Pixel detector back-up Document to support the ATLAS Technical Proposal

P. Fischer, B. Raith, N. Wermes
University of Bonn, Germany

A. Lankford, S. Pier and B. Schmid
University of California, Irvine, USA

M. Campbell, P. Jarron, E. H. M. Heijne, P.[1] Middelkamp, W. Snoeys
CERN, Geneva, Switzerland

C.Arrighi, L. Blanquart, V. Bonzom, J-C. Clemens, M. Cohen-Solal, P. Delpierre, A. Fallou,
E. [2]Grigoriev, M.C. Habrard, G. Hallewell, D. Labat, L. Lopez, A. Mekkaoui, T. Mouthuy,
R. Potheau, M. Raymond, A. Rozanov, D. Sauvage, L. Vacavant.
CPPM, Centre de Physique des Particules de Marseille,France

D. Bintinger, A. Ciocio, K. Einsweiler, M. Gilchriese, O. Milgrome,
J. Millaud, D. Nygren, M. Shapiro, J. Siegrist, H. Spieler, M. Wright
Lawrence Berkeley Laboratory, California, USA

G. Capannesi, M. Colucci, P.G. Pelfer, S. Sottini
INFN and Physics Department of Firenze University, Italy

D. Barberis, M. Bozzo, C. Caso, M. Dameri, G. Darbo, P. Morettini,
P. Musico, B. Osculati, L. Rossi, G. Sette
INFN and Physics Department of Genova University, Italy

R. Bates, S. D'Auria, S. Gowdy, V. O'Shea, C. Raine, K.M. Smith
Glasgow University, UK

G. Bellini, M. di Corato, A. D'Avella, P. Inzani,D. Menasce, L. Moroni,
D. Pedrini, L. Perasso, F. Ragusa, S. Sala, F. Tartarelli
INFN and Physics Department of Milano University, Italy

G. Cesura, D. Hauff, H. Hoernl, J. Kemmer, P. Lechner, G. Lutz, R. H. Richter, H. Seitz
MPI Halbleiterlabor, Munich, Germany

Y. Gao, J. Harton, R. Jared, M. Walsh, S. Wu
University of Wisconsin, Winsonsin, USA

K. H. Becks, K. W. Glitza, J. Heuser, S. Kersten
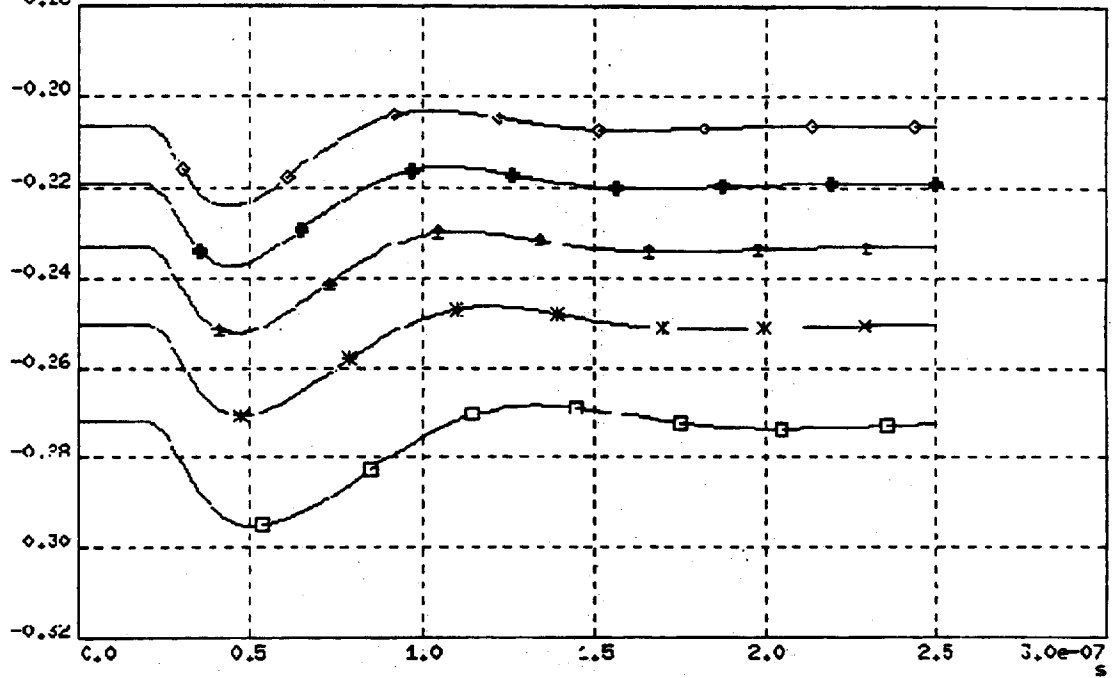*University of Wuppertal, Germany*

---

[1]Also at University of Wuppertal
[2]On leave of absence from ITEP, Moscow, Russia

Figure 97: Analogue outputs. Signal=6000e. Leakage current $0 \rightarrow 100$ nA.

Figure 98: Digital Output. Signal=6000e. Parameter: leakage current.

File : FeCell1.cou
Sign: 500->20000e-; Il=20n; th=500e-

◇V(XANALOGN.CUT)_1:2 ●V(XANALOGN.OUT)_2:2 ▲V(XANALOGN.OJT)_3:2
*V(XANALOGN.CUT)_4:2 □V(XANALOGN.OUT)_5:2 ○V(XANALOGN.OJT)_6:2
■V(XANALOGN.CUT)_7:2 ●V(XANALOGN.OUT)_8:2 ◆V(XANALOGN.OJT)_9:2

