Design and performance of a CMOS test sensor for a binary readout electromagnetic calorimeter

J. A. Ballin[*], J. P. Crooks[†], P. D. Dauncey[*], A.-M. Magnan[*], Y. Mikami[‡],
O. D. Miller[‡], M. Noy[*], V. Rajovic[‡], M. M. Stanitzki[†], K. D. Stefanov[†], R. Turchetta[†],
M. Tyndel[†], E. G. Villani[†], N. Watson[‡], J. A. Wilson[‡]

Primary contact: P.Dauncey@imperial.ac.uk

September 24, 2008

Abstract

We present a study of a novel concept for a sampling electromagnetic calorimeter. Using binary readout with very small pixels, then an estimate of the number of charged particles can be made. This is shown to give potentially better resolution than a more standard analogue readout of the deposited energy. A CMOS test sensor has been designed, fabricated and tested to investigate the parameters of such a binary calorimeter and results from the sensor are presented. The expected resolution of a full size binary calorimeter is shown based on the results from the test sensor.

[*]Department of Physics, Blackett Laboratory, Imperial College London, London.
[†]STFC, Rutherford Appleton Laboratory, Chilton, Didcot.
[‡]School of Physics and Astronomy, University of Birmingham, Birmingham.

Contents

1 Introduction and motivation

In this paper, we present results from a test sensor which was designed to investigate a novel approach to electromagnetic calorimeters (ECAL). In this approach, the readout of the ECAL is binary, meaning that each channel returns a yes/no result with no further amplitude information.

Although the potential applications for a binary ECAL are wide, the sensor discussed in this paper was designed for a specific application, namely the International Linear Collider (ILC) [1]. This is widely assumed to be the next worldwide particle physics frontier accelerator and it has a wide and rich physics programme [2]. The work presented in this paper has been done within the context of the CALICE Collaboration [3], which is studying calorimetry for the ILC.

Many of the ILC physics measurements require reconstruction of heavy particles decaying into hadrons. To identify the decaying particles through the invariant mass of the hadronic jets requires excellent hadronic jet energy resolution. The main techniques which should be able to provide this resolution are based on Particle Flow Algorithms (PFA) [4, 5], which require very high granularity calorimetry to allow the separation of the showers into the individual particles in the jet. For the ECALs, it is widely accepted that a sampling calorimeter will provide the best PFA performance, due to both transverse and longitudinal granularity. The current state-of-the-art has silicon sensors as the detecting layers and tungsten as the converter and this structure is being considered by both the major ILC detector concepts [6, 7].

An electron, positron or photon interacting in material produces an electromagnetic (EM) shower containing a large number of charged particles and photons. A sampling EM calorimeter (ECAL) measures the charged particle component of the shower at various positions in the shower depth by having sensitive layers interspersed with a dense material which efficiently produces the shower. The sensitive layers operate by measuring the ionisation created by the charged particles in these layers and normally the shower energy is estimated by measuring the energy deposited through the ionisation.

In this paper, we present a new technique for a sampling ECAL, in which an estimate of the number of particles in the shower is used as the measure of the shower energy, rather than the energy deposited. Clearly, these quantities are strong correlated, as each charged particle in the shower will deposit energy in the sensitive layer. However, the energy deposited depends on the speed and angle of the charge particle and also has fluctuations around the average deposit, as described by the Landau function. Hence, it would be expected that the energy deposited would have a larger spread for any given shower than the number of particles. This is illustrated in figure 1 which shows the energy deposited and the number of charged particles in simulated photon-induced EM showers using the GEANT4 program [8]. The ECAL geometry used for this example had 30 layers of $500\,\mu$m thick silicon sensors, with 0.6 radiation lengths (X_0) of tungsten in front of each of the first 20 layers and $1.2X_0$ in front of each of the last 10 layers. This is typical of the designs being considered [6, 7]. For all energy and resolution estimates below, the information from the last 10 layers is weighted by a factor of two compared to the first 20 layers. No experimental effects have been included in this simulation so the analogue energy deposits are assumed to be measured perfectly, with no resolution effects. The number of charged particles is determined from the perfect simulation information as the number of separate particles passing through the last $1\,\mu$m of the silicon layer.

Figure 2 shows the average value of each measure as a function of photon energy and shows both estimates demonstrate good linearity. Figure 3 shows the root mean square (RMS) of each measure as an estimate of the basic irreducible resolution in the two cases. It is clear the number of charged particles gives a very significant improvement over the energy deposited as an estimator of the incident photon energy. Fitting these resolutions to a function of the form

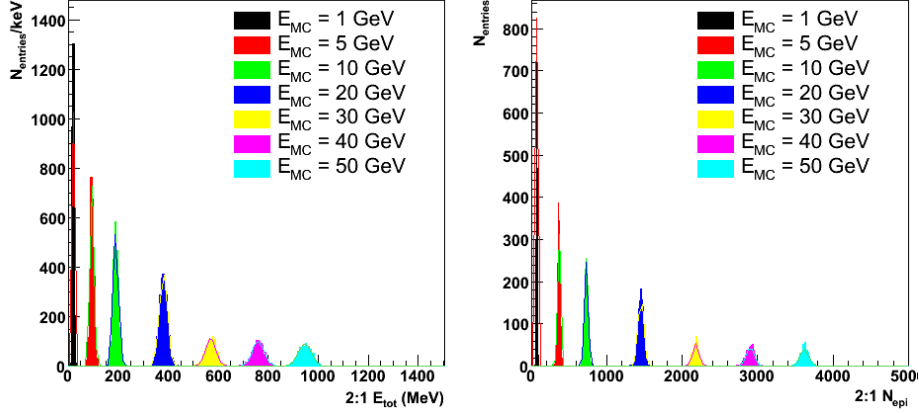$$\frac{\sigma_E}{E} = \frac{s}{\sqrt{E \text{ in GeV}}} \oplus c$$

Figure 1: *Comparison of the distributions of the quantity measuring shower energy using (left) the energy deposited in the sensitive layers and (right) the number of charged particles crossing the sensitive layers. Results for a range of incident photon energies from 1 to 50 GeV are shown.*
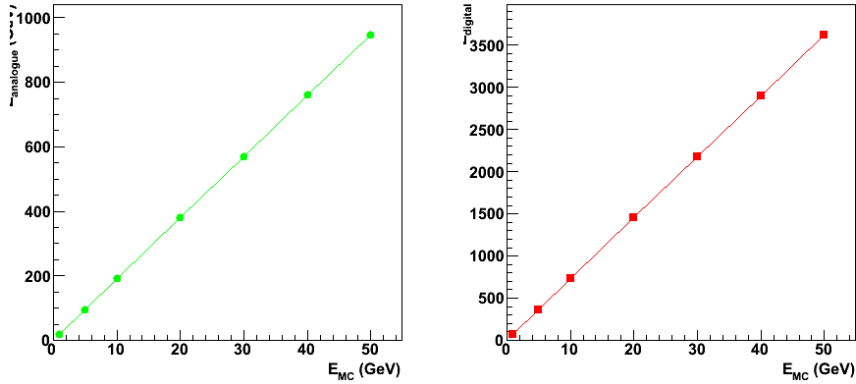

Figure 2: *Comparison of the linearity as a function of the incident photon energy, when measuring shower energy using (left) the energy deposited in the sensitive layers and (right) the number of charged particles crossing the sensitive layers.*

give values for the stochastic terms s of 14.6% for the "analogue" (energy deposited) case and 9.5% for the "digital" (number of charged particles) case, and for the constant terms c of 1.0% and 0.8%, respectively. Hence, the analogue case is degraded by a factor of around 1.5 compared to the digital case. It is known that a realistic analogue ECAL with low noise electronics can approach the basic resolution of the analogue case. The purpose of the study reported in this paper is to estimate the resolution of a realistic binary ECAL and compare it with this ideal case.

In the rest of this paper, Section 2 expands on the ECAL and sensor requirements for the ILC application. The design and fabrication of the test sensor are described in Section 3. Section 4 and Section 5 describe results from the test sensor for individual pixels and sensor-wide measurements, respectively. Section 6 gives the expected physics performance of a binary ECAL based on the results from the test sensor.
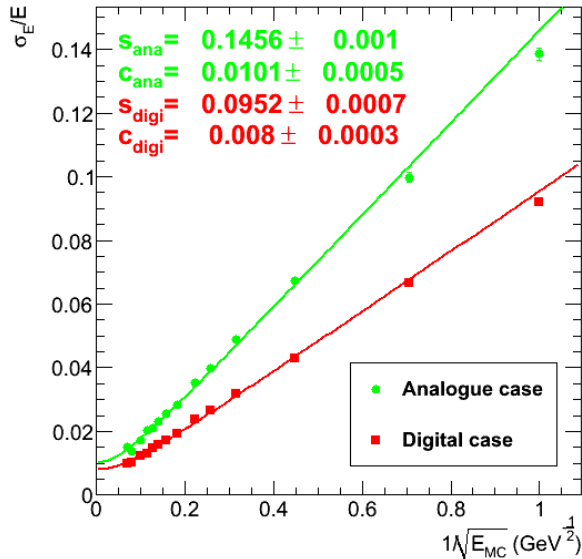
Figure 3: *Comparison of the resolution dependence on the incident photon energy, when measuring shower energy using (green) the energy deposited in the sensitive layers and (red) the number of charged particles crossing the sensitive layers. The lines show the results of the fits to the data points, as described in the text.*

2 Requirements

The results shown in the previous section motivate the development of a sensor which is capable of estimating the number of charged particles in a shower. The method considered here is to propose an ECAL with binary readout, such that each pixel of the calorimeter gives a yes/no result to the presence of a charged particle. Any charged particle gives an average energy deposit which has a universal minimum at a particular particle velocity and for normal incidence. Any other speed or incident angle gives a larger mean deposit, so a threshold set below the minimum ionising particle (MIP) level, allowing also for the Landau spread, should be efficient for detecting all charged particles.

2.1 EM shower requirements

To give an accurate estimate, a sensor for counting charged particles would need to have a low probability of two charged particles crossing the same pixel. The density of charged particles in the layer at which the density is a maximum for $100\,\mathrm{GeV}$ EM showers is shown in figure 4 for the example sampling ECAL used in Section 1. The peak value corresponds to 70 charged particles per mm^2; this would therefore require a pixel size significantly smaller than $0.014\,\mathrm{mm}^2$. The sensor discussed in the rest of this paper has a pixel size of $0.0025\,\mathrm{mm}^2$, implemented as a square pixel of size $50 \times 50\,\mu\mathrm{m}^2$.

2.2 ILC requirements

As shown above, the pixel size for a binary ECAL need to be of the order of $50\,\mu\mathrm{m}$. This is two orders of magnitude smaller than any of the analogue pad ECALs being considered for the ILC [6, 7] and will be more than sufficient to allow PFA separation of particles. The size of the ECALs assumed in the ILC detector studies is large with a total sensor surface area of around
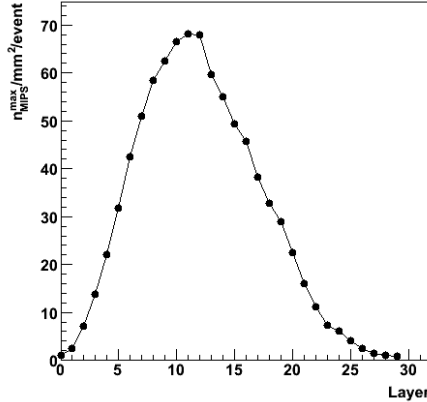
Figure 4: *Maximum density of charged particles per layer within $100\,GeV$ EM showers, as a function of the layer in the shower.*

$2000\,\text{m}^2$. For $50\,\mu\text{m}$ pixels, this would require a total of around 10^{12} pixels; hence this concept has been labelled the "tera-pixel" ECAL.

The ILC will run with bunch trains, whereby the beams will collide (a "bunch crossing") at a rate of around 200-400 ns but only for a short time, of around 1 ms (a "bunch train"). Each bunch train will therefore contain up to 5000 bunch crossings. The bunch train rate will be around 5 Hz, meaning that there will be 199 ms with no beam collisions following each 1 ms bunch train. Hence, the ILC detectors will buffer data during the 1 ms bunch train and then read out in the 199 ms of dead time.

The physics interaction rate for producing the high energy events of interest will be low; less than one per bunch train [2]. There will be significant levels of background from the ILC machine itself, as well as low energy physics events. However, with such a large number of pixels in a binary ECAL, then the overall data rate from most of the detector will be dominated by the rate of noise. Assuming a threshold of around five times the noise, then the noise rate will be $\sim 10^{-6}$ per pixel per bunch crossing. This means for the 5000 bunch crossings in a bunch train, each pixel will have an average of ~ 0.005 hits. Hence, the sensor needs to be able to buffer at least this number of hits per pixel during the bunch train.

The one area where other sources of hits will not be negligible will be in the most forward regions, where backgrounds from the machine will be substantial. The beam simulation programme Guinea Pig [9] allows such backgrounds to be estimated. Figure 5 shows the expected number of hits in the endcap region as a function of radius from the beam line for the ILD detector. The peak at the innermost radius corresponds to 150 hits per bunch crossing at 300 mm radius, within a 2 mm band in radius. This corresponds to a total area of $4000\,\text{mm}^2$ and so an average occupancy of 10^{-4}. Hence for the innermost endcap pixels, the probability of getting a hit in a bunch train could be as high as 0.5.

