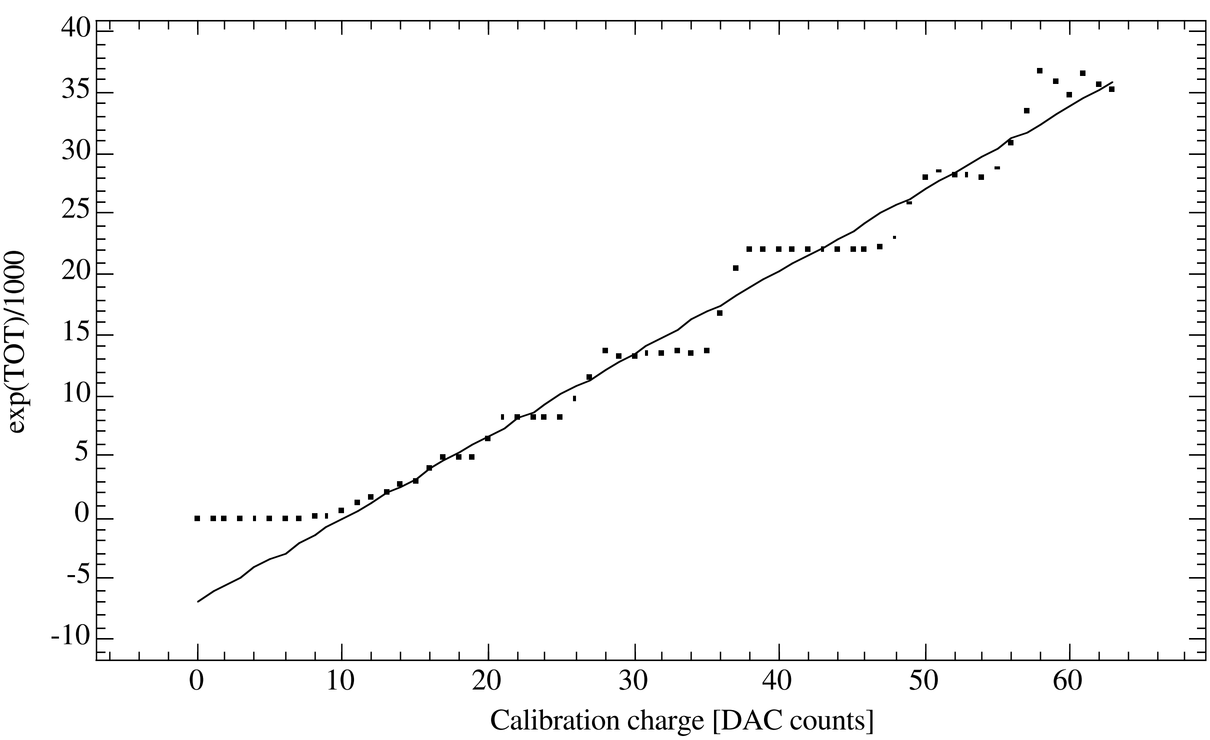


Figure 31 : The ToT transfer function compared to a fitted logarithm for threshold = 20

The ToT values have an upper limit of 11 due to the bug in the length of the calibration charge injection pulse. It is not entirely clear at which point this cut-of starts to influence the transfer function, because the steps make it impossible to extrapolate the function into the higher charge region. Figure 30 shows that the cut-of has little or no effect up to injected charges of 57 DAC units at a threshold setting of 20 DAC units.

Figure 31 shows a comparison of the transfer function to a logarithmic dependency. Here the steps are even more obvious than in the regular plot, so it is difficult to do a direct comparison. Still it becomes clear that the dynamic range compression is slightly less than logarithmic. The slope of the data set at the low end is slightly less than the slope of a linear regression fit. The comparison with the logarithmic behavior, however, is just to illustrate the shape of the transfer function better. No truly logarithmic behavior is expected. The dynamic range compression depends only on a variable slope, no specific gradient is necessary, as long as low amplitude hits are recorded precisely, and high amplitude hits are recorded with a comparable percentile error.

6.6. The Time Stamp Counter

The time stamp is the other value associated with each hit. While the ToT value is used to determine the exact particle position, the timestamp is used to determine if certain hits are part of the same event during the reconstruction.