◇V(OUT)_1:2 ●V(OUT)_2:2 ▲V(OUT)_3:2 *V(CUT)_4:2 ⊓V(OJT)_5:2
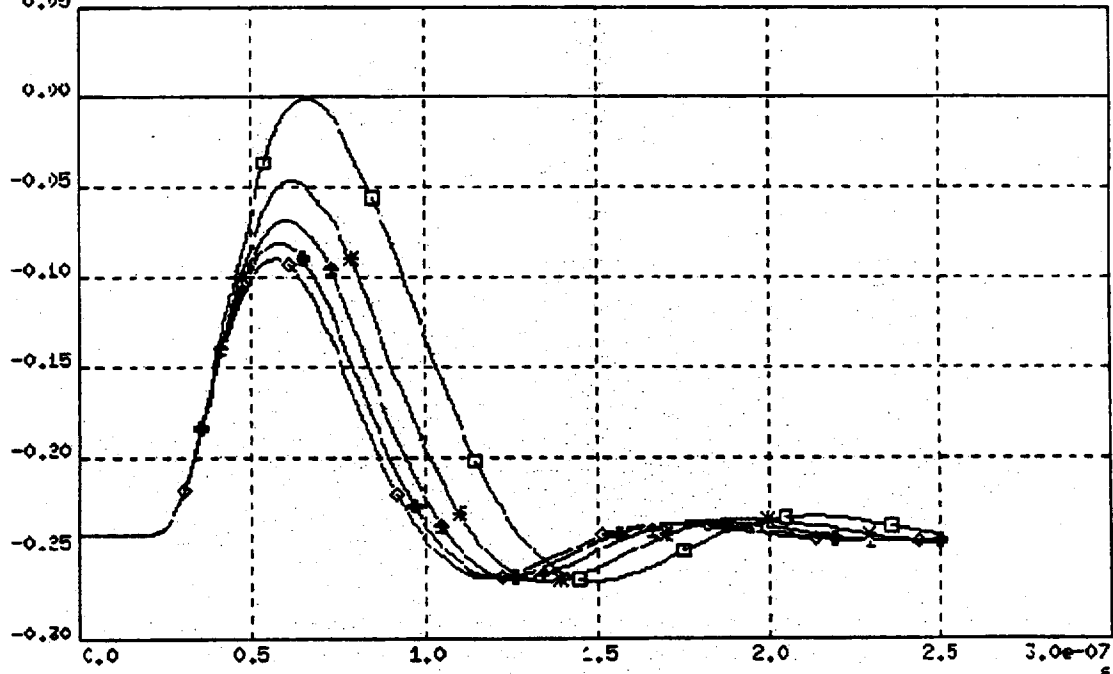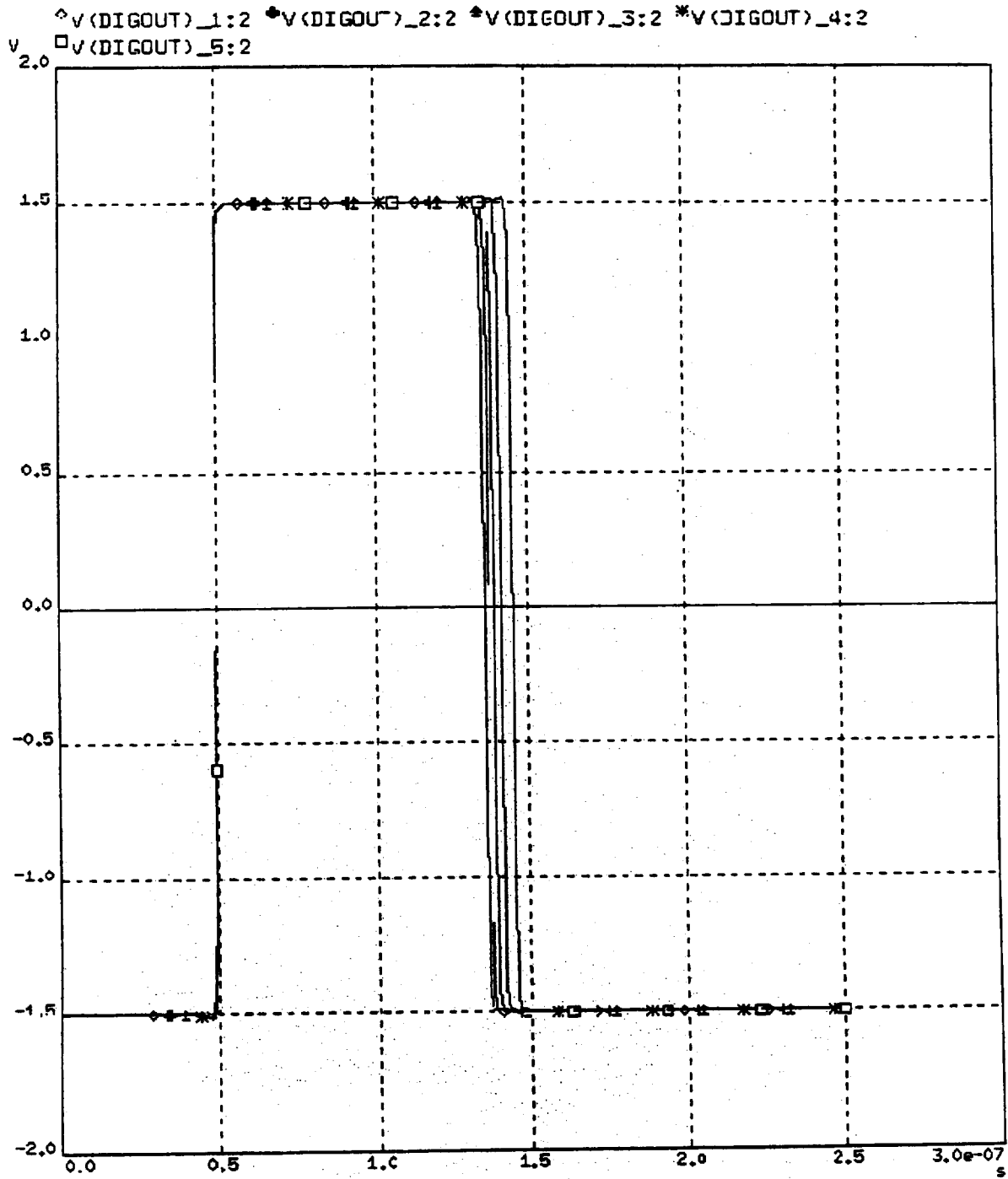○V(OUT)_6:2 ■V(OUT)_7:2 ●V(OUT)_8:2 ◆V(OUT)_9:2
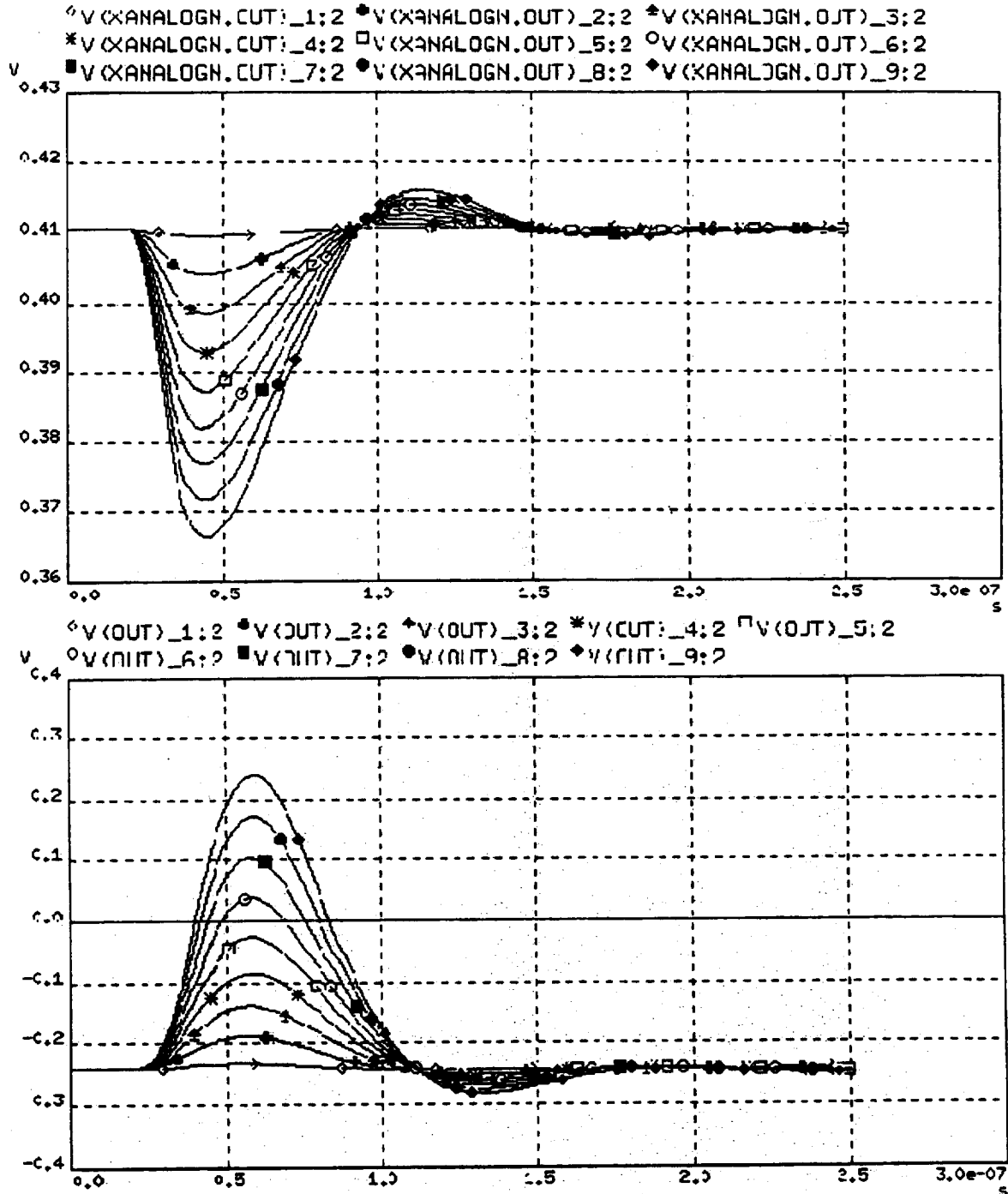
Figure 99: Analogue Output. Parameter: Input charge. Ileakage=20 nA

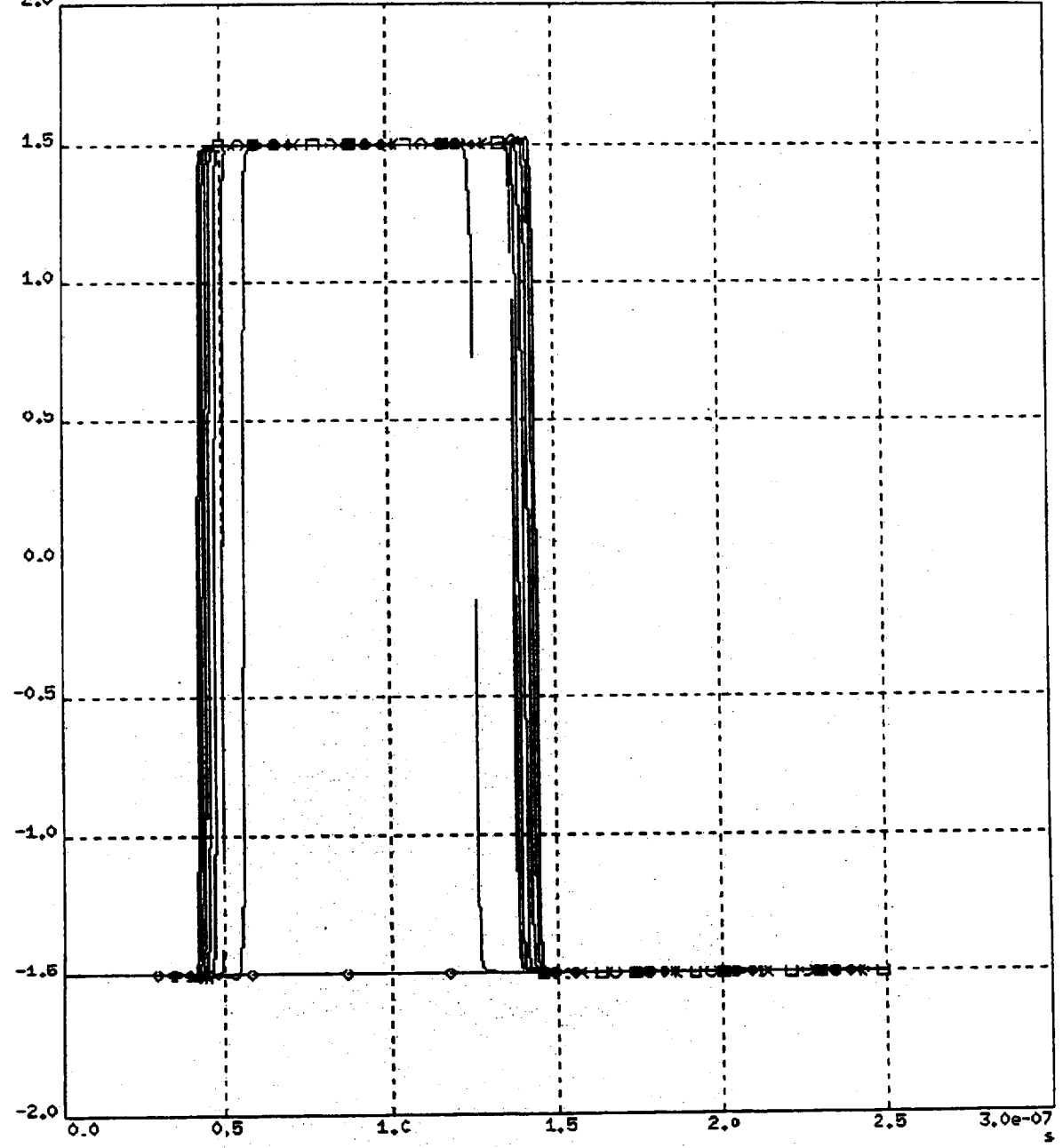Figure 100: Digital Output. Parameter: Input charge. Ileakage=20 nA