2.3 Other issues

An ECAL designed to deliver good PFA performance will also need to be compact, minimising any gaps between the converter layers. This keeps the transverse spread of the EM showers small, measured by the effective Molière radius, as this helps separate particles close to each other in jets. Hence, the sensitive layers have to be as thin as possible compared to the converter thicknesses. For example, tungsten has a radiation length of $X_0 = 3.5\,\text{mm}$. Most ILC ECAL designs [6, 7] have only $\sim 0.5 X_0$ in each of the first few layers, which therefore corresponds to
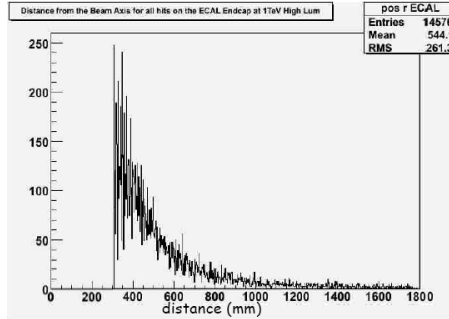
Figure 5: *Simulated number of ILC machine background hits per bunch crossing in the endcap region, as a function of the radius from the beam line.*

~ 2 mm of tungsten. Hence, sensitive detectors based on thin silicon sensors seem a sensible choice.

The compactness requirement also significantly limits the amount of cooling infrastructure which can be included between the converter layers. Any pipework will increase the space, degrading the effective Molière radius. Hence, this requires very low power sensors and readout electronics. The timing structure of the ILC means there are significant gains to be made by powering off any parts of the sensors which are not required during the 199 ms dead period. Even including this reduction of O(100) in power consumption, the requirements are tight. For example, one design [7] has no active cooling in the bulk of the ECAL, relying purely on conduction through the converter material and the electronics PCBs. The target average power consumption for this design is $1 \, \mu\text{W}/\text{mm}^2$.

A very significant issue with all such calorimeters is cost. With a sensitive area of $2000 \, \text{m}^2$, then the total cost will be dominated by the silicon detectors themselves. Hence, there is a strong motivation to reduce the sensor unit area cost. This pushes for simple sensors which can be produced through a large number of vendors and require no low-yield processing features.

3   Design and fabrication

In this paper, we present results from a test sensor which was designed to investigate a binary ECAL and to demonstrate the feasibility of this approach. Because a complete ECAL for this application would require a total number of pixels of the order of 10^{12}, this development has been named the Tera-Pixel Active Calorimeter (TPAC) sensor. This paper reports on results from the first test version of the sensor, TPAC1.0, which was based on a $0.18 \, \mu\text{m}$ CMOS process.

The use of CMOS structures for light-sensing applications was originally proposed in the 1970s. Numerous recent developments to the pixel design and processing, such as the pinned photodiode, have achieved high-grade performance in areas such as low noise and leakage current. Such performance can be applied beyond the commercial imaging domain, and CMOS sensors have been demonstrated in many alternative applications including particle physics [10].

A key advantage of CMOS imaging is the use of standard CMOS fabrication processes, in which a diverse family of NMOS and PMOS transistors, resistors, capacitors and diodes can be manufactured on the same silicon substrate as the sensitive pixel; hence these sensors are often referred to as Monolithic Active Pixel Sensors (MAPS). In typical imaging applications this allows for the integration of the readout electronics and control logic at the edge of the array of imaging pixels, offering a very compact one-chip imaging solution. The requirements for particle physics applications of CMOS sensors share many aspects with those for commercial imaging, such as low noise performance, but the pixel pitch needed is usually larger, with sizes up to

7

$100\,\mu$m. This area provides an opportunity to exploit the full range of components available in the CMOS process, with the possibility to add significant amounts of circuitry inside these larger pixels, rather than around the periphery of the sensor.

3.1 Process technology

The CMOS silicon substrate comprises a very low resistivity base material, over which a P-doped epitaxial layer is grown, generally up to $20\,\mu$m in thickness. For the application described here, an epitaxial layer thickness of $12\,\mu$m was used. The silicon substrate and epitaxial layer are predominantly free of electric fields, so any charge that has been deposited in the silicon will move randomly by diffusion, with typical carrier lifetimes of milliseconds. A small potential barrier exists due to the change in substrate doping between the silicon substrate and the epitaxial layer, which is sufficient to keep the majority of carriers within the epitaxial layer. A small electric field forms around a positively charged N-well diode, in which a diffusing electron will be swept towards the N-well and collected as signal. Figure 6 illustrates the basic principles.
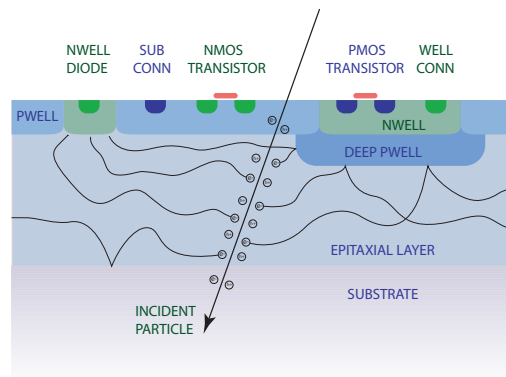

Figure 6: *Conceptual illustration of a charged particle crossing a CMOS sensor epitaxial layer. The N-well diode at the upper left is a signal collecting diode. The one at the upper right encloses a PMOS transistor which is part of the in-pixel circuitry. The latter is shielded from absorbing signal charge by the deep P-well implant between it and the epitaxial layer. UPDATE WITH VERSION WITH ARROWS.*

As stated above, the epitaxial layer is the detecting volume where charge is generated. A good approximation is to consider that the total number of electron-hole pairs generated by a MIP in silicon is equal to $80/\mu$m. For the $12\,\mu$m thick epitaxial layer used in TPAC1.0, this corresponds to only about 1000 electron-hole pairs. There will also be some contribution to the charge generation from the upper few microns of the substrate which increases the MIP signal charge to around $1200\,e^-$. However, given the charge diffusion between pixels, the number of charge carriers collected by any single pixel is smaller. Clearly, any further loss, specifically due to charge collection by unrelated N-wells, would make the efficient detection of MIPs in CMOS sensors very difficult, if not impossible.

Due to diffusion, the signal charge will not necessarily be collected by the nearest diode, in which case it continues to diffuse and may be collected by a diode in a neighbouring pixel. In particle physics tracking applications or in imaging sensors this effect is called "crosstalk" and represents an undesirable loss of position accuracy or image quality, respectively. Although this is less of an issue for an ECAL sensor, the pixel design presented in this paper implements

four diodes, placed toward the corners of the pixel for optimum crosstalk performance. The advantage of this approach is also a reduction in charge collection time, as the mean distance to any diode is shorter than the single-diode solution at the same pixel pitch.

The structure of a standard PMOS transistor presents difficulties for use in a pixel design, due to the N-well in which the PMOS device sits. Such N-wells, tied to a positive potential, are as equally likely to collect diffusing charge as the collecting diode. However, any charge which reaches these N-wells will not be collected as signal on the diode, and so this behaviour represents an inefficiency in charge collection. For this reason, commercial CMOS sensors normally only use NMOS transistors within the pixel, although this significantly limits the circuit functionality that can be implemented. For a complex pixel design with many PMOS transistors and small collecting diodes, this inefficiency would dominate the charge collection and the resulting signal size would be too small to resolve over the electronic noise.

To address the problem of inefficiency caused by PMOS transistors in the pixel, a deep P-well implant can be added to the standard CMOS process. This high-energy implant creates a region of higher doped P-type silicon beneath the N-well of a PMOS transistor, as shown in figure 6. The small potential barrier that forms, much like the boundary between the epitaxial layer and the substrate, is again sufficient to keep the majority of carriers within the epitaxial layer, and most importantly, from being collected by the N-well. This technique restores the charge collection efficiency of the sensor to be close to 100%, although some minor reduction in the initial signal charge deposited must be expected from the small proportion of the epitaxial layer that is now occupied by the deep P-well implant.

For the successful implementation of this project, a deep P-well module was developed by a leading commercial foundry, on a standard $0.18\,\mu$m CMOS process. This process is called "Isolated N-well MAPS" (INMAPS) and is described in detail elsewhere [11]. The INMAPS process features six metal layers, precision passive components for analogue circuit design, and may be stitched to manufacture sensors up to wafer scale. Whilst for this project a particular commercial partner was selected, the technology could be implemented in most modern CMOS processes.

The TPAC1.0 sensor was manufactured using the INMAPS process, and implemented deep P-well implants in the pixels to achieve a high charge collection efficiency. To further understand the device physics, the same design was manufactured with two thicknesses ($12\,\mu$m or $5\,\mu$m) of the epitaxial layer and with or without the deep P-well implant.

3.2   Overall architecture

The TPAC1.0 sensor comprises 28,224 pixels, row control logic, on-sensor SRAM memory banks and I/O circuitry in a $9.7\times10.5\,\mathrm{mm}^2$ die. The sensor collects the charge deposited by an incident particle in pixels arranged on a $50\,\mu$m pitch. This signal is compared with a global threshold and if above threshold, the time-code and location of the pixel "hit" is recorded in memories for readout at a later time.

Four different pixel designs have been implemented for evaluation, which fall into two distinct architectures, called pre-shape and pre-sample. These four designs are arranged in four quadrants of 84×84 pixels. All pixels contain four small N-well diodes for charge collection. Pixels may be individually masked, allowing any permutation of single pixels to be operated and evaluated.

A common control and readout architecture serves all pixel varieties, allowing the sensor to be operated as a whole or as sub-regions. Rows of 42 pixels share row control logic and SRAM memory, while columns of 168 such rows form a region which shares data readout. All four regions multiplex their data for external readout off-sensor. An overview of the architecture is shown in figure 7, together with a photograph of the TPAC1.0 sensor.
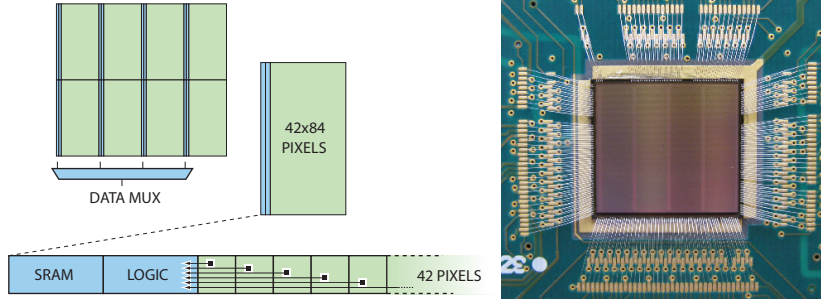
Figure 7: *(Left) Overview of sensor architecture; see text for a detailed explanation. (Right) Photograph of the sensor mounted on its PCB. The four pixel regions and memory areas are visible within the central area of the sensor.*

3.3 Pre-shape pixel

The pre-shape pixel is based on a conventional analogue front end for a charge-collecting detector and is shown in figure 8.
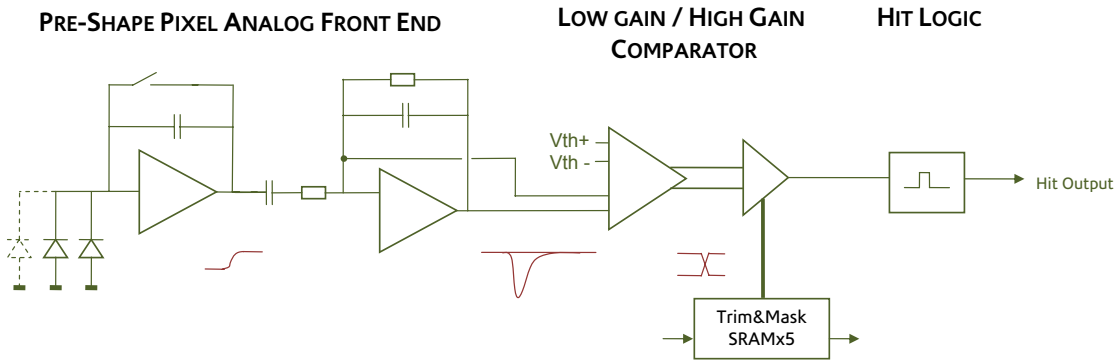

Figure 8: *Pre-shape pixel circuit; see text for a detailed explanation.*

Four diodes are connected to a charge preamplifier, which generates a voltage step output in proportion to the collected charge. A CR-RC shaper circuit with a peaking time of 200 ns generates a pulse output in proportion to the input signal with further circuit gain to yield $94\,\mu\mathrm{V}/e^-$ with respect to total input charge. The simulated response of the charge preamplifier to signals of varying magnitude is shown in figure 9. This signal, along with a local common-mode reference form a pseudo-differential input to the two-stage comparator.

Following a signal, the shaper circuit spontaneously returns to a stable state after a time depending on the signal size, and is then able to respond to another input signal. Saturation in the shaper circuit occurs for signal charge deposits greater than $2500\,e^-$, beyond which the shaper output becomes non-linear and takes a longer time to return to the steady state. Non-recoverable saturation of the pixel occurs when the preamplifier stage has integrated $10\,\mathrm{k}e^-$ on the diode node, beyond which the gain of the pixel deteriorates, reaching 50% after $22\,\mathrm{k}e^-$, until it will no longer respond to an input signal. The preamplifier reset is used to initialise the pixel before the start of a bunch train. The expected incident signal charge for a single pixel during a bunch train is small compared to these saturation limits.