A bug prevents the counter from counting properly: It counts 1,2, 3, 5, 5, 6, 7, 9,9,10,11,13,13,14,15,17,17,18,19,21,21,22. This has been well documented in [20]. A measurement of the timestamp response to the trigger delay is included here for the

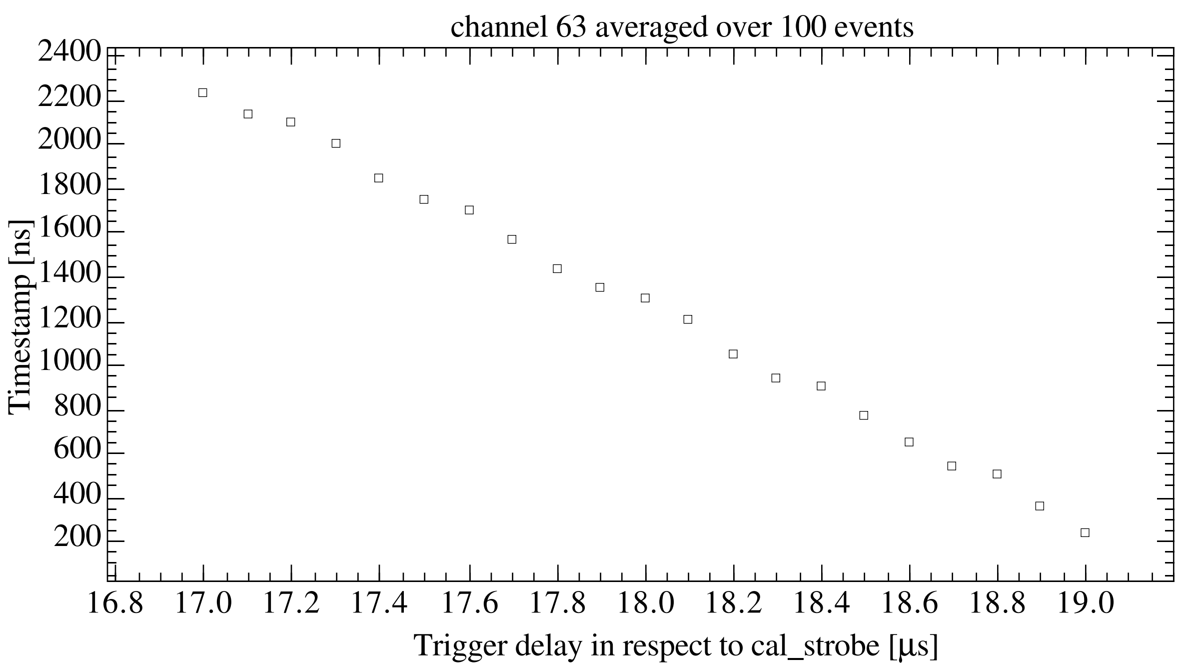


Figure 32 : This plot shows the nonlinear response of the timestamp counter due to an error in the graycode counter on the chip.

sake of completeness, even though this bug prevents any conclusive testing. Figure 32 shows the nonlinear response of the timestamp counter due the bug mentioned above. It makes it impossible to determine the true time of a hit from its timestamp value.

More interesting than the actual correspondence between the timestamp and time is the time walk due to variable signal heights. The time walk is dependent on the shaping time, because the shaping time influences the signal shape. The time stamp was measured 100 times at each charge injection setting and each shaping time for a moderate threshold (threshold setting = 24)