File : FeCell1.cou
In:2000,3000,100000e-; Il=20n; Th=2000=-

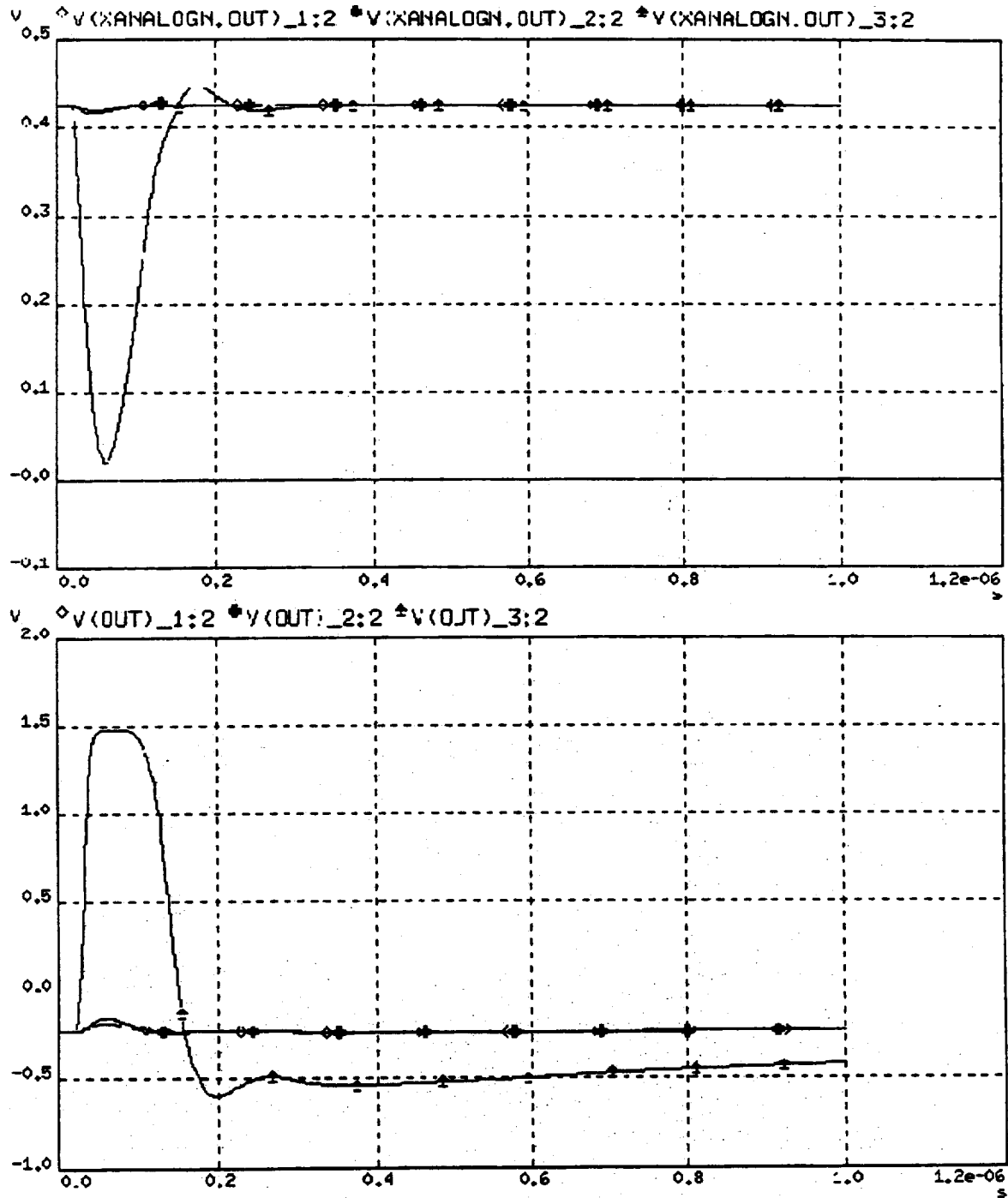V 0.5 ◇V(XANALOGN.OUT)_1:2 ●V(XANALOGN.OUT)_2:2 ▲V(XANALOGN.OUT)_3:2

V 2.0 ◇V(OUT)_1:2 ●V(OUT)_2:2 ▲V(OUT)_3:2

Figure 101: Digital Output. Parameter: Input charge. Ileakage=20 nA

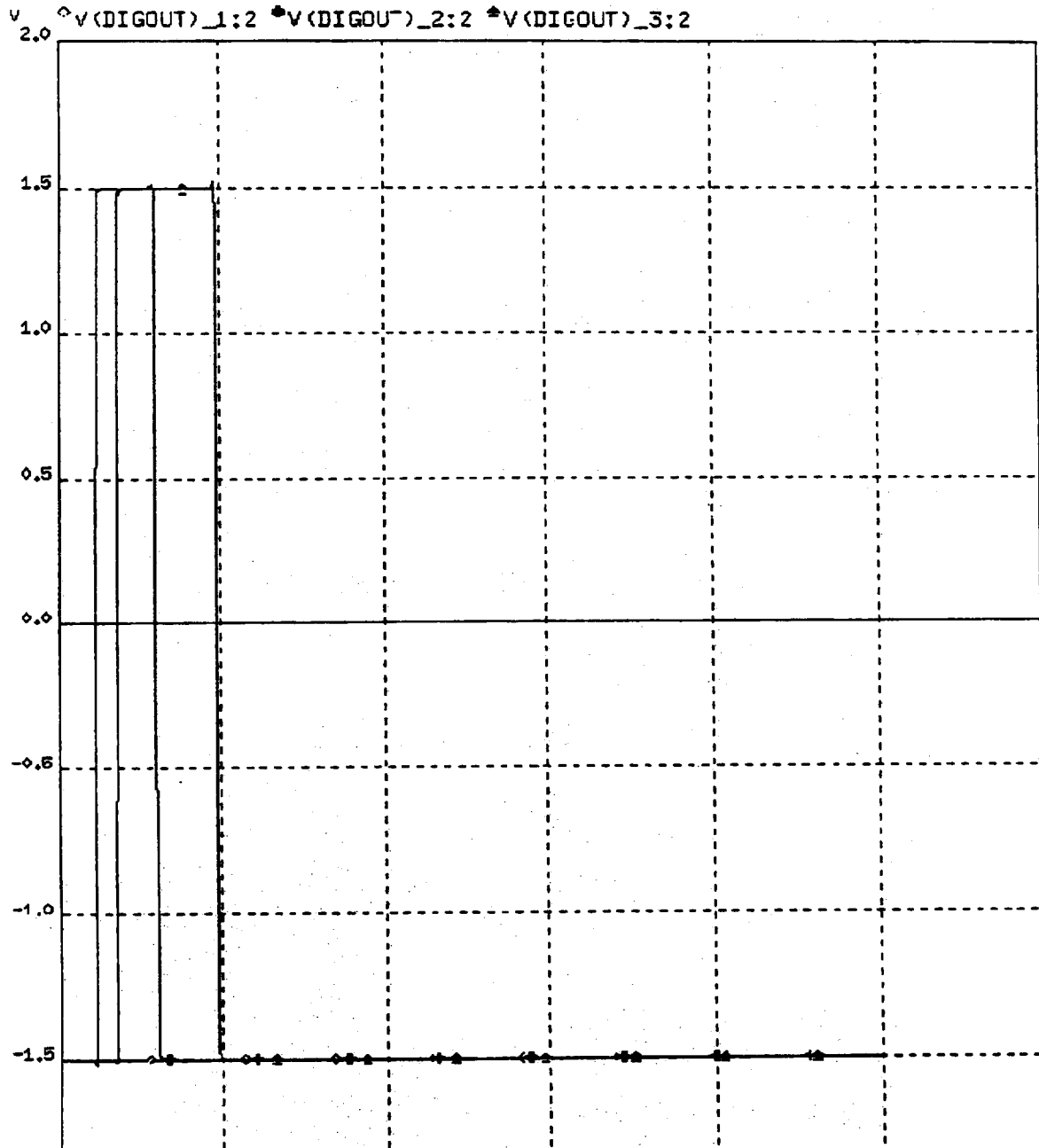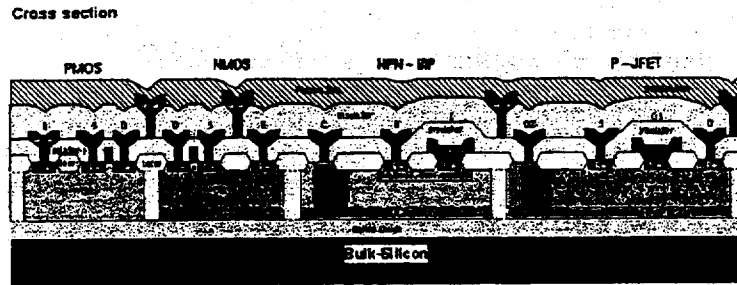Figure 102: Digital Output. Parameter: Input charge. Ileakage=20 nA