To achieve high circuit gain in the preamplifier, a small value of the feedback capacitance was required, which was made using two larger capacitors in series to comply with manufacturing
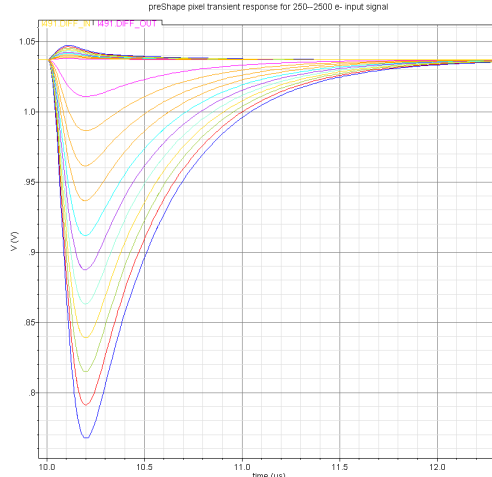
Figure 9: *Pre-shape pixel circuit response as a function of time to input signals injected at a time corresponding to 10 μs on the time axis. The input signals shown have various magnitudes up to a maximum of $2500\,e^-$, in steps of $250\,e^-$.*

design rules. Two different simulation tools were used to evaluate the optimum orientation of the series feedback capacitors, but the two tools selected different topologies for highest gain. Two capacitor orientations are therefore implemented on the TPAC1.0 sensor as subtle variants of the pre-shape pixel and these occupy quadrants 0 and 1.

The in-pixel comparator has two stages. The first takes two differential signals, and produces a real-time differential discrimination result, with some small analogue signal gain. The second comparator generates the full-swing discriminator output, and applies an offset trim adjustment with a four-bit resolution. The output of the second comparator is enabled with a one-bit mask which can be used to prevent the pixel from generating hit outputs.

Pixels which go above threshold generate a fixed length pulse using a monostable circuit, the output of which is connected to row control logic outside the pixel. The length of the resulting hit output pulse is independent of the signal size.

Electronic circuit noise is estimated at the input to the differential comparator and referred back to the diode node using the charge gain. The dominant noise source is the input transistor of the preamplifier circuit; the expected equivalent noise charge (ENC) for this pixel is $23\,e^-$ RMS.

The nominal power consumption for the pre-shape pixel is $8.9\,\mu$W during operation, although the circuit would be powered off between bunch trains at the ILC. Assuming a factor of 100 reduction in power by this method, this is equivalent to an average power consumption of $36\,\mu$W/mm^2 and hence is is significantly larger than the target power of the analogue ECAL designs (see Section 2). However, the main aim of the TPAC1.0 sensor is to investigate the concept of a binary ECAL and power was not regarded as a major constraint in its design. Lower power designs will be considered in future versions of the sensor.

3.4 Pre-sample pixel

The pre-sample pixel is based on a conventional MAPS sensor, with in-pixel analogue storage of a reference level, and is shown in figure 10.

Charge integrates on the four collecting diodes, causing a small voltage step proportional to the collected charge and the node capacitance. A charge preamplifier provides gain to yield $440\,\mu$V/e^- as a voltage step which, along with a local sample of the reset level, forms a pseudo-

PRE-SAMPLE PIXEL ANALOG FRONT END   LOW GAIN / HIGH GAIN COMPARATOR   HIT LOGIC
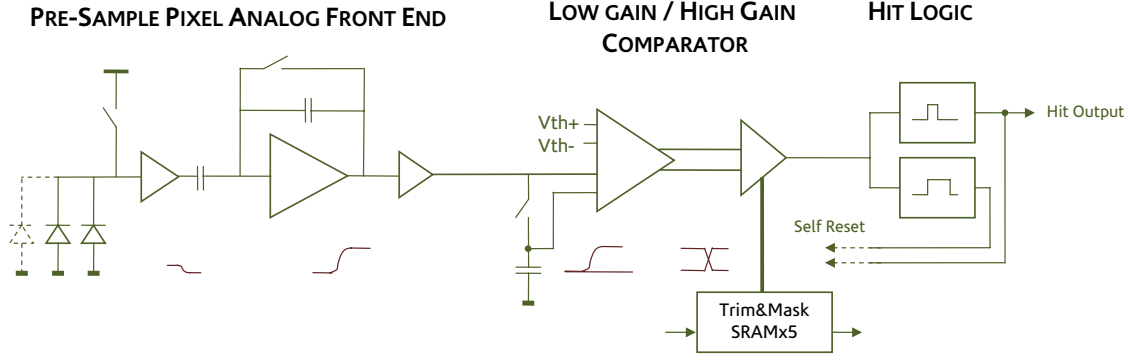
Figure 10: *Pre-sample pixel circuit; see text for a detailed explanation.*

differential input to the same two-stage comparator as for the pre-shape design. The charge amplifier and reference sample must be reset after a hit before the pixel can detect another hit; this is undertaken by the in-pixel logic.

The simulated response of the circuit to signals of varying magnitude is shown in figure 11. Saturation in the pre-sample pixel occurs when the diode node has integrated $64\,\mathrm{ke^-}$, beyond which non-linear operation is expected.
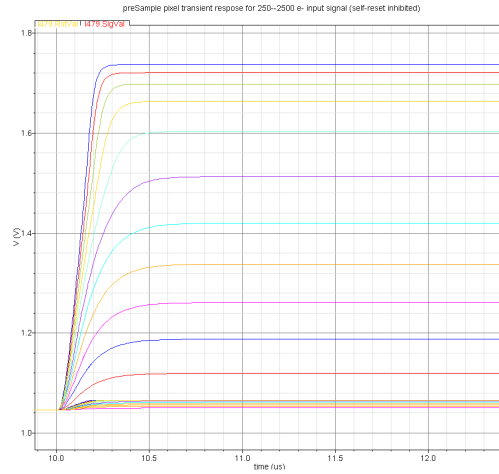

Figure 11: *Pre-sample pixel circuit response as a function of time to input signals injected at a time corresponding to $10\,\mu s$ on the time axis. The input signals shown have various magnitudes up to a maximum of $2500\,e^-$, in steps of $250\,e^-$.*

Similar to the pre-shape pixel, a small capacitance in the preamplifier feedback is made with two capacitors in series. This gives rise to two subtle variants of the pre-sample pixel, again based on results from different simulator tools, and these occupy quadrants 2 and 3.

The in-pixel comparator stage is common to all pixel architectures, but the pre-sample pixel includes an additional monostable circuit to generate the self-reset signals that are necessary to reset the amplifier and the reference sample in preparation for another input signal.

Electronic circuit noise is estimated at the input to the differential comparator and referred back to the diode node using the charge gain. The estimate of the observed noise requires an additional factor of $\sqrt{2}$ to be applied to the real-time simulation noise level to account for the uncorrelated sampling nature of this pixel. The dominant noise sources are the input transistor

of the first source follower buffer and the input device in the charge amplifier circuit. The expected ENC for this pixel is $22\,e^-$ RMS.

The nominal power consumption for the pre-sample pixel is $9.3\,\mu$W during operation, although again the circuit would be powered off between bunch trains at the ILC.

3.5   Mask and trim configuration

Each pixel contains a five-bit SRAM shift register, which is used to store a per-pixel trim (four bits) and a mask flag (one bit). The configuration shift registers are not used while the sensor is in active operation, but these configuration data are programmed during sensor initialisation, and are held indefinitely in each pixel until the sensor is powered down or the data are rewritten. The configuration data are loaded through a serial interface, which shifts single-bit data into each of 168 columns simultaneously, as illustrated in figure 12. Serial data outputs are available at two points to enable data read-back for error-rate monitoring. The data read is destructive, so in normal operation the read-back occurs after hits have been collected. The total configuration memory space on the sensor is 141 kbits.
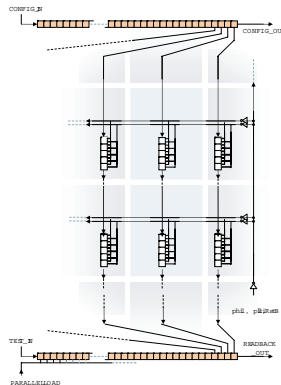

Figure 12: *Overview of logic for configuration data write and readback.*

3.6   Pixel layout

The pre-shape and pre-sample pixel layouts are illustrated in figure 13. The pre-shape pixel comprises 160 transistors, 27 capacitor cells and a large polysilicon resistor. The pre-sample pixel comprises 189 transistors and 34 capacitor cells. The two pixel architectures use the same diode positions, which were optimised for charge collection using the simulation described in section 4.1, and are located towards the corners of the $50\,\mu$m pixel.

The sensitive analogue front-end circuits are located in the very centre of the pixel with extensive substrate-grounded guard rings for signal integrity. Analogue signals are routed primarily on metal layer 1, with some plates of metal layer 2 used where necessary to shield the
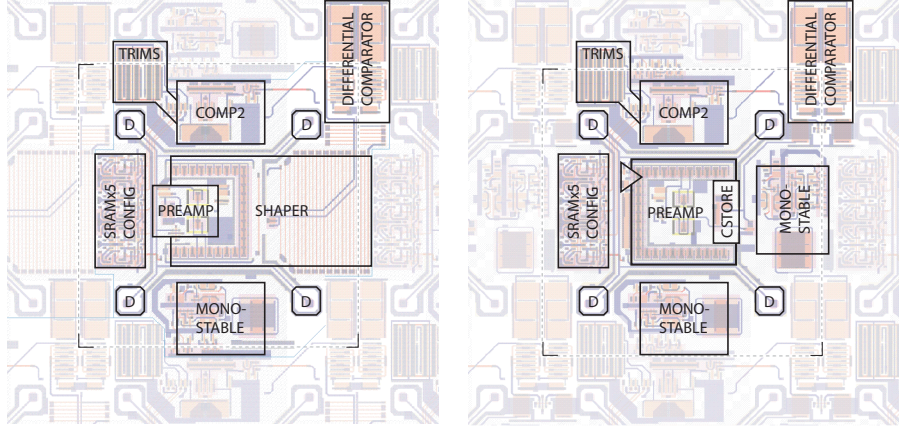
Figure 13: *Pixel layouts; (left) pre-shape and (right) pre-sample. The boxes labelled "D" indicate the diode positions. The nominal borders of the pixels are shown by the right-angled corners and dotted lines.*

analogue signals from switching signals passing overhead. The deep P-well layer is added as a symmetrical cross structure which leaves only the collecting N-well diodes exposed to the epitaxial layer. as shown in figure 14.
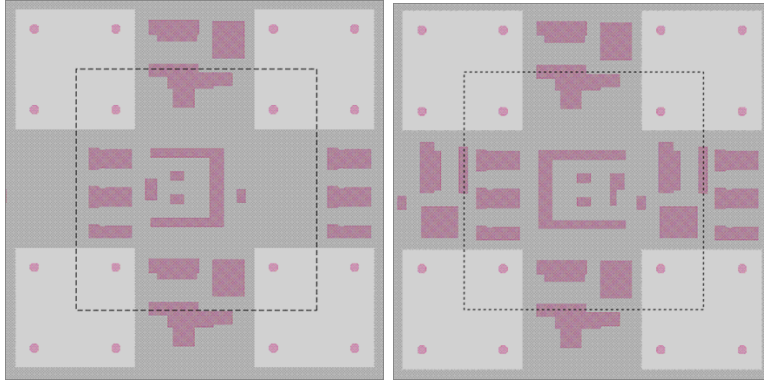

Figure 14: *Pixel deep P-well implant layouts; (left) pre-shape and (right) pre-sample. The nomnal border of the pixel is shown by the dashed lines. The grey shows the deep P-well implant area while the purple shows the N-well structures. All the N-well regions are protected by deep P-well implant except for the four signal diodes towards the pixel corners.*

Pixel power supplies are routed on the metal layers 3, 4 and 6 in horizontal and vertical directions to distribute power in a mesh structure. The various sub-circuits in each pixel design are mostly powered separately in this first sensor, so there are five independent power supplies routed to the pixels. The hit output signals from pixels are routed horizontally along a row on metal layer 5, which is fully shielded from the sensitive analogue electronics below by the metal layers in between.

3.7 Row control logic

The row logic is responsible for monitoring the individual hit outputs from a row of 42 pixels and writing details of any hits to local memory. An external clock defines the timing with

which hit signals are sampled. The hit signal from a pixel is asynchronous, but will have a fixed output pulse width defined by the in-pixel monostable bias setting. This pulse length would typically be set to be O(10%) greater than the hit sampling period, which would be matched to the bunch crossing rate of the target application, typically 200-400 ns. This regime ensures that an asynchronous hit will always be sampled by the synchronous logic, with a small probability that it will be sampled twice. This is an acceptable data overhead that allows for a reasonable spread in the length of the monostable pulses, with a minimal risk that an entire hit pulse occurs between sampling such that the hit is lost. The sampling of hits uses a "ping-pong" circuit architecture to ensure there is no dead time between samples.