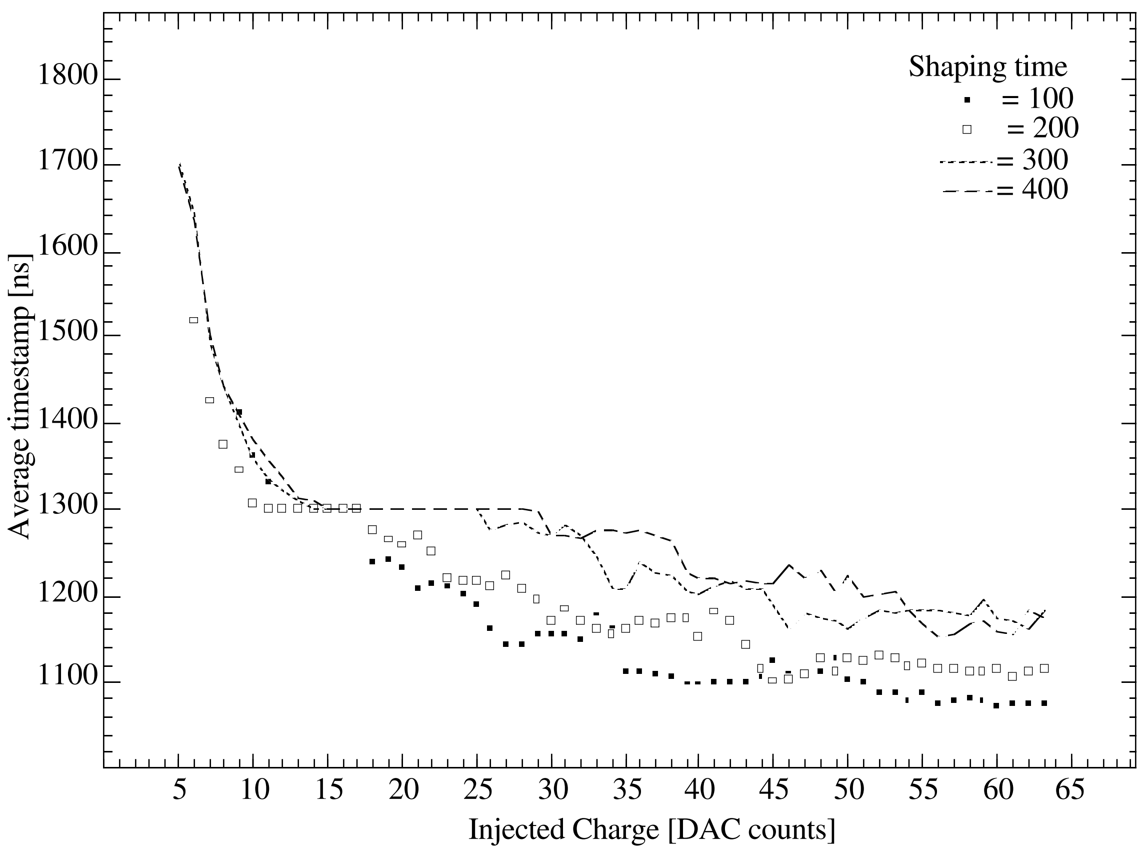


Figure 33 : Timewalk of the timestamp signal vs. signal height

Figure 33 shows the time walk of the timestamp value at different shaping times. The principal shape of the graphs is easily understood: At very small injected charges the timestamp is just the middle of the peak. There is a steep drop in the timestamp value as the injected charge rises, because the peak is very flat in that area, therefore a small change in peak height corresponds to a big change in peak width.

The strange details of the graphs are caused by the error in the timestamp counter. The plateau at timestamp = 13 is the most prominent feature. It is clearly caused by the doubling of the count of 13 in the timestamp counter. All 4 graphs eventually flatten out, due to the steep rising flank of the peak, but they do not converge to a certain limit, until the slew rate of the shaper is reached. This point can not be determined with the measurements above, because the bug in the timestamp counter prevents a predictable plot shape.

The total time walk is about 600 ns, 65% of the time walk occur in the range where the signal is only 3.0±0.5 fC above the threshold, the remaining 35% occur in a signal range of 3.0±0.5 fC to 33±6 fC above threshold.

6.7. Crosstalk Measurements

The high integration of the chip raises the question of adequate channel isolation. Two mechanisms couple the results of one channel to the other channels on the chip. Crosstalk, usually caused by capacitive coupling, causes signal sharing with the neighboring channels. The crosstalk influence on the neighboring channels is therefore of the same polarity as the signal. The other possible coupling is the amount of influence firing channels have on each other due to a drop in the supply voltage when multiple channels fire simultaneously. This effect has the opposite polarity of crosstalk influence: The readout influence causes lower readings in a certain channel if other channels fire simultaneously.

6.7.1. Crosstalk Measurement Procedure

To measure the crosstalk all channels are turned on. A charge is injected in a randomly selected channel. The innermost channel (63) was selected for this measurement. A threshold scan is performed on all other channels to determine if there is any charge injected into the neighboring channels. This procedure is repeated for every charge value.