Figure 103: Schematic cross section of DMILL technology transistors

The planarisation of the process has been achieved with the "BOSON" reticle. This batch, restricted for consortium use, has been submitted in March 1993 and was completed in August 1993. The planarisation ensures a better manufacturing yield, allows the reduction of the minimum gate length to 0.8 μm and also significantly improves the integration density. First results on small scale analogue and digital circuits relevant to LHC applications were issued from the batch "BOSON". In parallel, measurements have been performed on elementary devices to study the radiation hardness of the planar technology [65]

Two other batches are currently in processing:

- "FERMION" is restricted to the consortium use and allowing corrections to circuits processed in "BOSON". It also includes higher levels of circuit complexity and integration, particularly an array of cells for the ATLAS silicon pixel detector.

- "HADRON" is a batch opened to first users not belonging to the consortium but belonging to the high energy physics community.

These batches are scheduled to arrive no later than November 1994.

The main features of DMILL technology are:

- 4 transistors available (NMOS, PMOS, NPN and PJFET) (Fig. 103)
- 0.8 μm minimum gate length
- 18 nm oxide thickness
- Two metal layers and one polysilicon layer
- Rad-hard capacitances with low voltage coefficients
- Low, medium and high values of rad-hard resistances
- SPICE level 3 parameters for CMOS are available
- SPICE parameters for NPN and PJFET are available

2. Results on circuits implemented by CPPM in the "BOSON" reticle

In this reticle [66], we have implemented for the front-end electronics:

103

- A charge sensitive amplifier

- A D Flip-Flop

(a) The charge sensitive amplifier

The aim was to build an amplifier suitable for the silicon pixel detector for ATLAS, with in particular, a rise time of the order of 25 ns to ensure a time walk jitter less than a few nanoseconds. It includes all of the available transistors of DMILL technology. The PJFET is used for its natural radiation hardness and its low noise, the NPN is used for its high transconductance and CMOS transistors are used for biasing. The schematic of this amplifier is shown in Figure 104. The active feedback tunes the recovery time and sinks the leakage current.
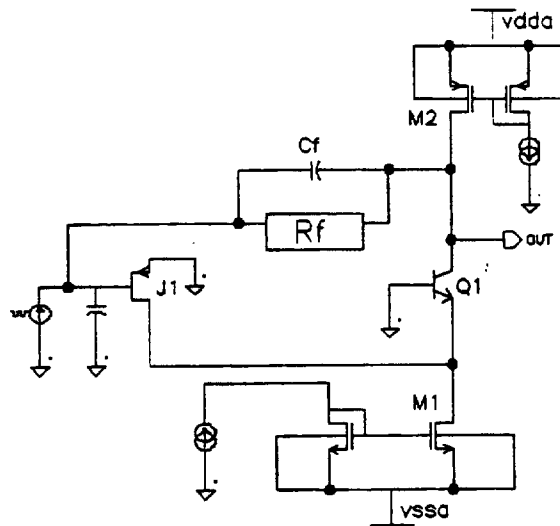

Figure 104: Schematic of the charge amplifier.

Measurements performed on this amplifier give a gain close to 15 μV/electron. The rise time is about 20 ns. Power consumption is about 36 μW and the silicon area is 50 μm $\times$ 50 μm. This amplifier has been irradiated up to $3.6 \ 10^{14}$ protons/cm^2, i. e., 18 Mrad(Si) at SATURNE. Transistors and circuits were biased during irradiation, and electrical measurements were made 30 min after beam stop. The beam intensity was measured continuously by counting the particles scattered from an aluminium target with scintillators connected to photomultipliers. Absolute dose calibration was measured from the activation of carbon cartridges. The beam intensity was about 10^9 protons/cm^2/s up to 10^{14} protons/cm^2 and then $5 \ 10^9$ protons/cm^2/s up to $4.5 \ 10^{14}$ protons/cm^2. Figure 105 shows the current impulse responses of the amplifier at different doses. The rise time, which is one of the most important parameters for the pixel detector, increases by about 15 % between 0 and 18 Mrad(Si). This shows that the amplifier works successfully even after a dose close to the dose expected for 10 years of LHC operation [4].
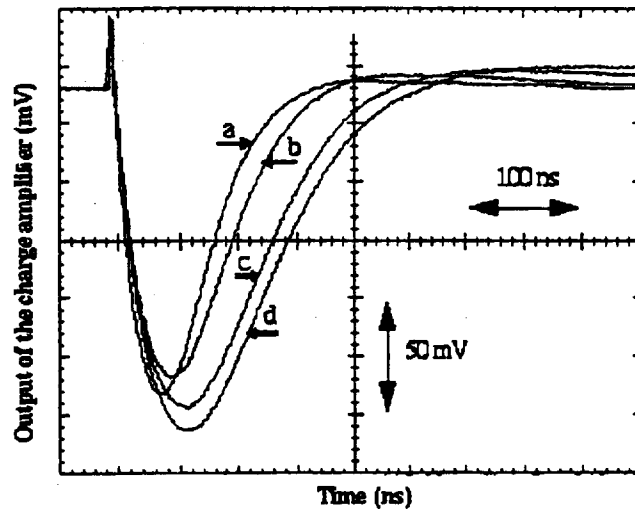
(b) The D Flip-Flop

104

Figure 105: Current impulse responses. a: before irradiation b: 150 Krad(Si) / 3 10^{12} protons(300 MeV)/cm^2 c: 9 Mrad(Si) / 1.8 10^{14} protons(300 MeV)/cm^2 d: 18 Mrad(Si) / 3.6 10^{14} protons(300 MeV)/cm^2

D Flip-Flops are extensively used in pixel electronics, particularly to hold and shift the signal at the output of the discriminator. **Figure 106** shows the schematic and the layout of a two phase dynamic D Flip-Flop implemented on the "BOSON" reticle.

The silicon area is 34 μm $\times$ 30 μm (to be 50 μm $\times$ 15 μm in the new version). A shift register including 33 dynamic stages operates correctly even at a frequency as high as 240 MHz.

Protons and pions with high energy ($\geq$ 100 MeV) may induce transient effects. DMILL technology is free from latch-up because of complete dielectric isolation between transistors. In addition, the "Single Event Upset" (SEU) sensitivity is reduced due to the SOI wafer [67]. To estimate this effect and its consequences for detector operation, we have carried out some SEU tests on the dynamic shift register. The device has been submitted to proton irradiation with beam intensity of 2.4 10^9 protons/cm^2/s. Figure 107 depicts the total number of errors per bit (per D Flip-Flop) as a function of integrated flux. This consequently gives a cross section of 3.3 10^{-13} errors.cm^2/bit, i. e., an error per dynamic D Flip-Flop every 139 hours in ATLAS at 11.5 cm radius from the beam (at the LHC nominal luminosity of 10^{34} protons/cm^2/s) [4].