The 42 sampled hit signals are subdivided into seven "banks" for most efficient processing and storage. Each bank is selected for inspection in turn with a three-bit multiplex (MUX) address code. An OR circuit then tests the six pixel outputs of the bank to see whether the bank contains any hits. If so, the six-bit pixel hit pattern and the three-bit MUX address code are stored, thus identifying a single location in the full row of 42 pixels. This is illustrated in figure 15. A key feature of this approach is in the case of a dense particle shower, for which multiple nearby hits are stored in a single register, rather than multiple registers. The hit-seeking circuits operate at eight times the bunch crossing rate, i.e. up to 50 MHz. Multiplex signals are gray-coded for reliable high speed operation, and the reserved address value 0 deselects all banks for additional testing provision.
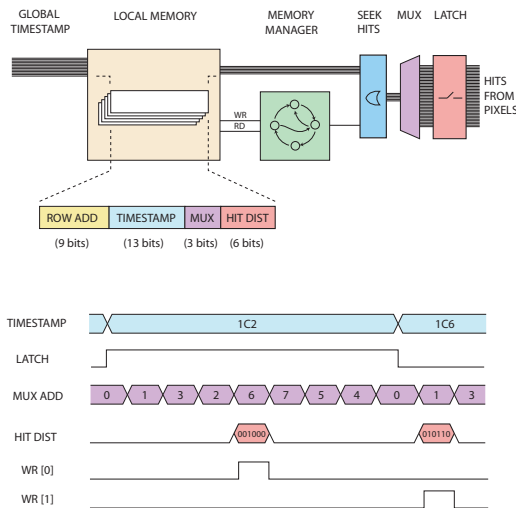

Figure 15: *Overview of the row logic for hit storage; see text for details.*

3.8 Data storage

The row control logic has 19 SRAM registers available for storage of hit data. A memory controller is implemented to organise the use of these registers, such that registers are not overwritten once used, and only those with valid data participate in readout. This memory controller is implemented as a bidirectional shift register, with 20 cells. The memory control register initialises with a token in the first position, which enables the first SRAM register for write access. During data write the shift register advances to the next position, filling the SRAM registers in order for the first 19 hits. A further hit in the row then moves the token into a holding cell, which raises a global overflow flag indicating the memory full status. This hit, and any further hits on that row, will be discarded.

The row control logic may be operated in "hit override" mode, whereby the result of the OR circuit is ignored and the value of the hit pattern in each bank is always stored. This operating mode fills the memories in less than three complete cycles of the standard control sequence, and so is intended as a test feature.

The 19 SRAM registers occupy the full $50\,\mu$m row pitch. The hit pattern and corresponding MUX address are stored in the first 9 bits of a register, with a further 13 bits used to store the global timestamp code, which is incremented each time hit signals are sampled. The cross-coupled inverter structure of a SRAM cell ensures the data will be held indefinitely provided the cell is powered, so there is no requirement to refresh the data and no maximum hold time after which data are corrupted.

The full TPAC1.0 sensor comprises four columns of row logic, each with 168 rows, hence there are 12,768 SRAM registers of 22 bits each in total. The row control logic and the SRAM register bank occupy a $250\,\mu$m wide region adjacent to the 42 pixels, equivalent to the area of five pixels. There is no sense circuitry in this region and, as no deep P-well is added here, charge arising from particles that pass through these regions will be collected by local N-wells associated with PMOS transistors in the logic and SRAM. This charge will not be collected as signal and the logic and SRAM are therefore insensitive to incident particles. In addition, an inactive $50\,\mu$m wide strip across the centre of the sensor corresponding to a single row of pixels, is used to distribute bias and reference voltages, and to re-buffer control signals. These inactive regions result in the sensor having an inherent 11% dead area.

3.9   Readout

During readout, the memory controller is switched into the reverse direction and clocked once to initialise the token to enable the most recently written register for readout. A combinational read-enable signal propagates to the first register that has valid data for readout, and enables the connection to the parallel readout bus. On each subsequent clock of the memory controller the next valid register is selected until no further registers remain, at which point a "done" output flag is asserted. The off-sensor control software uses this flag to initiate readout of each logic column in turn.

In addition to the 22-bit SRAM registers, a 9-bit ROM cell is activated during readout of each row. These extra cells encode a unique row address that forms part of the parallel data bus.

The SRAM and ROM readout is implemented with a current sense amplifier which was optimised to operate over long distances with minimal read time. An activated SRAM or ROM cell pulls current from the parallel data bus depending on its state, which is sensed by the circuit at the column base. A total of 31 of these current-sense amplifiers operate in parallel to create a 31-bit digital output data word, which is multiplexed and driven off-sensor with no serialization. The maximum read time from the furthest cell is $150\,$ns and parallel data readout is typically operated at a $5\,$MHz rate. A full sensor readout, in which every register contains valid hit data (such as when operating in hit override mode) therefore takes approximately $2.6\,$ms, and generates $50\,$kBytes of data.

3.10   Additional test features

In addition to the main design presented above, three pre-sample test pixels have been implemented which allow access to internal nodes for evaluation. The design includes additional analogue buffers to monitor internal analogue signals in the pixel circuit. These facilitate evaluation of the performance of monostable circuits, comparators, trim adjustment of threshold, and the analogue front end circuits for the pre-sample pixel architecture. The signal pulse and the reset sample are available for two adjacent pixels, and the internal differential comparator

output is available from one test pixel. A third pixel allows evaluation of other in-pixel circuits, including the two monostables and the full comparator chain. The circuit is shown in figure 16. These test pixels are included at the edge of the main pixel array.
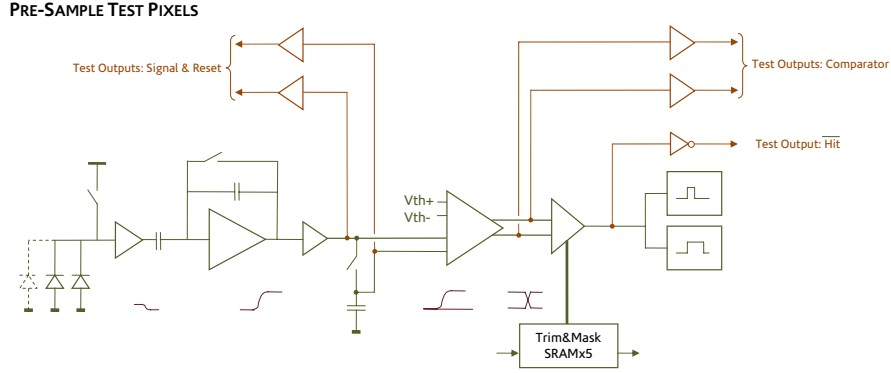

Figure 16: *Pre-sample test pixel circuit, showing the internal nodes which are accessible externally.*

Key digital signals, such as control clocks and the least significant bit of the multiplex address and time-code are driven off-sensor at the farthest point from their initial distribution. This debug feature allows the timing of critical signals to be evaluated during operation.

All bias currents are independent and are generated off-sensor to evaluate the performance of sub-circuits in different operating modes. The two pixel architectures can be operated independently, with separate threshold voltage, bias settings and power-down control.

3.11 DAQ system and operation

For all the tests described in this paper, the sensor was mounted on a custom-designed PCB and read out with a USB-based readout system using a custom software data acquisition (DAQ) system.

The PCB design included 31 DACs to allow all the bias voltages and currents for the sensor to be software controlled. It also held a temperature sensor. Each PCB was given a unique ID through a dip switch which could be read out to the DAQ. Below each sensor, a hole approximately $6 \times 6 \, \text{mm}^2$ was cut in the PCB to allow access to the substrate for use with a laser (see below). An additional notch approximately $1 \times 1 \, \text{mm}^2$ was also cut below the test pixel structures. The PCB hole left a strip approximately $1 \, \text{mm}$ wide for gluing the sensor to the PCB. Of the O(20) sensors mounted on PCBs, no mechanical failures of this glue joint were found. Conductive glue was used to allow the substrate to be optionally grounded. The sensor required 265 wirebonds to connect to the PCB. The PCB is shown in figure 17.

Three 64-way flat ribbon cables connected the sensor PCB to a generic readout board, the USB_DAQ card, which provided the USB interface to the computer. The USB_DAQ card required firmware specific to the TPAC1.0 sensor to generate the required control and timing signals, as well as the configuration load and readback, and the data memory readout. It also controlled the PCB DAC settings and performed the temperature and PCB ID readout.

The sensor was operated with pseudo-bunch trains to mimic ILC operation. The timing structure of the bunch train was controlled by the USB_DAQ and was to a large extent under software control. For all following measurements, the operating parameters used were to have $400 \, \text{ns}$ between bunch crossings and a bunch train up to the maximum timestamp storable in the memory of 8192 bunch crossings.
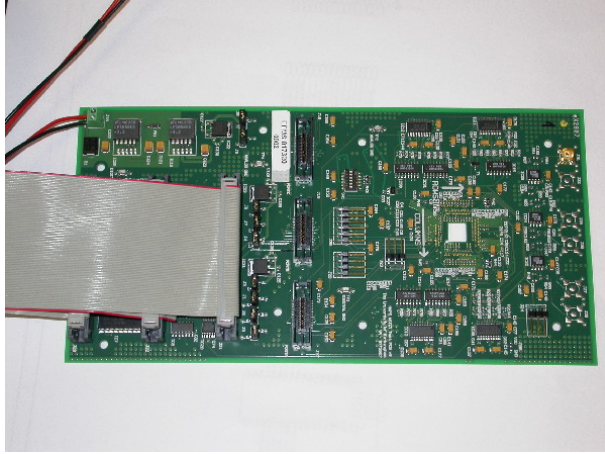
Figure 17: *Photograph of the PCB used to hold and operate the sensor. The hole for the sensor is just right of centre.*

4   Single pixel performance

4.1   Charge spread simulation

The signal from a minimum ionising particle (MIP) arises as the particle passes through the epitaxial layer of the sensor. Physically, the electron-hole pairs are created throughout the silicon sensor, but only within (or near) the epitaxial layer are the liberated electrons able to diffuse to the collecting diodes and hence be seen as signal. Some electrons from the substrate close to the epitaxial layer can also diffuse into the epitaxial layer and so be collected. Conversely, electrons in the deep P-well implant will tend to be lost. Overall, in a $12\,\mu$m thick epitaxial layer, then for a MIP at normal incidence to the sensor, around $1200\,e^-$ will be available to contribute to the signal. Once liberated, the charge will diffuse and so may be absorbed by the collection diodes of the neighbouring pixel. In addition, other N-well structures can collect the charge and hence charge will be lost in terms of being signal. This last effect is the reason for which the deep P-well implant was developed.

To estimate these effects quantitatively, a simulation of the sensor pixel was performed using the Sentaurus [13] package. The pixel GDS design file was used as input to the simulation, ensuring the details of the pixel were correctly simulated. Charge equivalent to that expected for a MIP was deposited at various locations within the pixel and then simulated to investigate where it was absorbed or lost. A 3×3 array of pixels was simulated, to allow for neighbour pixel charge collection, from the sensor surface down to $32\,\mu$m below ($150 \times 150 \times 32\,\mu$m^3 in total). The simulation was time-consuming so a limited subset of positions were simulated. Because of the approximate eight-fold symmetry of the sensor, all positions simulated were within a triangle with the pixel centre, corner, and side centre as its three corners. Within this triangle, the positions simulated were spaced $5\,\mu$m apart, resulting in 21 points in total. Using the pixel symmetry, these could be translated and rotated to give a $5\,\mu$m regular array across the whole of the central pixel. Figure 18 shows the layout of the 21 points simulated.

The simulation results are shown in figure 19 for sensors with and without the deep P-well implant. Here the fraction of the signal charge predicted in all nine pixels is shown. It is seen that the signal in all cases is predicted to be significantly larger with the deep P-well implant. In particular, in the central pixel, the fraction of charge seen as signal varies between 20% and 50% with the deep P-well implant, whereas without this, the signal can be less than 1% and only reaches a maximum of 30% when near the diode. Similar large differences are seen for the
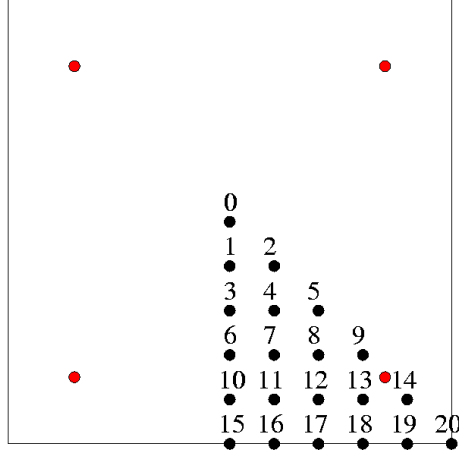
Figure 18: *Layout and numbering of the 21 points (black) used for simulation. The red circles indicate the four diode positions.*
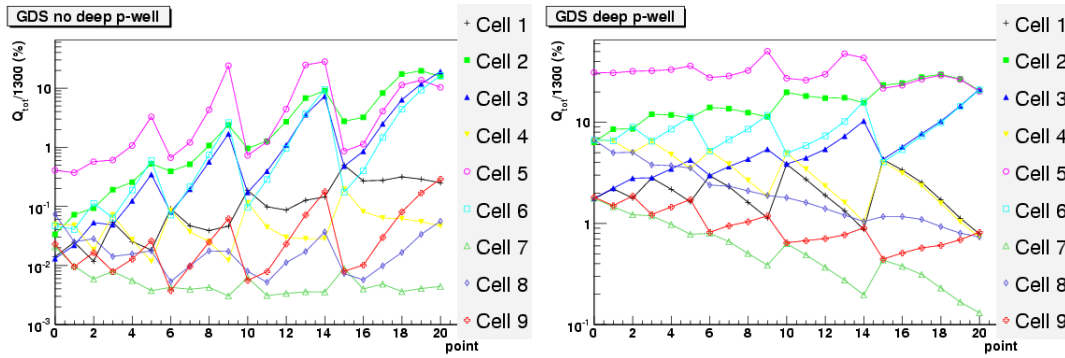
neighbouring pixels also.