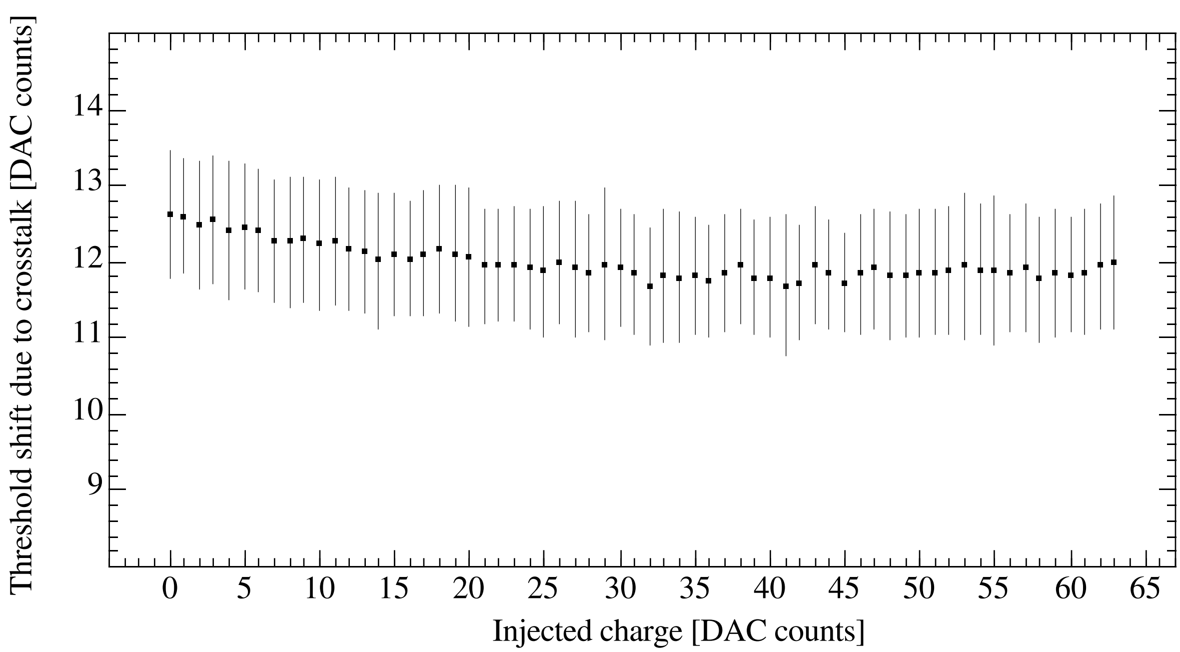


Figure 34 : Crosstalk observed on channel 64 when charge is injected into channel 63.

The crosstalk for any channel is derived by subtracting the threshold without charge being injected from the threshold when charge is injected into channel 63. Only channel 64 showed any significant crosstalk influence. Other channels show no visible crosstalk influence. If there is any influence in the other channels it is much smaller than the accuracy of the measurement. Not even a slight tendency of increased thresholds could be recorded.

Figure 34 shows a strong crosstalk influence on channel 64, that is essentially independent of injected charge. This extremely high crosstalk is caused by the charge injection switch itself, it is not visible when external charge injection is used. When external charge injection was used on channel 112 during the chip calibration a crosstalk of 1.5±4 mV was observed in channel 113. This result can be interpreted as no measurable crosstalk.

6.7.2. Readout Influence Measurement Procedure

The readout influence is measured by injecting charge first in the middle channel (channel 63), then channel 63 and its two neighbors, and so on until charge is injected in all channels. The change in threshold of the channel 63 is observed.

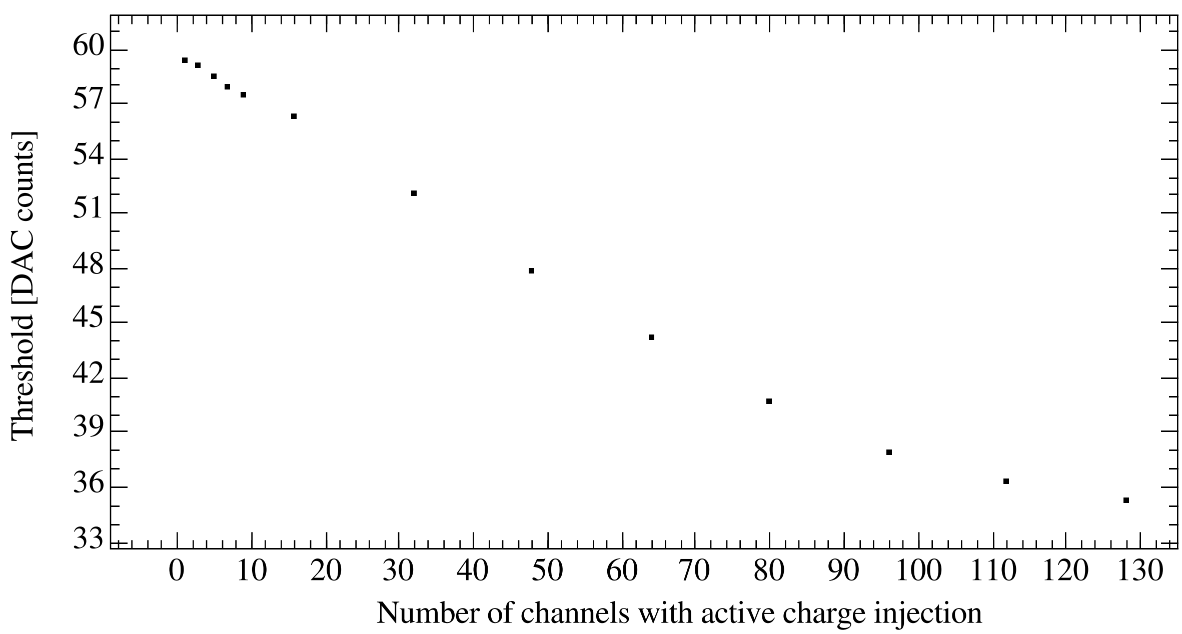


Figure 35 : This plot shows the drastic drop in measured injected charge when several channels are fired simultaneously. The measurement was made with a charge injection seting of 5 calibration DAC units.