In a detector of 10^8 pixels with 2000 pixels hit per BCO, 10 D Flip-Flops per pixel and a latency of 2 μs for the trigger level 1, the number of 'true' pixels not seen in one BCO is 8.1 10^{-8} and the number of 'false' pixels seen in one BCO is 14.2 10^{-5}. This is totally negligible. Further investigations to determine the dependence of the SEU rate on the rise and the fall time of the clock and also the
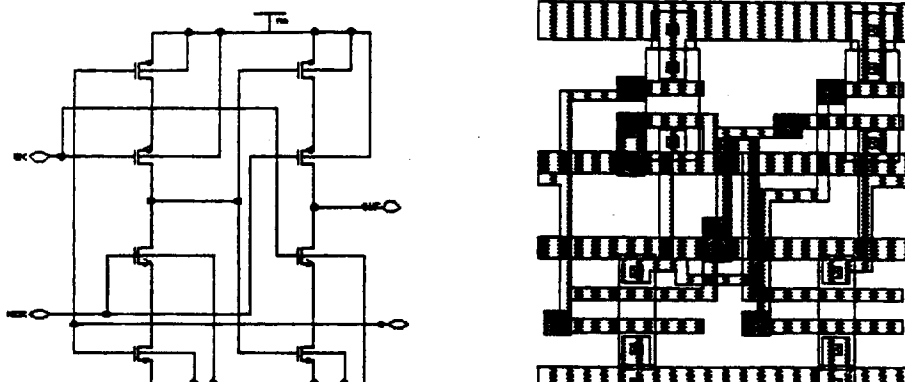
Figure 106: Schematic and layout of the D Flip-Flop.

dependence on the protons incidence angles are planned.
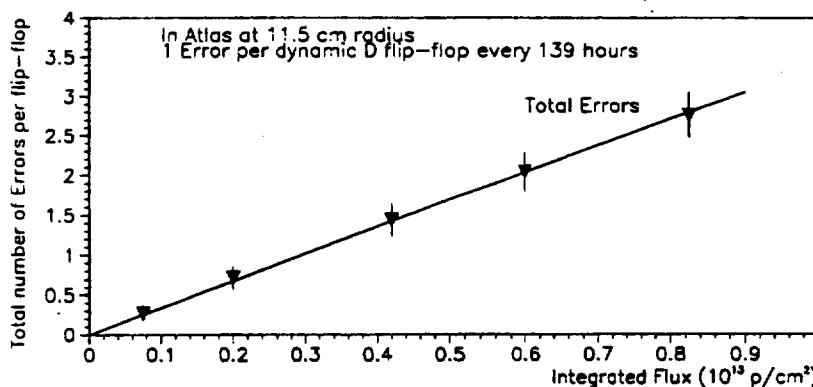
Figure 107: Total number of errors per bit (per D Flip-Flop) as a function of integrated flux.

(c) Circuits implemented by CPPM in the "FERMION" reticle

In this reticle are implemented:

- An enhanced charge amplifier with its expected power consumption is ≈ 30 μW.

- A first complete analogue pixel cell (amplifier + discriminator) called DMILLPIX1. This cell includes the previous charge amplifier and a very low offset discriminator. AC coupling between the first stage and the discriminator allows the comparator to be insensitive the DC level variations at the input amplifier. Furthermore, this improves the signal to noise ratio.

  The expected (simulated) features are:

  - Power consumption: close to 70 μW.

- Silicon area: 50 μm x 240 μm.

- Input stage gain: 15 μV/electron or 30 μV/electron.

- Rise time of the input stage: 20 ns.

- A second complete analogue pixel cell called DMILLPIX2. This cell has been developed to comply with power consumption and silicon area requirements. It has totally autonomous behaviour and all voltage variations (i.e. technology dependent variations and parameter shifts during irradiation) are completely self compensated.

The expected (simulated) features are:

- Power consumption: close to 45 μW.

- Silicon area: 50 μm x 130 μm.

- Input stage gain: 15 μV/electron or 30 μV/electron.

- Rise time of the input stage: 20 ns.

- A prototype (32 × 8) pixel array. It includes DMILLPIX1 and DMILLPIX2, 4 test-in and input transistors (bipolar structures) used for photo-injection. The central part (16 × 8) is built to be bump-bonded to a silicon detector. The purpose is to investigate the cross-talk between the analogue front-end and the detector, and the coupling between 2 analogue front-ends. Furthermore, we want to study threshold mismatch in putting the pixel cell in its real environment (detector, pad and neighbouring cells). To be totally independent of the digital part, a simple shift register based read-out system has been implemented. Trenches and shields ensure a complete isolation for this part and eliminate logic signal crosstalk. Figure 108 shows one of the chips submitted to the "FERMION" batch. On the left part is the pixel array and in the right several channels of DMILLPIX1.
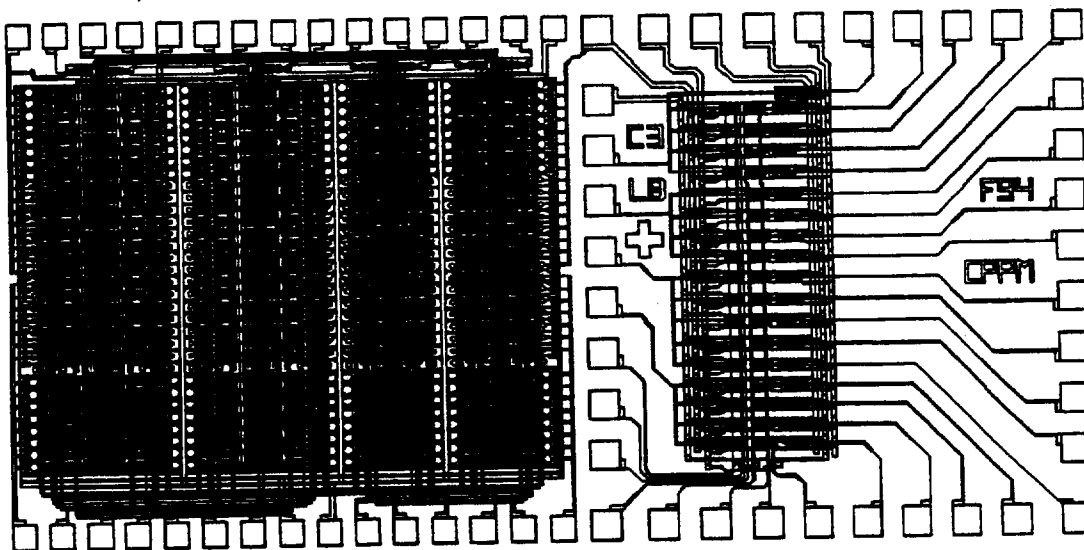

Figure 108: One of CPPM chip from "FERMION" reticle.

(d) Circuits implemented by CPPM on the "HADRON" reticle

Figure 109 shows the chip submitted to "HADRON" batch. It is a 16 × 8 array with read-out having LHC functionality. The aim is to validate an ATLAS specific prototype. This array includes DMILLPIX1 and DMILLPIX2 and is built to be fully bump-bonded to a silicon detector. Several different shielding configurations are implemented for study (see section 3.8.2).
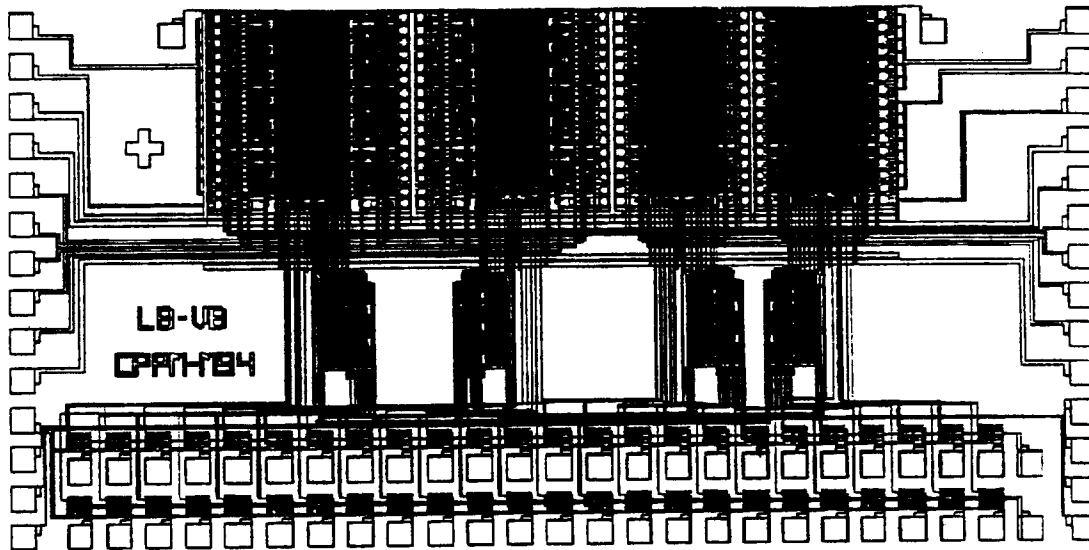
Figure 109: CPPM chip from "HADRON" reticle.

3.7 Pixel Array Validation and Tests

Many complex elements are put together in the construction of a ladder or of a wedge. In a typical ladder 176 readout chips should be bump bonded on 16 detectors on which there is a total of ≈675000 pixels plus all the necessary bussing for data transmission, and, eventually, few other control chips for data handling. It is obvious that only known good dies (KGD) can be assembled. Otherwise yield at the ladder level could be unacceptably low (a few percent). Testing a bare die before assembly must not impair it's "bumpability" and prevent assembly. The technology of KGD assurance is not well understood yet, but rapid progress is being done. One should be aware that the cost of quality assurance may be substantial and is likely to be higher than hybridization costs. In particular, tests should be done at each level of detector hybridization. Individual elements or dies should be tested, tests should then be repeated at each hybridization step (each module should be tested and only modules with number of faults below a given threshold can be used to form a ladder, then the complete ladder should be tested and charachterized). The design of each element of a ladder should take into account its "testability", therefore one should foresee to inject charge at each pixel input and be able to readout electronics chips using probe cards. Use of infrared laser beams and β sources should be considered when detector and electronics are conneted together.