Figure 19: *Simulation of the fraction of the charge seen as signal for (left) no deep P-well and (right) deep P-well. The colour code for each of the nine pixels ("cells") is shown on the plots, where Cell 5 is the central pixel.*

4.2 Test pixel performance

The sensor included three test pixels outside of the 168×168 "bulk" pixel array. Several internal nodes of two of these pixels were accessible externally, allowing analogue measurements of their levels.

A low noise pre-amplifier was connected to the output node from the charge preamplifier (the "Test Output Signal" shown in figure 16) and the signals were measured on an oscilloscope. Allowing for a gain of 0.8 in the external output buffers, the RMS of the signal gives a direct measure of the noise on the signal input line to the comparator. The raw RMS value was found to be $3.5 \, \text{mV}$, implying the noise at the comparator input was $4.4 \, \text{mV}$.

To get an absolute calibration of the gain and hence of this noise value, the sensor was exposed to an intense ^{55}Fe source. This isotope produces gamma radiation with a photon energy of $5.89 \, \text{keV}$. Photons of this energy can deposit all the energy within a small $O(1 \, \mu\text{m}^3)$ volume within the silicon. As shown above, in most cases, it would be expected that a significant fraction

of the charge would diffuse out of the test pixels or be absorbed. However, a small number of the γ particles will interact directly inside the collection diodes. In this case, effectively all the charge liberated, around $1610\,e^-$, will be collected by the diode, giving a well-defined signal size. The signal observed is shown in figure 20.
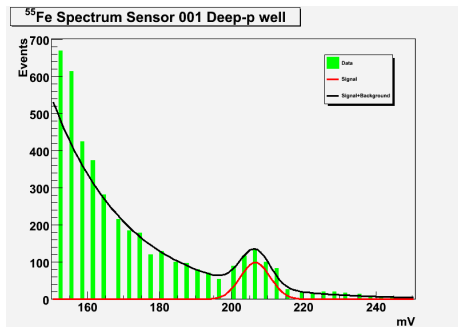

Figure 20: *Test pixel:* ^{55}Fe *signal.*

The signal peak is at $207\,\text{mV}$ and this includes the gain of 0.8 from the test pixel external buffers not in the standard pixel signal path. Hence, this determines the test pixel charge preamplifier gain to be $160\,\mu\text{V}/\text{e}^-$. This also allows the $4.4\,\text{mV}$ noise discussed above to converted to an ENC of $27\,e^-$. This is the noise on one leg of the comparator input. The total noise when operating the pixel is larger by a factor of $\sqrt{2}$ (see Section 3.4) so this would imply a total noise level of $38\,e^-$. This is higher than the predicted $22\,e^-$ for the pre-sampler design (see Section 3.4); the difference is not yet understood.

4.3 Bulk pedestals and noise

The performance in the bulk pixels is less straightforward to measure as there is no analogue readout available. However, the rate of hits in each pixel as a function of the applied threshold effectively allows the integral of the analogue spectrum to be measured and hence the spectrum itself can be estimated. The results in this section are based on performing such "threshold scans" and measuring the response.

One further complication is that pickup was observed between pixels when O(100) were enabled and firing at the same time (see Section 5). This is not a major problem when operating the sensor for its design purpose as only O(10) pixels per sensor are expected to fire in each event. However, for the basic performance measurements presented below, this pickup would prevent the response of individual pixels from being measured. Hence to remove this sensor-wide effect, only small number of pixels are unmasked at any time for the following results.

Specifically, only one pixel of the 42 which shared SRAM memory was unmasked in each run. In addition, runs were performed with only 19 bunch crossings per bunch train. Both of these together ensured that the memory could never overflow and so distort the threshold scan, even when very close to threshold.

The pixel comparator only fires when the input crosses through the threshold from below. Hence, a threshold scan performed with no external stimulus will result in hits only when near the pixel pedestal. The results of such a scan for typical pixels of each type are shown in figure 21. Here, the scan is specified in the (arbitrary) units of the DAC threshold setting, called "threshold units" (TU) in this paper.

It is seen that the pre-sample pixel gives a distribution which is not well-behaved and is rather asymmetric. In addition, it was found that other properties of the pre-sample pixels (such as the rate of noise hits being strongly dependent on the time since the reset) made them
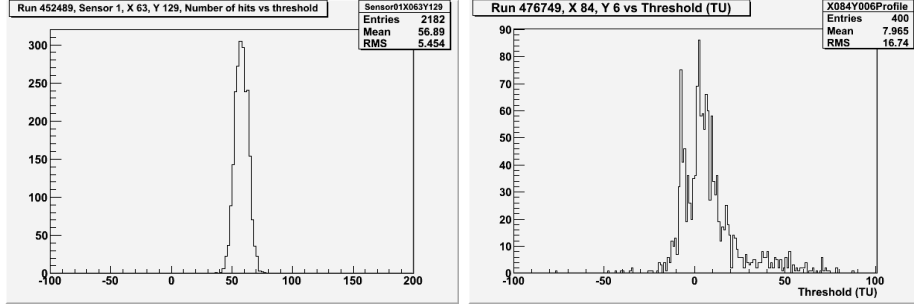
Figure 21: *Typical pixel threshold scans for (left) a pre-shape pixel and (right) a pre-sample pixel.*

significantly harder to use and interpret. In the rest of this paper, results are shown only for pre-shape pixels. Only pre-shape pixels were implemented in the next version of the sensor.

For the pre-shape threshold scan, the resulting distribution is well described by a Gaussian and so can be characterised by the mean and RMS. The mean gives the pedestal for the pixel and the RMS gives the noise [12]. The distribution of the pixel mean and RMS values for the two pre-shaper quadrants of a typical sensor are shown in figure 22. It is seen that there is a significant spread of pedestals within a quadrant. The pedestal spread RMS is around 20 TU and is the same for quadrants 0 and 1. This is much bigger than the noise on the pixels, which has an average value of around 5 TU for quadrant 0 and 6 TU for quadrant 1.
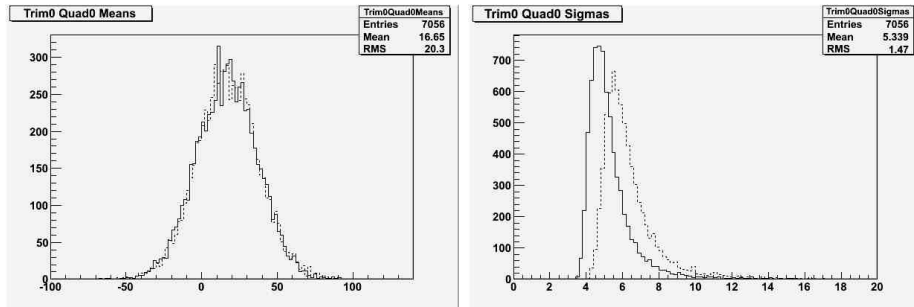

Figure 22: *Distributions of (left) pedestals and (right) noise for quadrants 0 (solid histogram) and 1 (dashed histogram), in threshold units, for a typical sensor.*

The values of the pedestal and noise for each pre-shape pixel as a function of the position within the sensor are shown in figure 23, where the difference in noise for the two quadrants is visible. The values of the pedestals and noise show no correlation with position within the sensor and their spreads appear genuinely random pixel-to-pixel.

The spread of pedestals can be reduced using the per-pixel trim setting. This allows a value in the range 0-15 to be loaded to each pixel to adjust the pedestal position upwards. The effect of this on the pedestal of a typical pixel is shown in figure 24.

The spreads of pedestals before and after trimming a typical sensor are shown in figure 25. It is seen that the RMS of the pedestals is reduced by a factor of around five. The resulting spread is a little smaller than the size of the per-pixel noise and so gives a small but non-negligible contribution to the apparent noise rate when setting a common threshold for all pixels. The ability to improve on this effect is limited by the number of trim bits available. With four bits, then a granularity of the trim of 16 values is possible. However, the total range of the
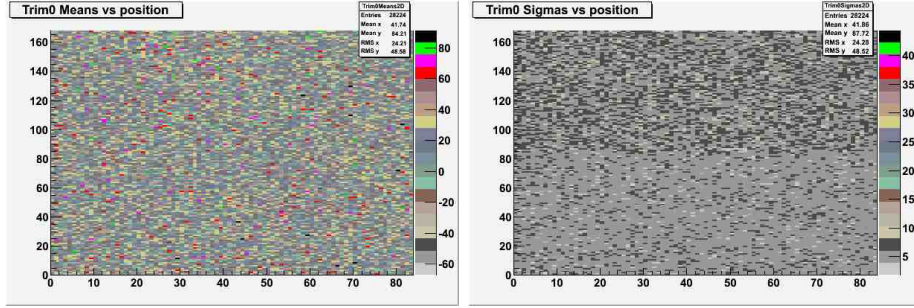
Figure 23: *Two-dimensional map of (left) pedestals and (right) noise for the two pre-shape quadrants of a typical sensor. The lower half of each plot corresponds to quadrant 0 and the upper half to quadrant 1.*
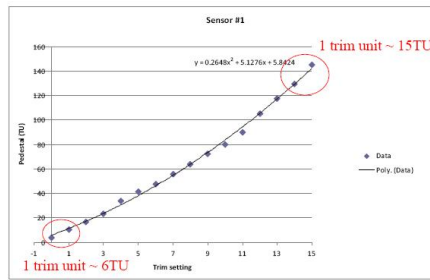

Figure 24: *Change of pedestal for a typical channel as a function of the trim setting.*

uncorrected pedestals is around 100 TU, so that even with careful matching of the trim range, the trim LSB would be around 6 TU. Two additional trim bits have been implemented in the next version of the sensor.
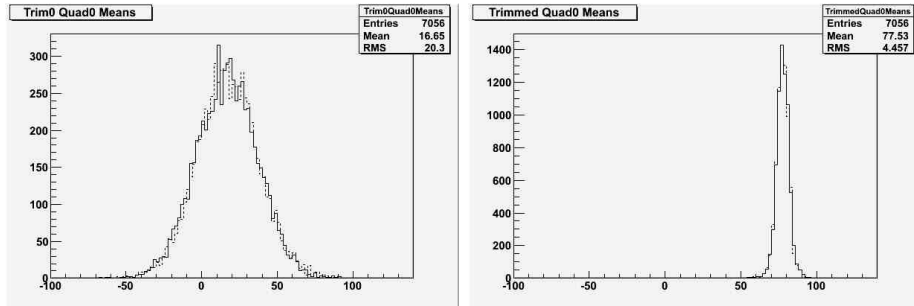

Figure 25: *Pedestals in quadrants 0 (solid line) and 1 (dashed line) for (left) a trim setting of 0 for all pixels and (right) a optimised per-pixel trim setting, for a typical sensor. The left plot is the one shown in figure 22.*

In principle, the pixel noise can be influenced by several factors. The sensor substrate can be grounded or operated with no explicit bias connection to the substrate. This could in principle change any environmental noise pickup. The distribution of noise for a subset of pixels both with and without the substrate being grounded are shown in figure 26. It is seen that leaving the substrate unconnected has no significant effect on the noise. In addition, the sensor may be sensitive to light and it is normally operated with a cover to keep it in the dark. However,

the six metal layers used effectively completely cover the top surface of the sensor so little light reaches the epitaxial layer. Figure 26 also shows the noise distribution of the same pixels with and without a strong lamp being shone on the top surface of the sensor. Again, there is no significant difference.
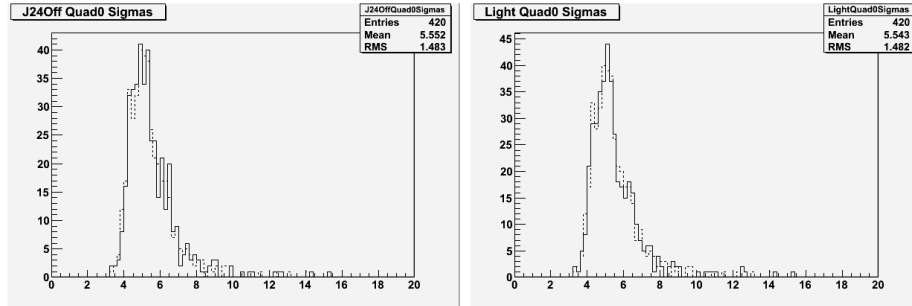

Figure 26: *(Left) Noise distribution with (solid histogram) and without (dashed histogram) grounding the substrate. (Right) Noise distribution with (solid histogram) and without (dashed histogram) a strong light source shining on the sensor.*

4.4 Bulk pixel calibration

The calibration of the threshold setting in TU to a physical scale was again performed using an ^{55}Fe source using the same principle as described in Section 4.2. A very powerful source, O(GBq), allowed a sufficient rate of source hits that a calibration signal could be seen in individual pixels. Again a threshold scan was performed and the rate plotted as a function of the threshold setting. A numerical derivative was then taken of this rate plot to get the basic spectrum and so to find the calibration peak. The results for a typical pixel are shown in figure 27.
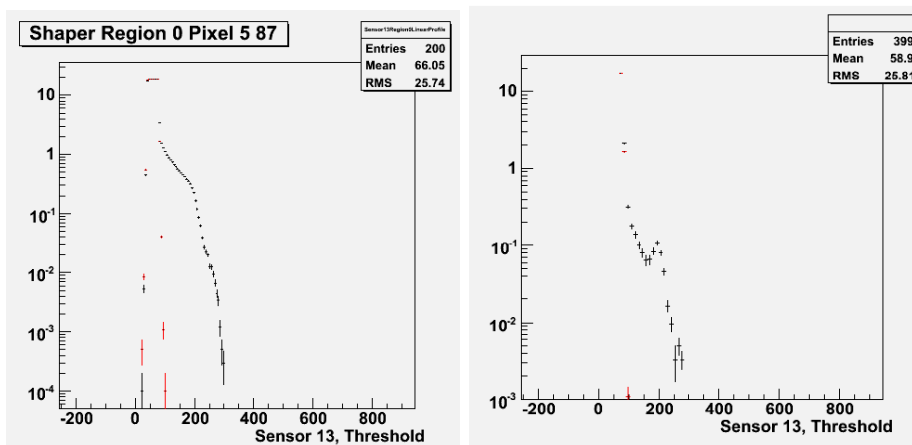

Figure 27: *(Left) Rate of hits as a function of the threshold setting. (Right) Numerical derivative of the rate plot. The ^{55}Fe calibration signal peak is visible at around 200 TU.*

The average peak position relative to the pedestal was found to be approximately 200 TU. This corresponds to 5.89 keV of deposited energy and to $1610\,e^{-}$ of signal charge. This can be used to interpret the noise measurements above. Hence, the average noise value of 6TU corresponds to an equivalent in terms of deposited energy of 180 eV, or in terms of signal charge

23

of $50\,e^-$. The expected value of the noise for the pre-shape pixels was $23\,e^-$ and so the measured noise is again significantly above the predicted value.