Figure 35 shows the drop in measured injected charge when several channels are fired simultaneously. The measurement was made with a charge injection seting of 5 calibration DAC units. The change in threshold is fairly linear with the number of simultaneously firing channels. At high channel count it flattens out slightly. The difference between just a single channel and all channels firing is on the order of 40%. This is much higher than expected, but it is probably caused by a drop in the calibration charge injection voltage due to the sudden load, rather than a change in gain due to loading of the analog power plane. A drop in the calibration charge injection voltage does not compromise functionality of the chip, because the drop can be corrected by increasing the calibration DAC setting according to the number of channels that will be fired. With the external charge injection setup on the test bench a well controlled charge can not be injected into several channels simultaneously, so there is no way to verify if the calibration charge injection voltage is the cause. The assumption is based on the fact that the calibration charge injection line is loaded heavily by switching all injection capacitors simultaneously, so a drop is to be expected, while the influence on the preamplifier and readout electronics is much less direct. The linear drop makes it more likely that only one effect is causing the difference, if two effects superimposed the result is usually a graph with a dual slope, as each effect dominates in a certain region.

7. Conclusion

7.1. The Performance of the Digital Part

The digital section of the chip works very well, except for the known bugs. Even with those bugs the chip is fully functional for a test-bench at 40 MHz. Both DAC converters have a linearity that is so good, that no deviations could be measured. Their offset and slope is within design value range.

The command decoding section is somewhat sensitive to both signal swing and command/clock phasing. In a final setup with fixed delays this will be much less of a problem than on the test bench, so this sensitivity is acceptable.

The ToT analog to digital converter shows the desired dynamic range compression. The exact response has to be determined once a detector is bonded to the chip, using a laser diode to inject a well known energy into the detector. This will be done during the system test in early 1997.

The timestamp counter has a large amount of time walk, this is not a problem, however, because the time stamp is used only during reconstruction, where the time walk can be compensated using the ToT value.

Many problems on the prototype chip have been identified, and fixes have been implemented. The next version of this chip is expected to perform at the full operating speed of 59.5 MHz

7.2. The Analog Section

The noise on unbonded channels is negligible compared to all other noise factors. With just 288±8 electrons ENC it does not influence the total system noise noticeably. The preamplifier noise due to input capacitance agrees both with other measurements, as well as simulations [12] [22].

Assuming a 15 pF strip to strip detector and fanout capacitance (as simulated by the test board capacitors) the noise would be roughly 1200 electrons (see 6.4.3.3). This would roughly correspond to a layer 3 Z-strip [21]. This is in accordance with the simulation results in [22]. The measurement is just a rough approximation of the actual detector setup, while the simulation included a full model of the detector and amplifier. Therefore the experiment can not really be seen as a validation of the simulation.

This rough estimate would result in a signal to noise ratio of about 18 to one for a minimum ionizing particle orthogonal to the detector. For the particles at very low angle to the detector the signal is lower by a factor of 3 (for layers 1 to 3) due to the detector geometry (readout pitch vs. thickness). With a signal to noise ratio of only 6 to 1 the median finding becomes much less precise, reducing the resolution for such tracks to be little better than the readout pitch.

No detailed simulations of the other chip parameters exist. Especially the threshold variation was an unknown factor. Similar threshold variation measurements performed at Pavia [23] using analog probes to determine the shaper output voltages arrived at a variation of 9 mV. The results obtained in this thesis are 25% higher. This is most probably caused by the additional variation introduced by the comparator and digital readout.

One of the only remaining problems on the chip seems to be the excessive channel to channel influence in the form of both crosstalk and diminishing ToT value if many channels are fired. Both of these effects seem to be an artifact of the calibration charge injection circuitry, so the performance for data taking is not diminished.

The crosstalk using external charge injection seems to be nonexistent. This outstanding performance can only be attributed to the careful and consequent decoupling of the analog and digital sections.