Tests of hybrids with a ^{90}Sr β source has been successfully done during the preparation

108

of the Omega pixel detectors for the WA97 experiment. Irradiation was done connecting the hybrid to the readout system through specially designed probe cards. This test has allowed to reject malfunctioning hybrids and has increased considerbly the quality of the detectors ($\leq$ 3% dead area over 90 cm^2 active area). However, we still miss experience on very large scale systems (we have to go from $\approx$100 cm^2 to $\approx$2000 cm^2) and we believe that testing and qualification of the hybrids is a substantial task which remains to be adressed in detail.

3.8   Readout Architectures

The average number of pixels per layer at ATLAS will be in the range 10^8, while the average number of hit pixels (at a nominal luminosity of 10^{34} cm^{-2}s^{-1}) during a 2 μs level 1 trigger latency will only be in the range 10^4 - 10^5. For acceptable readout time and power dissipation, it is therefore important to address only the hit pixels, and only those for BCOs (25 ns time stamps) with a valid level 1 trigger. Essential features of the pixel detectors are their timing resolution and their capability of resolving hits inside high multiplicity jets. Both of these potential advantages should be fully exploited by an adequate readout architecture.

A number of data-driven and clock-driven readout architectures have been proposed or are under development [38, 35, 68, 29]. Two of these are already implemented for the Omega WA97 project [29] and for the DELPHI-VFT project [38]. Practical experience will soon be gained with the use of these architectures in real experiments. A third architecture is at the test chip level.

The OMEGA architecture [29] has autonomous pixel operation with storage of the binary hit data in a pipeline memory in each pixel until the arrival of a level 1 trigger pulse, which is broadcast to all pixels in the system. This strobes the data into a long - time memory location which is read out or can be reset depending on the level 2 trigger. Sparse data readout and pixel masking are currently implemented on the VME card and will eventually be migrated onto the chip. The level 1 trigger pulse itself generates the time stamp to be encoded with each hit position on the chip.

In this architecture, there is no clocking activity on the analogue part of the detector until a level 1 trigger. More studies are required to establish whether the pixel pipelines can be made sufficiently precise to allow 2 μs depth with 25 ns resolution. Recently, 72 chips have been operated in a pair of detector arrays in the OMEGA WA97 test run. The timing precision was found to be limited by voltage drop in the power supply lines on the ceramic detector support. Although in the lab, clocking at up to 16 MHz worked well, only 2 MHz could be used in the experiment, illustrating that large scale testing of prototypes is indispensable to discover problems related to overall system implementation.

The DELPHI VFT architecture with sparse readout [37] is shown in fig 23. After a pixel is hit, it closes two switches; one connects the "row" bus and the "column" bus and the other generates a signal that opens another switch in the row daisy chain. During the readout, a hit pixel can be identified and its address stored on the chip periphery, together with a time stamp. In its present form the chip does not have time stamping capability.

Prototypes of 1024 pixels have been already produced and tested by flashing different configurations of light spots. Readout chip operation with clock frequencies up to 12 MHz has been verified.

3.8.1 Pixel read out and architecture design in LBL

3.8.1.1 LBL–4 Array LBL–4 is a 12 column by 64 row pixel unit cell array, combined with 12 columns of read-out logic. This is the first two-dimensional pixel array implementing the column read-out (CRO) architecture, developed at LBL. The pixel unit cell (PUC) circuitry is a third generation design, and is expected to meet many of the performance requirements for ATLAS. LBL–4 employs data-driven logic with continuous-time signal filtering within the pixel unit cell to minimize interactions between the pixel array and the associated peripheral logic.

LBL–4 incorporates some PUC test features and hybridization compromises that boost the pixel unit cell area requirements somewhat. The active circuit area of the PUC is $50 \times 350 \mu m^2$. To permit hybridization at the prototype die level, rather than at the wafer scale level, extra area was devoted to increase the spacing between bump pads. The total area of the PUC for the LBL–4 prototype is $50 \times 536 \mu m^2$. The total active area of the corresponding 12×64 detector array is 3.2×6 mm^2. A streamlined "final" PUC design is expected to be in the range of $50 \times 300 \mu m^2$, with an array size of perhaps 32 columns $\times$ 192 rows, for a total of 6144 active pixels, and an active area of 9.6 mm $\times$ 9.6 mm.

3.8.1.2 General Goals of LBL–4 The overall strategic goal of the LBL–4 chip is to demonstrate most of the performance requirements for the ATLAS environment; more specifically, to demonstrate:

1. the upgraded pixel unit cell (PUC) design functionality;

2. the column read out (CRO) architecture when coupled to a PUC array;

3. adequate channel isolation in a large two-dimensional (2D) array;

4. full speed operation and time resolution;

5. operation in a hybridized form with a detector diode array;

6. particle track reconstruction in a test beam *milieu*;

This experience is needed to provide the basis for a more advanced prototype design with improved I/O and control characteristics.

3.8.1.3 LBL–4 Performance Goals The performance specifications for an ATLAS pixel array are strongly influenced by the extremely challenging operating conditions of the LHC as well as the underlying physics goals. This section deals mainly with the electronic requirements imposed by these two demanding aspects, including the innovative strategies developed to realize a practical system.

The full set of performance requirements for an ATLAS pixel vertex detector include radiation hardness specifications that have not been addressed in current work due to cost considerations. The specifications for LBL–4 have been chosen as follows:

- **Power dissipation: ≤ 0.25 W/cm^2.** This specification is chosen to minimize the contribution of the pixel cooling system to the overall material budget. It is also chosen with a desire to roughly match the average power dissipation of a strip system of similar area. The specification was, however, chosen with knowledge that low-power pixel circuitry meeting other performance requirements is possible.

- **Electronic noise: ≤ 200 electrons, rms, referred to input.** This specification is a tradeoff between power dissipation, and circuit bandwidth and yields anegligible noise hit rate for a large pixel system. The very small detector capacitance, relative to strips, makes this a practical goal. This value represents a reasonable choice for an imposed power dissipation limit of ≤ 0.25 W/cm^2.

- **Minimum signal: 1000 electrons;** This number refers to the actual threshold of the comparator in the pixel unit cell. It is taken to be large compared to the intrinsic electronic noise ($\simeq 5\sigma$) to reduce noise hits to a negligible level but is small enough to yield high efficiency after radiation damage. This specification must be considered together with threshold variations among pixels.

- **Pixel Cell Threshold Uniformity: ± 200 electrons maximum excursion from mean threshold.** This is a challenging electronic design task in CMOS. Meeting this goal has required an innovative design approach.

- **Time-walk less than 14 ns for input charge from 1 fC to 8 fC.** The time-walk spec involves a tradeoff between power dissipation (*i.e.* speed) and efficiency for small pulses resulting from charge-sharing. It is known that small pulses near the threshold will fall into subsequent time-stamp buckets. Presumably that pixel receiving the larger part of the shared charge will result in a proper time-stamp for the hit and the next time bucket can be read out if better resolution is desired.

- **Efficiency: not specified *a priori*.** The overall efficiency will be affected by several small factors, the largest being dead-time associated with pixels waiting for the Trigger level I latency. Nevertheless, it is expected that efficiency should approach 99%.

- **Linearity of Charge information: $\leq 10\%$ over the charge range 1 fC to 8 fC.** This fairly relaxed spec comes from the forgiving nature of the charge interpolation formula, and is not particularly difficult to achieve.

- **Maximum dark current at input: 20 nA.** This is due to radiation damage to the detector, after several years at a luminosity of 10^{34}cm^{-2} s^{-1}. Some uncertainties affect this spec, which depends on radius of operation, detector temperature, machine backgrounds *etc.*

In addtion to these goals, other specifications relevant to the specific circuit design were established, and will be discussed in context.

3.8.1.4 Column-based Architecture Consideration of several distinct architectural concepts for array organization led us to reject traditional x–y schemes for a variety of reasons. By x–y scheme, we refer to the approach where peripheral registers store uncorrelated x and y data as hits occur, with recorrelation being made as valid data are requested, all on a chip-wide basis.

The major limitation of the traditional x–y scheme in the ATLAS context is an excessive rate of activity in the circuitry recording hits. While not viewed as an impossible approach, the approximately 5–10 MHz average hit rate for a chip complicates simulation of chip behavior enormously, and appears to introduce mechanisms for loss of efficiency and registration errors. Other objections are the inefficient use of peripheral registers, which would store mainly zeroes, and the dead-time associated with readout.

Alternatively, we have chosen an architecture in which each column acts as as independent logical element, detecting, time-stamping, and reading out valid data independent of other columns. The advantages in this approach are several. First, the maximum activity of the circuitry is reduced by a factor equal to the number of columns. Second, there is no peripheral circuitry except on one edge; this allows several pixel array readout IC's to be abutted on one larger detector element. Third, the logical organization and activity of the chip is now one-dimensional, greatly simplifying the design task.