4.5 Bulk laser signal response

The sensor was illuminated using a laser of wavelength $1064\,\mathrm{nm}$. The laser enclosure contained a computer-controlled moveable XY stage, on which the sensor PCB could be mounted. This allowed the sensor to be moved in two dimensions below the laser, allowing a large number of measurements which are described in this section. For these measurements, the signal was injected into the bulk pixel area and the response measured using a threshold scan. A typical response is shown in figure 28, where the laser signal is clearly seen. The actual size of the laser signal was found by fitting the falling edge of the response curve to an integrated Gaussian (erf) function. In terms of the integrated Gaussian parameters, this gives a pseudo-analogue signal response estimate from the mean and an estimate of the signal spread from the width. The absolute intensity of the laser was uncalibrated so the signal size for the following measurements is arbitrary.
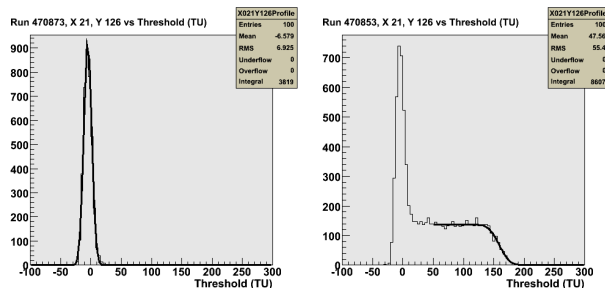

Figure 28: *Typical threshold scan (left) without and (right) with the laser. The laser plot shows the fit to the erf function and the signal was determined from the fit to be 160 TU in this example.*

The laser must be focussed onto the epitaxial layer. This was done by focussing the microscope optically on the substrate surface and then recording the observed laser signal size at various focus depths relative to this. Figure 29 shows the results of such a scan, where there is seen to be a strong dependence on the focus depth. The maximum signal was found to be $60\,\mu\mathrm{m}$ into the substrate relative to the optical focus. In addition, it can be seen that there is a range of only around $20\,\mu\mathrm{m}$ where the signal size is effectively unchanged. This sets a limit on the degree of tilt acceptable when performing such laser measurements. A focus on one side of the sensor will only be within this tolerance at the other side if the sensor is flat to around $2\,\mathrm{mrad}$. For this reason, the focus was checked when large movements were made and, when possible, only small movements were then performed around this point.

The laser system was used to find the working point of the sensor in terms of the external bias voltages and currents. These are set using DACs on the sensor PCB; see Section 3.11. Of the 31 DACs used, 7 are used to set the threshold and monostables, 4 are common to pre-shape and pre-sample pixels, 8 are only used by pre-shape pixels and the remaining 12 are only used by pre-sample pixels. Hence, for pre-shape pixels, there are 12 DAC settings which determine the working point. For each DAC in turn, the value was scanned through a reasonable range and the noise and laser signal size determined for each value. Figure 30 shows the results of these DAC scans in terms of the laser signal divided by the noise value. In almost all cases, a broad maximum was found and the values were set to be within this maximum region. These standard values were used for all the measurements, including those discussed previously in this
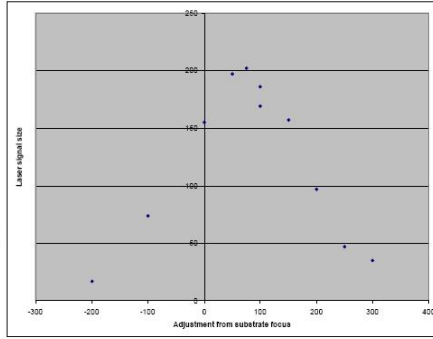
Figure 29: *Laser signal size as a function of the focus depth relative to the optical focus position. A positive value means the focus position was moved into the sensor.*

section.

Figure 30: *Laser signal/noise as a function of each of the 12 DAC settings relevant to the pre-shape working point. The DAC values used for standard operation are shown by the arrows in each case.*

The laser system was also used to determine the relative gain of the pixels. By firing the laser into the centre of each pixel in turn, then the signal size for each was determined. Figure 31 shows the distribution of the observed laser signals for a subset of pixels in the two pre-shape quadrants. A clear difference in gain for quadrants 0 and 1 was observed, with quadrant 1 being around 40% higher than quadrant 0. As shown above in figure 22, quadrant 1 also has a higher noise than quadrant 0, by around 20%, so quadrant 1 has the better signal/noise ratio. This is shown explicitly in figure 31. For this reason, only the quadrant 1 variant of the pre-shape pixels was implemented in the next version of the sensor.

The laser system was also used to measure the charge diffusion so as to compare with the simulation described in Section 4.1. The laser was focussed to give a illuminated spot sixe of less than $2\,\mu$m and the stage could be moved with a $1\,\mu$m precision, so that the charge could be generated accurately to match the positions of the 21 points used in the simulation. Because the laser was uncalibrated, the absolute fraction of the laser signal observed could not be determined, but a single scale factor was used to give reasonable agreement to the simulation expectation for the central pixel. Figure 32 shows the results of these measurements using sensors with and without deep P-well implants. It is seen that the general trend of the data is well reproduced by the simulation (shown in figure 19) with the deep P-well sensor signal varying between 20% and 50%. However, the sensor without deep P-well implants gives somewhat higher signals than the simulation predicts, being always above 1%, although the general shape is similar. Overall, it is clear that the sensor without deep P-well implants would give a much smaller signal.
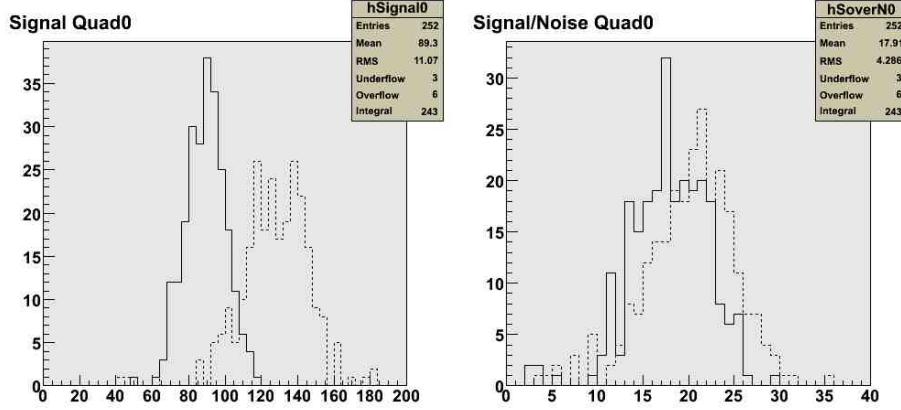
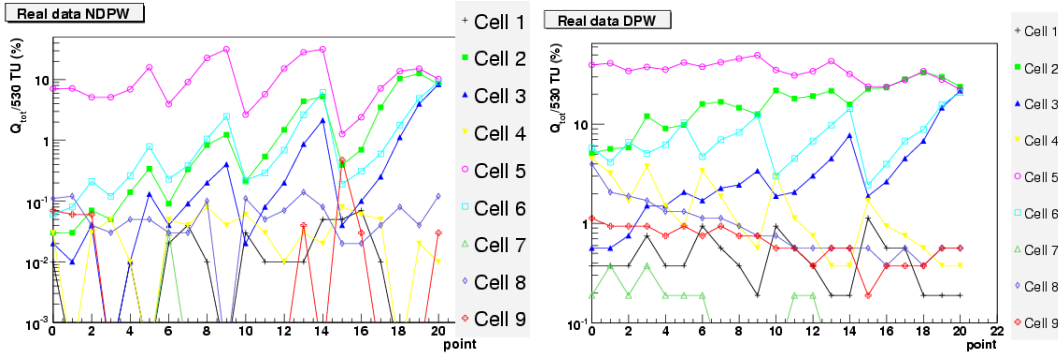Figure 31: *Laser (left) signal and (right) signal/noise for quadrants 0 (solid line) and 1 (dashed line).*

Figure 32: *Charge seen at each of the 21 point positions for each pixel (left) without deep P-well and (right) with deep P-well.*

5 Sensor performance

5.1 Configuration load

The per-pixel five-bit configuration data were loaded and read back to check for bit errors. Each load transfered 141 kbits into the sensor configuration registers. The load and readback cycle was repeated 25,000 times and the data written and readback were compared. No bits were seen to have been corrupted in this process. This sets an upper limit on the bit error rate of 1×10^{-9} at 90% confidence level.

5.2 Pickup between pixels

As mentioned in Section 4.3, pickup was observed between pixels when large numbers fired at the same time. For this reason, all the results described in Section 4 were done with only the pixel(s) being studied being unmasked. Typical pickup behaviour is illustrated in figure 33 where the threshold scan for a single pixel is shown with just that pixel unmasked and also with all pixels unmasked. It is clear the distribution is uncorrelated with the actual pedestal and depends on the pedestals of the other pixels firing. The scan with pickup has a much larger RMS and this has been used to characterise the onset of pickup.
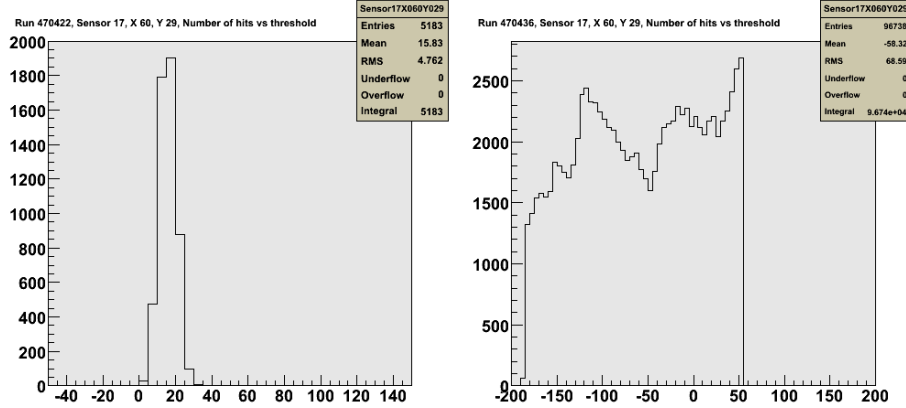
Figure 33: *Effect of pickup on a typical pixel with (left) the single pixel enabled and (right) with all pixels enabled. Note the difference in the x axis scale.*

Many studies to analyse this effect were performed. It was found that the level of pickup only depended on the number of pixels unmasked but not their geometric positions relative to each other. One example of these studies which shows this result clearly is shown in figure 34. In this study, a single pixel on one edge of the sensor was unmasked. The RMS of this pixel from a threshold scan was then determined as an increasing number of pixels were unmasked along the other edge of the sensor, i.e. at the furthest point from the single pixel. It is seen that pickup on isolated pixel has the same behaviour as that of all the pixels along the far edge. Hence, the pickup is seen to be unrelated to position.
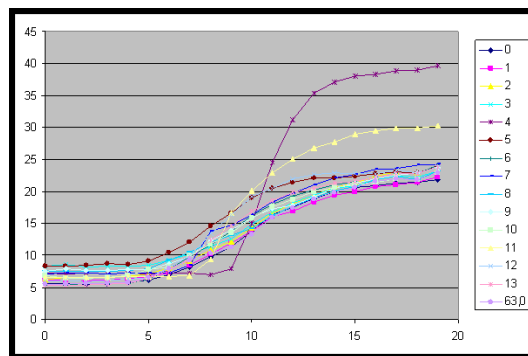

Figure 34: *Onset of pickup, as measured by the RMS determined from a threshold scan, as a function of the number of unmasked channels. The x axis corresponds to the number of rows of 42 pixels which were unmasked, in addition to a single edge row which was always active. The line labelled "63,0" corresponds to an isolated pixel on one edge of the sensor, while all the other lines correspond to a subset of pixels in the row along the opposite edge. It is seen that pickup starts around an x axis value of 6, which corresponds to 294 pixels unmasked.*

It is thought that the pickup is caused by a power mesh being shared between the pre-shape comparator and monostable components. It is clear it is not coupling directly between monostables as this would not be affected by masking pixels. If a large number of pixels fire and so give hits in the monostables, the common power level may droop and can cause the comparators in other pixels to fire. The power meshes to these two components have been separated in the next version of the sensor.