Appendix

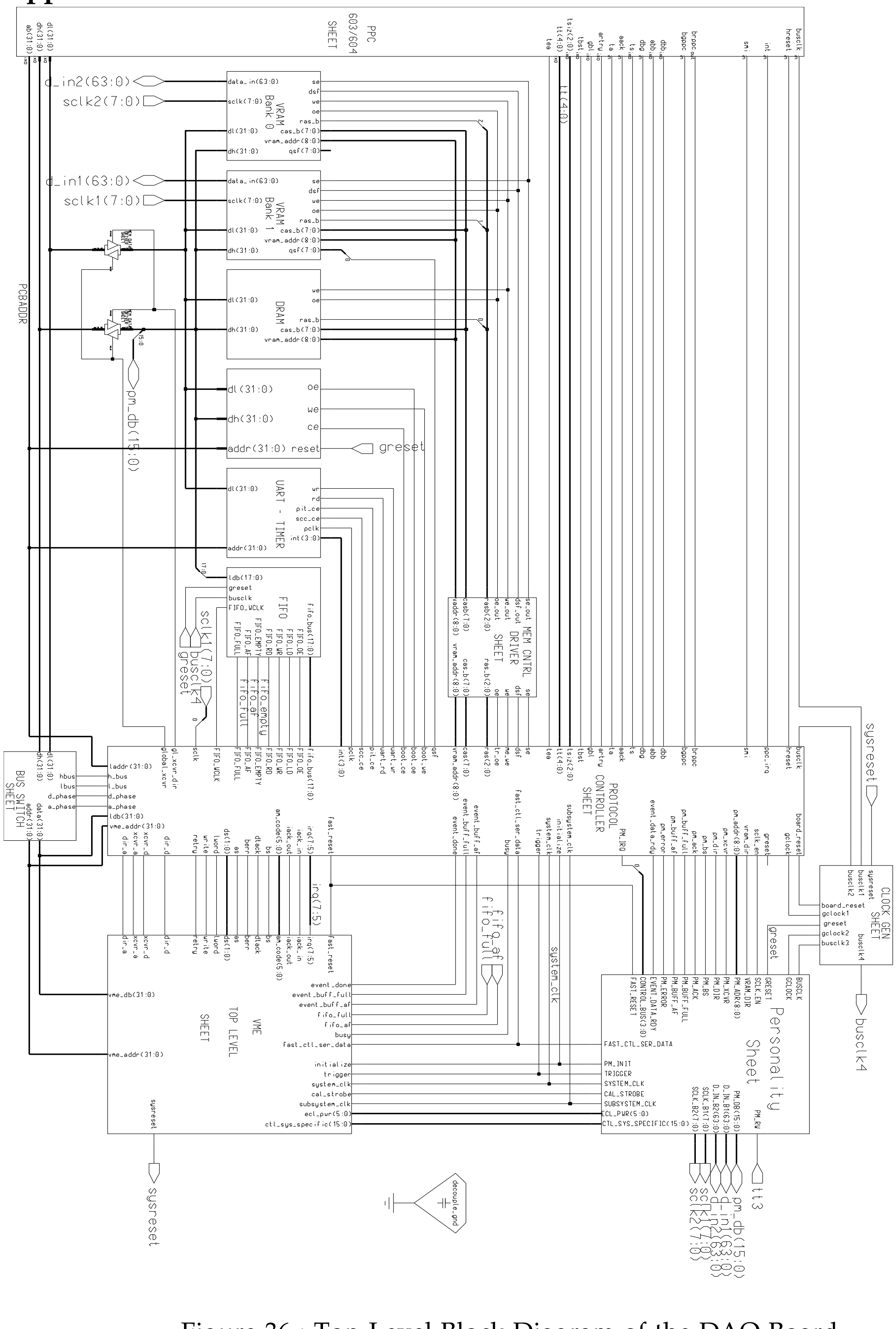


Figure 36 : Top-Level Block Diagram of the DAQ Board

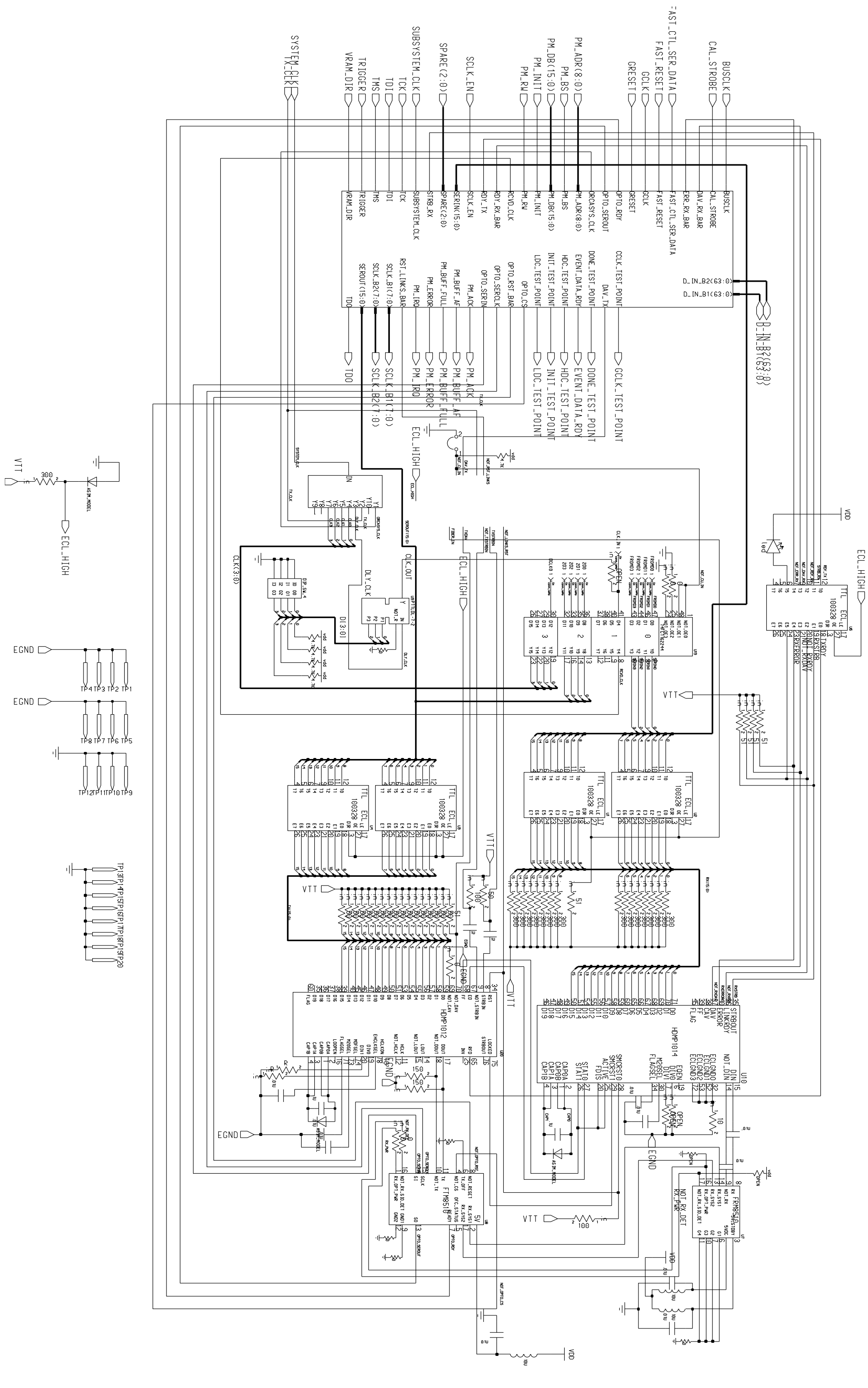


Figure 37 : Block Diagram of the Personality Module

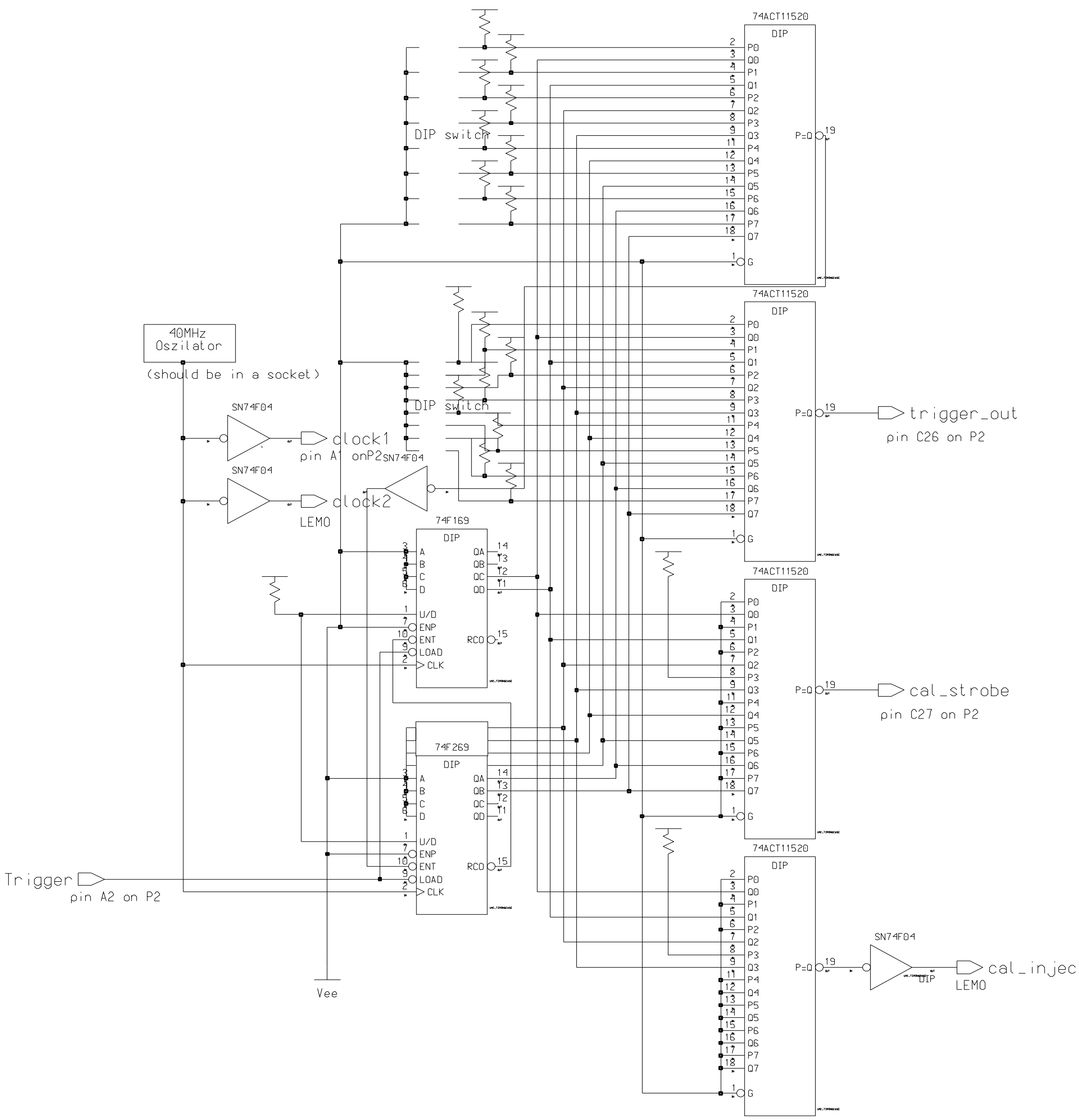


Figure 38 : Schematic of the minimal Fast Control System

Acknowledgments:

I want to give special thanks to :

- Michael Levi for a large measure of patience and an even greater number of suggestions on how to improve this thesis. He also gave me all the time I needed to write down my thesis by keeping other pressing work from me.
- Natalie Roe for giving me confidence in my thesis and her extremely good choice of literature for me to read.
- Claudio Campagnari for his help with the software, especially for his quick answers even on weekends.
- Robert Minor and Jimmy Johnson for getting hold of all those hard to find parts.
- Helen Chen for performing incredible feats with her soldering iron.
- Michael Momayezi for patiently answering all my questions in great detail.
- Windell Oskay for his support, both moral and practical during those darkest hours when nothing seemed to work.
- And last but not least to Prof. Norbert Wermes for letting me roam abroad to write my thesis.