The column read-out (CRO) architecture is regarded as a major advance toward the goal of an array capable of processing the extremely high rates associated with LHC luminosity levels of 10^{34}cm^{-2}s^{-1}. In simplest terms, the CRO eliminates all transverse communications within the pixel array. The sole penalty appears to be a modest increase in the PUC area due to the need to provide an addressing function for each pixel.

The more detailed discussion of LBL–4 functionality is given below, beginning with the end-of-column logical elements, the column addressing logic, and concluding with the pixel unit cell.

3.8.1.5 End of Column Logic Description The present LBL–4 pixel array contains a relatively complex prototype of the circuitry required to transfer data from the individual hit pixels off the pixel array chip to the outside world. As the present design is column-based, the column with its associated row of pixels and the data collection logic at the end provides the fundamental unit of the data acquisition system. The individual columns operate in parallel, and are connected together with multiplexing circuitry so that each may be read-out in sequence. In the ATLAS context, it is anticipated that the data would be transferred off the pixel array chip after each Level 1 Trigger.

The descriptions that follow are a summary of the functionality presently implemented in the LBL–4.2a prototype chip. This design is evolving as progress is made towards a pixel array which could be used in a large-scale pixel detector such as that proposed for ATLAS. The present design does not claim to solve all of the neccesary problems, but should allow rather sophisticated measurements to be performed in a test-beam environment.

The pixel array must carry out three basic tasks concurrently in order to operate in the LHC environment. The characteristics of this environment from the data acquisition point of view can be summarized in a few basic numbers (given here for the design luminosity of 10^{34}). The occupancy of a pixel is estimated to be about 10^{-4} per beam-crossing, leading to occupancies of 1% for a column, and roughly 1 (that is one hit pixel) per pixel array chip.

112

Note that once an individual pixel is hit, it must either reset itself after some delay, or if it is associated with a Level 1 Trigger, it must hold-off its reset until it has been read out. Thus, the deadtime caused by the 10^{-4} occupancy will be about 1% (assuming that a pixel must wait typically 2–4 μsec before it is reset). This makes it essential that the remainder of the data collection process be highly concurrent to avoid further deadtime generation.

In the present ATLAS DAQ architecture, all of the data relevant to a particular beam crossing must be transferred from the pixel arrays after a Level 1 Trigger. In ATLAS, the Level 1 Trigger requires a fixed latency of roughly 2 μsec, and can occur at a rate of up to 100 KHz.

The necessary concurrent tasks are:

- Acquire Data. The individual pixel unit cells must create small deadtime during this phase, and hence operate almost entirely asynchronously.

- Trigger Filtering. The first level trigger for ATLAS is expected to provide a rejection of roughly 400 compared to the total crossing rate of 40 MHz. Thus, most of the data that causes hits in the pixels is of no interest and must be ignored.

- Data Sparsification and Readout. The present LBL–4 prototype chip contains 12 columns of 64 pixels. The pixel array chip proposed for use in ATLAS contains 6144 pixels in the one cm^2 active area of the 2-D array. Only one of these pixels is likely to be hit in a typical beam-crossing while operating at the LHC design luminosity of 10^{34}. Sparsification is essential.

It is essential that there be minimal digital activity throughout the pixel array (typical threshold settings are likely to be about 1000 electrons), in order to avoid inducing many false hits. This has led to a design which is data-driven, that is, no activity occurs within the pixel array until pixels are hit. It also leads to a design which seeks to minimize the signal traffic along the columns during operation.

The signals required for operation of the End of Column logic are simple. The inputs are:

- An 8-bit beam-crossing number (in Grey code). This code must have a period equal to the Level 1 Latency in order for the present logic design to operate correctly. It is possible to implement such a Grey code for any even integer (e.g., 80 crossings) using a lookup table.

- A TriggerAccept pulse which should occur precisely one Level 1 latency after the beam crossing of interest.

The outputs are:

- The 8-bit beam-crossing number for each hit pixel.

- The 6-bit row address for each hit pixel (the prototype LBL–4 chip has 64 rows per column).

113

- The analog pulse height stored on the pixel. It can easily be imagined that this signal would be digitized to 6 bit precision using a simple 10 MHz ADC in future versions of the chip.

The basic control signals are:

- ReadRequest and ReadAcknowledge. These are a pair of signals used to transfer data from different beam-crossings from the Output Buffer. Once data from a given beam-crossing is read from the Output Buffer, the next step is to transfer the pixel data for that crossing using the sparse scan logic, using the following signals:

- LaunchRead and ReadDone. These are a pair of signals to transfer pixel data from a given beam-crossing. They initiate and signal the end of a sparse scan of one column.

- ReadNext and PucOutput. These are a pair of signals to transfer individual pixels from a given crossing. They are clocked to synchronize the data transfer during sparse scan of a column.

3.8.1.6 Functional Blocks The following list gives a brief summary of each of the functional blocks required to implement the capabilities present in the LBL-4 pixel array.

1. The pixel unit cell itself is the source of data, and it stores the following information: a bit to define whether the pixel is hit, two bits to define a pointer that associates a hit pixel with a given beam-crossing number (the full 8-bit Grey Code beam crossing number is only stored in the end of the column Event Buffer, the 2-bit pointer to this number is stored in each hit pixel), and an analog charge representing the total signal collected on the pixel within the sensitive time of about 20 nsec.

2. The pixel unit cells inform the End of Column logic that they have been hit by using a wire-OR of the individual pixel HIT signals. This information causes the End of Column logic to register the presence of data in the column, and to store the current beam-crossing number in an Event Buffer. There are two additional data fields which are passed up the column from the End of Column logic to the individual pixels. The first is a WritePointer which is latched by the pixel cell when it is HIT in order to associate itself with a beam-crossing. The second is a ReadPointer which is used subsequently by the pixel cell to decide when its beam-crossing is being read out (the match between the latched WritePointer and the current ReadPointer causes a hit pixel to participate in the sparse scan readout).

3. The End of Column logic contains an Event Buffer to store beam crossing numbers for pixels within a column which have been HIT within the last Level 1 Latency Period. Data is automatically removed from this buffer after the latency period to avoid using stale hits. This buffer contains storage for four hits, and refuses to increment the WritePointer once it becomes full.

4. A comparator circuit is constantly checking for matches between the present beam crossing number and that of the oldest hit in the Event Buffer. When a match is

114

found, this indicates that a HIT occurred in this column exactly one Level 1 Latency Period before the present time.

5. A coincidence between the TriggerAccept signal and the comparator MATCH described above initiates the transfer of the beam-crossing number and corresponding buffer pointer into a second level of buffering, referred to as the Output Buffer. This process filters the information stored for HIT pixels so that only the information for beam-crossings with associated Level 1 Triggers is kept.

6. The column contains sparse scan logic to facilitate the transfer of information only from pixels which have been hit in a given beam-crossing. This readout involves cooperation between the Output Buffer, which sends the ReadPointer up the column to notify the pixels which beam-crossing is presently being read out, and the sparse scan logic distributed along the column. The sparse scan is then clocked by the ReadNext signal which requests the next hit pixel in a column. The sparse scan logic is capable of skipping over un-hit pixels at a 1 GHz rate, before reaching an interesting pixel. The ReadDone signal terminates the scan of a given column. The column select multiplexer is then shifted to the next column with information, and the data from that column is transferred out.

3.8.1.7 Operation In order to present a more coherent picture of the End of Column logic, the following description provides a sequential description of the operation of the logic. The first two items below are the acquisition phase, the third is the trigger filtering phase, and the final one is the readout phase. All must operate concurrently.

1. When a pixel unit cell is hit, the resulting charge is integrated, and fires a discriminator to signify that the pixel is HIT. The pixel then also latches the present value of the WritePointer from the column, thereby associating itself with the beam crossing whose ID will be stored at the end of the column. The total charge is stored, using a Time-over-Threshold method, on a capacitor.

2. The pixel HIT signal travels down the column, arriving at the end of column, where it causes the present beam-crossing number (an eight bit Grey code value) to be stored in the Event buffer. The WritePointer in the End of Column logic is then incremented in preparation for another HIT.

3. A comparator is continuously examining the beam-crossing numbers stored in the Event Buffer. When it finds a MATCH between the present beam-crossing number and the oldest entry in the Event Buffer, it checks for the presence of a TriggerAccept signal. If this coincidence exists, the corresponding beam-crossing information is transferred to the Output Buffer. A Flag is set to notify the outside world that this column contains valid data with an associated Level 1 Trigger.

4. When the presence of data in a column is noticed, the read-out sequence can be initiated. This involves reading an entry from the Output Buffer, which simultaneously provides the beam-crossing number for the data to be read out, and sends the Read-Pointer up the column to notify any hit pixels that their beam-crossing is being read

115

out. The ReadNext signal may then be clocked to transfer data from each hit pixel. A corresponding row address and analog charge for each hit pixel will appear on the outputs. This process continues until all hit pixels from the column have been transferred out.