5.3 Hit corruption

The "hit override" mode described in Section 3.8 allows hit data to be stored in the sensor memory without requiring any pixels to be above threshold. In this mode, all seven banks of pixels in each row of 42 pixels will store data for every timestamp and hence the 19 data registers will be filled within the first three timestamps.

This test mode results in predictable values for the MUX address, timestamp and row address for every data word stored. These can be compared with the values read out to check for corruption. (Note, the six-bit hit data themselves are not predictable as noise can cause them to be randomly set.) A low level of corruption, at around 0.1%, was seen for the MUX address and timestamp bits, although the actual rate varied from sensor to sensor. This was traced to a design fault; the signal to write to the SRAM did not have a wide enough voltage range to guarantee the value would be written correctly. The rate of corruption was reduced by applying an external voltage level of 2.7 V to increase the range and this was used for all the results shown in this paper. This design fault has been corrected in the next version of the sensor.

$$TODO : ADDTWINNING$$

5.4 Beam test

Four sensors were exposed to an electron beam at DESY in December 2007. This was relatively early in the sensor testing and the sensors had not yet been fully understood. In particular, the pedestal measurement and trim setting were not yet in use so all sensors ran untrimmed. This required a much higher threshold than would normally be used, due to the spread of pedestals as shown in figure 22, and hence a low efficiency.

The sensors were mainly run in a "tracker" mode, with no tungsten converter plates as would be required for a calorimeter. This allowed charged tracks to be reconstructed using the hits observed in each sensor. By connecting hits in the outer two sensors, then the hits in the inner two sensors could be compared with this interpolated track. Figure 35 shows the distributions of the difference between the observed hit and the interpolation for these two sensors, following alignment. The residual widths are consistent with that expected given the $50\,\mu$m pixel pitch. It is clear that hits due to charged particles were observed.
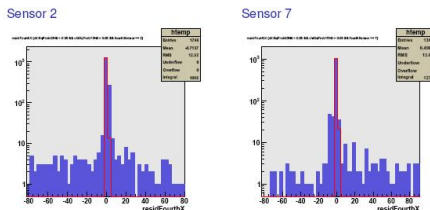

Figure 35: *Distribution of residuals for inner sensor hits compared to outer sensor interpolated tracks. Note the y axis log scale.*

Following the beam test, the pedestals for each pixel were measured and so the thresholds set in the beam test data was corrected pixel-by-pixel to give the actual thresholds used for each pixel. These actual thresholds had a wide spread due to the uncorrected spread of the pedestals. This in principle allows the hit efficiency to be found over this wide threshold range. However, in practise, the overall low effiency meant that there were insufficient track statistics available to extract an absolute efficiency per pixel. Instead, a relative efficiency could be found by comparing hits in neighbouring pixels for spatial and time correlations. The absolute normalisation is not known but assuming uniform exposure of all pixels (which was true to a good approximation

due to the small size of the sensor) the relative rate in each pixel gives a relative efficiency measurement. Figure 36 shows the results of this analysis. A strong fall-off of efficiency with threshold is observed.
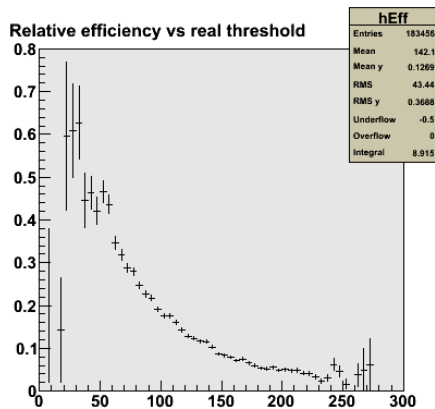

Figure 36: *Relative efficiency per pixel as a function of the actual threshold for the pixel in TU.*

5.5 Cosmics

$$TODO : WAITINGFORSTATISTICS$$

6 Physics performance

Given the results shown in the previous sections, it is possible to simulate the performance of a full scale binary ECAL. This study was done with the same sampling geometry as discussed in Section 2. These results supercede those presented previously [14].

6.1 Simulation

The underlying EM shower simulation used was based on the GEANT4 programme [8]. This simulation has been used extensively and in particular has been checked to high accuracy in an analogue silicon-tungsten sampling ECAL by the CALICE collaboration [15]. However, this ECAL had a pad granularity of $1 \times 1 \, cm^2$ and it is not known if the shower simulation is accurate at the much smaller granularities relevant for the binary ECAL simulation.

The energy deposited within each pixel was recorded on a grid of $5 \times 5 \, \mu m^2$ subpixels. These corresponded to squares formed with the 21 simulation points described in Section 4.1 at the corners. Within each subpixel, the energy deposited by all the particles in the event was summed. The simulation used units of the deposited energy throughout, so the plots below are in terms of these units.

Following the completion of the shower simulation, then this energy was shared between the hit pixel and the eight surrounding nearest neighbour pixels according to fractions determined from the charge diffusion simulation. The energy per pixel was then smeared by applying Gaussian noise and a threshold was applied to each pixel. For pixels with no energy (whether from the original shower or due to diffusion), then hits were added to random pixels with a rate corresponding to the noise and threshold values chosen. The effect of dead area was included by removing any hits in the inactive areas following the simulation process. The spectra of energies in each pixel at each stage of this process are shown in figure 37.
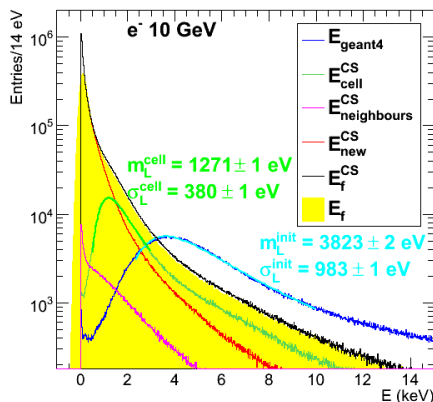
Figure 37: *Spectra of energy in each pixel at various stages of the sensor simulation. The spectra correspond to the original deposited energy (dark blue), the energy in the hit pixel after charge diffusion (green), noise in neighbouring pixels which were not hit (pink), the energy in the neighbouring pixels due to charge diffusion (red), and the total (black). The yellow histogram corresponds to the total smeared by the noise.*

A nominal noise value of $120\,\text{eV}$ was used; this corresponds to the low end of the noise distribution observed but is a realistic target for a future sensor. For a single MIP at normal incidence, the resulting efficiency for generating at least one hit when passing through a layer is shown in figure 38 as a function of the threshold applied. This should be compared with figure 36, for which the calibration of TU means the range of the threshold axis shown in that figure corresponds to $9\,\text{keV}$, twice the range of figure 38. Both show a rise of around a factor of five in the range $1\,\text{keV}$ to $4\,\text{keV}$, so giving some confidence that the simulation is reasonably reproducing the sensor performance.
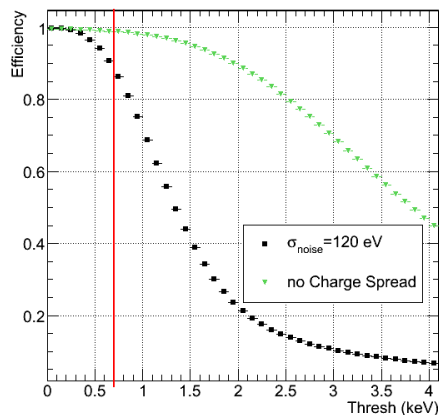

Figure 38: *Simulated of the efficiency for obtaining at least one hit from a MIP at normal incidence, as a function of the threshold. The squares correspond to the full simulation and the inverted triangle to a simulation with no charge diffusion. The vertical line indicates a typical threshold which would be used.*

6.2 MIP counting

Due to the charge diffusion between neighbouring pixels, it is possible for a single simulated MIP passing through a layer to result in anything from zero to nine hit pixels, depending on position, noise level and threshold. Hence, simply using the number of hit pixels as an estimate of the number of MIPs incurs a large error from this fluctuation. By clustering the neighbouring pixels and counting clusters rather than hits, it is possible to remove most of the effect of such fluctuations and so get a better estimate of the number of MIPs. This clustering is called "MIP counting" here.

A simple MIP counting algorithm was used for the following results. The pixels were grouped into isolated clusters, which were defined to be contiguous clusters of hit pixels with no further hits in any of the nearest neighbouring pixels. If the cluster contained no hit pixels for which all eight nearest neighbours also were hit, then the cluster was counted as one MIP. If one or more hit pixels had all eight nearest neighbours also hit, then the cluster was counted as having a number of MIPS equal to the number of such completely surrounded pixels. Clearly, this algorithm is not unique and is unlikely to be the best possible for this purpose. However, the optimisation depends on details of the EM shower behaviour at very fine granularity. These details have yet to be verified with real data and so finding the best MIP counting method cannot be done with any degree of realism at this stage.

6.3 Electromagnetic shower resolution

Using the MIP counting algorithm decribed above, the resulting EM shower resolution predicted by the simulation is shown in figure 39. It is seen that the resolution of the binary ECAL modelled here is degraded significantly with respect to the ideal digital case of perfect MIP counting. The stochastic term is increased from 9.5% to 13.0%, around 35% higher. However, note the binary ECAL is still predicted to have a better resolution than the ideal analogue ECAL which uses the same sampling fraction.
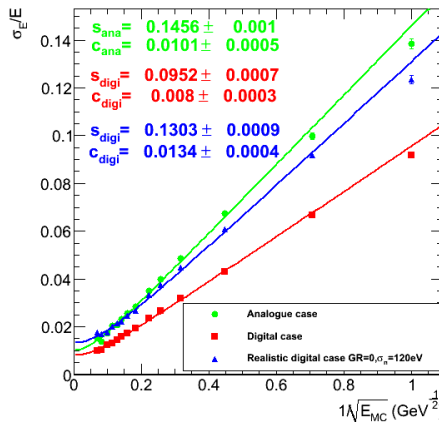

Figure 39: *Comparison of the resolution dependence on the incident photon energy. The red and green lines are the ideal cases, as shown in figure 3. The blue line shows the result for a realistic binary ECAL after the full sensor simulation described in the text, including MIP counting.*

The various effects included in the simulation can be varied to observe their individual contributions to this degradation. For the following studies, an incident photon energy of $10\,\mathrm{GeV}$ was used, for which the value of σ_E/E is 3.2% for the ideal binary ECAL and 4.5% with the full simulation.

The effect of the deep P-well implant on the sensor is shown in figure 40. It is seen that for a sensor with no deep P-well, the loss of charge to the unprotected N-well circuit diodes causes a very significant degradation of the resolution from 4.5% to 8.1%, more than doubling the ideal value and making the performance much worse than the ideal analogue ECAL.
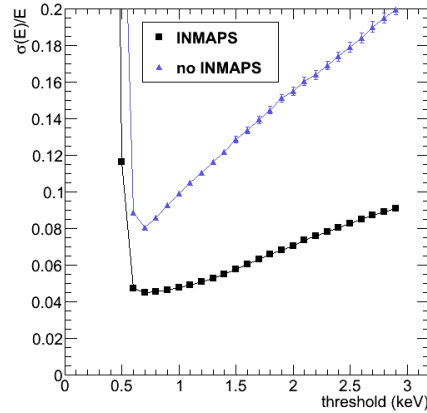

Figure 40: *Resolution for 10 GeV photons with (squares) and without (triangles) the deep P-well implant, as a function of the applied threshold.*

The effect of the noise is shown in figure 41. Noise affects the resolution through adding
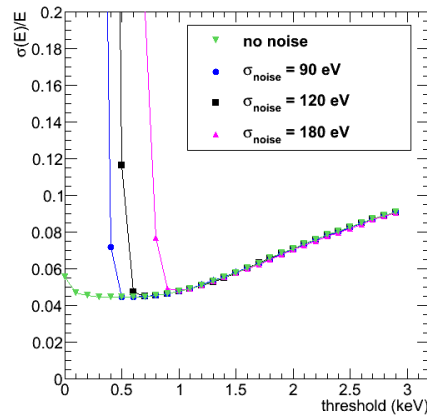

Figure 41: *Resolution for 10 GeV photons with simulated noise equivalent to 90 eV (cirles), 120 eV (squares) and 180 eV (triangles) energy deposited, as a function of the applied threshold. These correspond to the expected, best observed and average values of the sensor, respectively. Also shown is the resolution in the absence of noise (inverted triangles).*

isolated hits which are counted as MIPs. While a correction can be made for the average number, these fake MIPs increase the resolution error due to the statistical fluctuation on the number of noise hits. The number of hits depends very strongly on threshold but, for high enough thresholds, the effect of the noise can be made small. Since the threshold can be chosen at will given a particular noise level, then by adjusting the threshold, the increase in resolution due to the noise can be reduced. The figure shows that higher noise gives an optimal resolution at a higher thresholds but that the resulting resolution degrades only slightly compared to the case with no noise. Note, the wide minimum for the case with no noise; 90 eV noise gives the same

resolution as no noise, so indicating the level as which the noise becomes negligible. The optimal thresholds for the three noise values shown were found to be 0.5 keV for 90 eV noise, 0.7 keV for 120 eV noise and 1.0 keV for 180 eV noise. The resulting resolutions were 4.4%, 4.5% and 4.8%, respectively. Hence, the noise contribution to the degradation from the ideal resolution of 3.2% is small.