References Cited

1 The BaBar Collaboration, BaBar Technical Design Report, chapter 4, SLAC, March 1995

2 I. Kipnis et al. : A Time-over-Threshold Machine: the Readout Integrated Circuit for the BaBar Silicon Vertex Tracker, Presented at the IEEE Nuclear Science Symposium in Anaheim, California, November 1996; Submitted for publication.

3 A.S. Schwarz: Silicon Strip Vertex Detectors at Z^0 Factories, Max Planck Institut für Physik, München, Germany, July 1992

4 The BaBar Collaboration, BaBar Technical Design Report, chapter 2, SLAC, March 1995

5 A.S.Grove: Physics and Technology of Semiconductor devices, John Wiley & Sons 1967

6 C.J.S. Damerell: Vertex Detectors - Lectures Presented at the Advanced Study Institute on Techniques and Concepts of High Energy Physics, St Croix, June 1986

7 The Particle Data Group, Review of Particle Physics, Physical Review D, volume 54, July 1996

8 B. Rossi, High Energy Particles, Prentice Hall, Inc. Englewood Cliffs, NJ, 1952

9 P. Horowitz, W. Hill, The Art of Electronics, Second Edition, Cambridge University Press, Cambridge, 1980

10 F. Manfredi: Low Noise Devices and Amplifiers, lecture notes