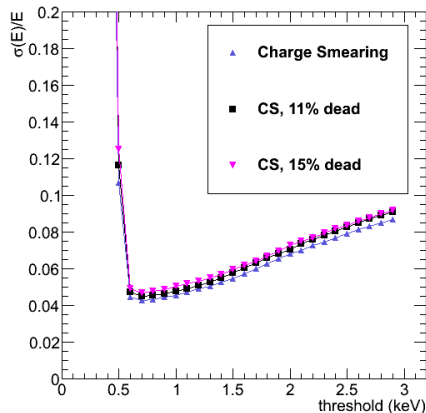The effect of the dead area is shown in figure 42. It is seen that the effect of the SRAM

Figure 42: *Resolution for 10 GeV photons with dead area equivalent to 11% (squares) and 15% (inverted triangles), as a function of the applied threshold. These correspond to the on-sensor memory area and an estimate of the total dead area in a full-scale ECAL including edge effects around the sensor. Also shown is the resolution for a sensor with no inactive areas (triangles).*

memory dead area is to increase the resolution from 4.3% to 4.5%, again a small effect. By design, the dead area due to the on-sensor memory is spread out across the sensor. The EM showers have a width determined by the Molière radius of 9 mm and the memory dead areas are spaced out by 2.4 mm. This means every shower loses approximately the same fraction of hits and so the resolution degrades by only the statistical fluctuation on the number, not due to a variation in the mean number as a function of the impact position of the incident particle. Although somewhat uncertain, an estimate of the effect of extra dead area around the periphery of the sensor, and in gaps between sensors, for a realistic-sized sensor has been made. This may contribute a further 4% of dead area. The effect of this has also been modelled and is shown in figure 42 to give a very small contribution.

The effect of the charge diffusion is shown in figure 43. Although it is not physically possible to turn off the charge diffusion in the sensor, it is interesting to see how the charge diffusion affects the resolution. The figure shows the the resolution with and without simulating charge diffusion. The amount of charge per hit pixel is around three or four times higher on average without charge diffusion and so the effect of the threshold cut in the range considered is much smaller. However, at the optimal threshold, the resolutions with and without charge diffusion are very similar at 4.3% and 4.5%. Hence, the crosstalk from charge diffusion has a small effect for a binary ECAL application.

The effect of the MIP counting is shown in figure 44. The figure shows that using a naive hit count to estimate the number of MIPs gives a very large resolution, namely 5.9% compared with the 4.5% using a MIP counting algorithm, so using such an algorithm is essential to get good performance. However, the algorithm used does not bring the resolution down close to the ideal case value of 3.2%; there is still a significant difference between the ideal and realistic cases which is bigger than any of the effects previously considered.

Hence, the dominant contribution to the degradation of the resolution, even in the absence
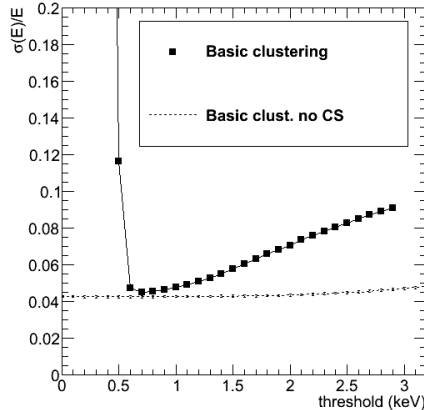
Figure 43: *Resolution for 10 GeV photons with (squares) and without (dashed line) charge diffusion, as a function of the applied threshold.*
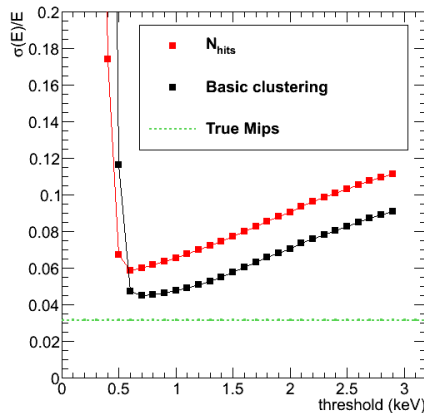

Figure 44: *Resolution for 10 GeV photons with no MIP counting (red) and with the simple MIP counting algorithm (black) as described in the text, as a function of the applied threshold. The horizontal line (green) shows the resolution obtained with an ideal MIP count.*

of noise, dead area and charge diffusion, is due to the confusion of interpreting the observed hits in terms of numbers of MIPs. It is clear that a major study of MIP counting algorithms will be needed to find the limit of this approach. However, as stated above, the details of such an optimising MIP counting algorithm are likely to depend heavily on the details of EM showers. In particular, they will be sensitive to the core hit density in each layer and potentially to correlations between layers. As stated above, the GEANT4 shower simulation has not been verified at the very fine granularities relevant for these studies. Hence, it is essential to get real experimental data on showers at these granularities in order to know the real limits of the binary ECAL approach.

6.4 Hadronic jet resolution

The current state-of-the-art in PFA performance is exemplified by the PandoraPFA programme [4]. This programme implements a very detailed algorithm which has been developed using simulations of analogue ECALs with pad sizes of 4-10 mm. To reoptimise the whole algorithm to take best advantage of the small pixel size of a binary ECAL would be a major task and is beyond

the scope of this paper. Hence, the ultimate performance of a binary ECAL for PFA is not yet established.

However, to check the PFA performance would not be degraded by changing from an analogue to a binary ECAL, then a "no-harm" study was done. In this, the PandoraPFA programme was run on an analogue ECAL with $5 \times 5 \, \text{mm}^2$ pads within the ILD detector framework [7]. To compare directly with no changes to the PFA programme, the binary ECAL MIP counts were combined into pseudo-pads of the same size. Specifically, the total MIP count of all clusters centred in the same area as each equivalent analogue ECAL pad was found and this value used to give an ADC-equivalent value to that of the analogue ECAL. By adjusting a single calibration factor multiplying the MIP count, then the binary ECAL data were put into an identical scale and format as the analogue ECAL and the PandoraPFA programme could be run on both with no changes.

The results for reconstruction of pairs of jets from the process $Z \rightarrow q\bar{q}$, for $q = u$, d or s quarks, are shown in figure 45. This shows the resolution on the Z mass is effectively unchanged
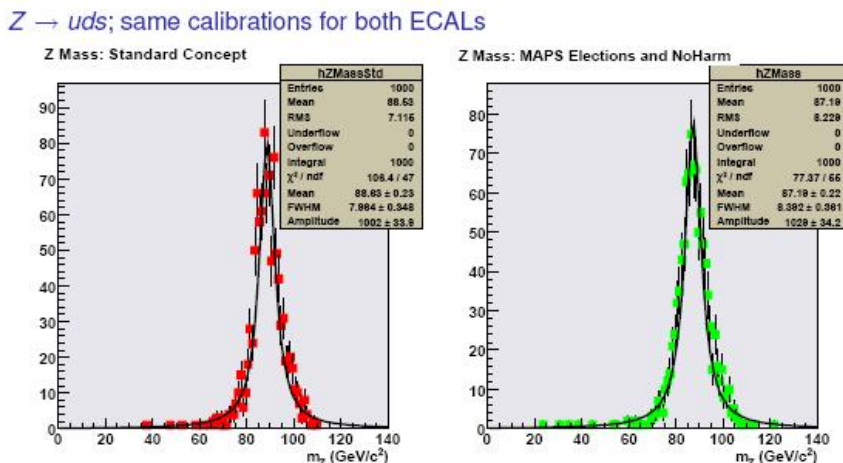

Figure 45: *Z mass resolution using hadronic jet reconstruction for (left) an analogue ECAL and (right) the equivalent binary ECAL with hits grouped according to the "no-harm" case.*

between the analogue and binary ECALs. This demonstrates the binary ECAL will be able to give at least equal performance to an analogue ECAL for PFA. Indeed, as much of the fine granularity information is discarded in the above study, then it may be that a significantly improved performance would be possible if the PFA is tuned to use this information.

7 Future prospects

This work is planned to continue with a new sensor, TPAC1.1, which has been designed and fabricated and is about to undergo first tests at the time of writing. This sensor is the same size as TPAC1.0 and has the same number of pixels. It is I/O compatible with TPAC1.0 and so can use the same sensor PCB and DAQ system.

The new sensor will only contain the pre-shape quadrant 1 pixel design. It will also implement several fixes, such as the decoupling of the comparator and monostable power meshes and correcting the SRAM memory write corruption. Small adjustments have been made to the pixel layout to improve the pixel uniformity for noise and gain.

If this sensor functions as expected, then it should be possible to use it in an electron beam test to generate real EM showers and sample the hit densities following various thicknesses of

tungsten. This would give invaluable data on the fine granularity properties of EM showers and so would allow much firmer predictions of the realistic performance of a binary ECAL.

The actual verification of a binary ECAL resolution could not be done with TPAC1.1 since the sensor is only $1 \times 1\,\mathrm{cm}^2$. A real demonstration of a binary ECAL stack would require a large scale sensor but this is not available at this time.

8   Conclusions

We have presented a study of a novel concept for sampling electromagnetic calorimetry; namely a calorimeter with binary readout. This requires very fine granularity and a high level of sensor integration to handle the very large number of pixels which are required.

We have produced a CMOS test sensor which demonstrates many of the basic parameters required for a binary ECAL. Test results show the sensor functions close to the expected level. As part of the sensor development, we worked with the CMOS foundry to develop the INMAPS process, which enables us to use a deep P-well implant to shield the signal charge from the active PMOS circuit elements. Measurements show this improved the signal efficiency by large factors and makes the sensor viable.

The overall performance of a simulated full scale binary ECAL based on the test sensor results has been shown to give a better resolution for electromagnetic showers than an equivalent analogue readout ECAL. The dominating issue for the resolution is the confusion of counting charged particles from the hits and this will require real data from EM showers at very fine granularity to be sure of the limiting resolution of a binary ECAL.

A second round sensor is planned for the near future and this will have a uniform pixel design throughout the active area. This should enable the critical EM shower measurements to be performed.

9   Acknowledgements

References

[1] "International Linear Collider Reference Design Report," (2007) http://www.linearcollider.org/rdr.

[2] "International Linear Collider Reference Design Report Volume 2: Physics at the ILC," ed A. Djouadi *et al.*, (2007) http://www.linearcollider.org/rdr; "TESLA TDR Part III: Physics at an e$^+$e$^-$ Linear Collider," ed. R.-D. Heuer *et al.*, DESY 2001-011 (2001).

[3] CALICE Collaboration, http://polywww.in2p3.fr/flc/calice.html.

[4] M. A. Thomson, "Particle flow calorimetry at the ILC," Proceedings of Hadronic Shower Simulations Workshop, Batavia, Illinois (2006), AIP Conf. Proc. **896** (2007) 215; M. A. Thomson, "Progress with Particle Flow Calorimetry," Proceedings of 2007 International Linear Collider Workshop (LCWS07 and ILC07), Hamburg, Germany (2007), arXiv:0709.1360.

[5] J. E. Brau *et al.*, "Calorimetry for the NLC detector," Proceedings of the 1996 DPF/DPB Summer Study on New Directions in High-energy Physics, Snowmass, Colorado (1996) eConf C960625 DET077.

[6] SiD Detector Concept, `http://silicondetector.org/`.

[7] ILD Detector Concept, `http://www.ilcild.org/`.

[8] GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Inst. Meth. **A506** (2003) 250.

[9] Guinea Pig beam simulation,
`http://www-sldnt.slac.stanford.edu/nlc/programs/guinea_pig/gp_index.html`.

[10] G. Deptuch *et al.*, "Simulation and measurements of charge collection in monolithic active pixel sensors," Nucl. Inst. Meth. **A465** (2001) 92.

[11] J. A. Ballin *et al.*, "Monolithic Active Pixel Sensors (MAPS) in a quadruple well technology for nearly 100% fill factor and full CMOS pixels," to be published in Sensors (2008), arXiv:0807.2920.

[12] A. Papoulis and S. U. Pallai, "Probability, Random Variables and Stochastic Processes," Section 11.4, 4th edition, McGraw Hill (2002).

[13] Sentaurus TCAD, `http://www.synopsys.com/products/tcad/tcad.html`

[14] J. A. Ballin *et al.*, "A MAPS-based Digital Electromagnetic Calorimeter for the ILC," Proceedings of 2007 International Linear Collider Workshop (LCWS07 and ILC07), Hamburg, Germany (2007), arXiv:0709.1346.

[15] CALICE Collaboration, C. Carloganu *et al.*, "Calice Si-W EM Calorimeter: Preliminary Results of the Testbeams 2006," Proceedings of 2007 International Linear Collider Workshop (LCWS07 and ILC07), Hamburg, Germany (2007), arXiv:0709.2516.