11 A. Karcher , F. Kirsten, M. Levi, R. Minor, W. Oskay, S. Wunduke: Readout Module Prototype : Functions and Operation, LBNL, Sep. 1996

12 A.Leona, M.Magna: ENC measurements on the AToM chip, Universita di Pavia, Sep. 1996

13 E. Gottlieb et al.: An Object-Oriented Software Interface for VME, University of New Mexico, November 1995.

14 J.K. Ousterhout: Tcl and the Tk Toolkit, Addison Wesley Publishing Company, Reading, Ma, 1994

15 MVME1604 Single Board Computer User's Manual, Motorola, Inc. Computer Group, 1994

16 The VxWorks operating system reference manual, Wind River Systems, Mountainview, California, 1995.

17 E. Mandeli, M. Melani and M. Pedrali-Noy :The AToM die testing report, LBNL, Sept. 1996

18 I. Kipnis, informal note, November 1996

19 F. Lani, informal note, December 1996

20 T. Collins : Minutes of the IC meeting , LBL, September 1996

21 G.Batignani et al.: Specifications and Quality Control Procedures of Silicon Detectors for SVT, BaBar Note 312, July 1996

22 W. Dabrowski, R. Johnson: Simulations of the Performance of the Proposed BaBar SVT Modules, BaBar Note 263, November 1995

23 P.F. Manfredi : Pedestal variation of the AToM chip, Talk presented at the BaBar Detector Collaboration Meeting, October 1996.