3.8.1.8   Pixel Unit Cell   Some important aspects of pixel unit cell functionality have already been mentioned in the previous end-of-column section. Recapitulating this discussion, it is useful to divide the operations of the pixel unit cell into four main areas:

1. **Hit Sensing/Dark Current Rejection;**

2. **Data Storage and Conversion;**

3. **Valid data recognition;**

4. **Data transmission/reset.**

Of these, the last two have been discussed as part of the end-of-column logic. This section focusses mainly on the analog signal processing design approach.

Hit Sensing/Dark Current Rejection:

The HIT signal itself, as noted above, is a very short 4 ns current pulse ORed together with all other pixel cells in the column. It is generated by the pixel unit cell comparator as the input charge exceeds the expected threshold of 1000 electrons.

Hit sensing is accomplished by a high gain single-stage integrator with an unprecedentedly high conversion ratio of charge to voltage, $\sim$2 V/fC. The conversion of 2 V/fC is achieved through the use of an extremely small feedback capacitor $\sim$0.5 fF. This value is about one order of magnitude smaller than typically attempted in such circuits. The resulting high voltage gain presents a large swing to the comparator even at the intended threshold input charge level of 1000 electrons ($\sim$ 1/6 fC), a rather small charge value. The capability to achieve the high Q/V conversion ratio in the single-stage integrator provides four very beneficial results.

First, the realization of high gain in a single stage reduces circuit complexity, power, and area relative to a multistage configuration. Second, the standing power in the comparator can now be made very small, as the relatively large voltage swing at the input results in rapid transition in the comparator to a high gain-bandwidth point of operation. Third, the comparator can be made with a small number of simple circuit elements, *i.e.*, cascaded inverters. Fourth, the variation in threshold offset due to unavoidable CMOS process fluctuations can be made small relative to the intrinsic electronic noise referred to input. The last point is perhaps the most compelling argument in favor of this design approach, since uniformity of thresholds is of paramount importance in a system with 10^8 elements.

On the other hand, the realization of such a small feedback capacitor with its attendant high Q/V conversion ratio leads to saturation of the integrator at relatively low Q values, less than one fC. As the integrator saturates, the input is no longer a virtual ground. In this case capacitive coupling to neighboring pixel detector elements can lead in principle to non-negligible cross-coupling of signal, an undesirable situation. An extensive cross-talk model was developed, that included all identifiable capacitive contributions and realistic values for

116

inter-pixel couplings. The results of simulation with this model and the pixel cell show that adjacent pixels are not typically triggered by the effects of integrator saturation.

The reset of the integrator is accomplished by a circuit element that functions as a synthetic inductor for DC, a resistor for small AC signal noise analysis, and as a constant current drain for large signals. The design of this part represented perhaps the most challenging part of the analog section. The DC capability is driven by the desire to use direct-coupled detectors, a great simplification in the detector fabrication process.

However, integrator stability concerns limit the magnitude of the DC component that can be compensated. It appears that up to about 25 nA (representing the effects of several years of radiation damage at maximum luminosity) can be safely absorbed.

Data Storage and Conversion:

During the trigger Level 1 latency, the pixel cell must store the fact that it was hit, the hit charge information, and the pointer information that allows the pixel cell to associate the hit with a particular beam crossing. It must also be prepared to forget stale hit information after the latency period has expired.

The integrator stability issues affecting DC compensation also limit the drain current that resets the integrator to about 6 nA, or 6 fC/μs. This means that the integrator will return to near-normal equilibrium within the alleged two μs trigger Level 1 latency for input pulses not exceeding 12 fC. This is noteworthy, since the measurement of input charge depends on the time-over-threshold of the comparator, as noted above. For valid trigger level 1 hits depositing more than 12 fC, readout may begin before conversion (*i.e.*, time-over-threshold) is complete. Such instances are rare, and the distortion of the charge datum is not regarded as a significant source of error for spatial resolution.

The presence of a hit is stored in a mono-stable during the latency period. The monostable period is externally adjustable, and is nominally set for a period of ~3 μs for initial testing. The additional time beyond the latency interval is to permit readout, and is set to allow the most complex likely event to be readout before reset occurs. This is not an ideal arrangement, but is adequate for test and prototype purposes.

For ATLAS, the use of the monostable could introduce unwanted extra deadtime. Subsequent designs will enable pixel cells to retain hits associated with a valid level 1 trigger based on a pointer comparison (read pointers from the EoC match those in the PUC) trigger level 1 condition. This would result in pixel reset automatically after the L1 latency unless valid data were present.

The Hit pointer is stored as a two-bit digital number in the current design. Any increase in buffer depth requirements will obviously introduce an additional bit. While not seen as a complication, this will increase circuit area non-negligibly. Earlier studies showed the feasibility of an analog pointer, and this approach will likely be reviewed for future design work.

Effort has been expended on making the PUC performance robust and repeatable. The specs should be maintained even after a radiation dose of 10 MRad SiO_2. In general, circuits were designed to 6 σ design rules.

The input signal polarity is chosen to be negative, anticipating the impact of radiation damage on detectors. Rather than pass through inversion from the familiar n-type to p-type, p-type detectors will be used *a priori*. The electron signal is faster than the hole current,

117

and contibutes to the excellent time-walk results.

3.8.1.9 Prototype Process and Layout

The radiation-soft Hewlett-Packard CMOS26 bulk CMOS, N well, 1 polycrystal Si, 3 Al metal layer, digital process was accessed via MOSIS prototyping. This is a superior low parasitic capacitance and resistance inter-connect process, with 1 μm (drawn) FET gate length, 2 μm poly pitch at 3 Ohm per square, 3.5 μm metal 1 pitch, 4 μm metal 2 pitch, 5 μm metal 3 pitch. The entire 3rd metal layer is devoted to power supply distribution such that the top of the chip looks like an AC ground in the expectation of reducing signal coupling into the detector inputs to a level below the electronic noise of each channel, at least for the case of only a few dozen PUC collecting hit charge on any one time.

The circuit design is directly mappable to a radiation hardened N well process such as UTMC or Motorola, but would require some re-design for a P well process, such as Honeywell's, because the present design puts some selected N wells at other than the most positive supply, as well as being optimized for multiple positive supplies while assuming only one negative supply to the substrate (ground).

There are 113 wire-bond pads (100 by 100 μm on 188 μm centers) around the perimeter of the prototype chip, and another 768 PUC pads in the array of 12-by-64 PUC. The PUC pads are 48 μm dia. metal pads under 40 μm diameter holes in passivation, on 100 μm centers, and expected to be big enough to wire-bond gold balls. A $50 \times 186\mu$m area in each PUC is spent on wire-bond pads, as opposed to the near-zero extra area needed for indium bump bond pads, because it is wished that prototype die be hybridizable, and prototype wafers are not available. The pads are cuts through 1/3 μm SiN, 2/3 μm SiO2, exposing the 2 μm thick AlCu alloy metal over 2 μm SiO$_2$ field oxide and Si wafer. The pads on the chip periphery are spaced 500 μm away from the PUC array, measured edge-to-edge.

3.8.1.10 Extensions to Future Prototypes

Our goal was to, first, build a prototype that serves as an proof of the architecture in its simplest form. For example, a near full-complexity PUC and End-Of-Column Read Out and 2 Level buffer are implemented, but the time stamp and much of off-chip read-out was left off the chip. This provides maximum flexibility, and also will allow the interface to the data collection and transmission chips to be defined when those chips are designed. It is expected that we will: double the End-Of-Column buffer depth, bring the Grey code counter into the chip, add logic to avoid timeout of PUC holding valid trigger level I data, remove bond pads to the bias sections, greatly reduce bond pad count, and move all bond pads to one edge of the die- the edge ajacent to the EOC read out.

In addition, there are other changes and additions that will be implemented in future submissions. They would predominantly be to support the interface to a data collection chip, and to put a more efficient read-out sequencer on the pixel array chip.

3.8.2 Read Out Architecture under development at CPPM

1. Abstract

The digital readout system under development at CPPM is based on the use of pixels linked into columnar digital shift registers. Hit pixels enter their wire-coded addresses

into the shift register. These address words shift down the column one step at each BCO. A non-active zone at the periphery receives these pixel addresses, implements a time tracking algorithm with a memory and identifies pixels hit in the BCOs with a valid level 1 trigger.

The proposed algorithm is robust and modular, and can accommodate changes in trigger latency in a wide band width around a nominal value.

Low power consumption, "self killing" of "stuck-at-one" pixels and shielding techniques are also under study.

The functionality proposed and described has been implemented on different chips.

The first has been designed in AMS 1.2 μm technology with digital input and output, and successfully operated at a frequency of 50 MHz.

Present efforts are devoted toward hard rad implementations: a matrix of pixels, designed in DMILL technology, including analogue front end and a complete digital read out system, was sent to the foundry in May 1994; it will be tested in November. An identical chip is under design in HSOI3HD (THOMSON) technology.

2. General principles

In this system there are 2 different parts:

- the active zone of the chip, organised in columns of pixels,

- the periphery of the chip (non-active for physics) containing decoders and a memory to perform the time tracking and data selection.

(a) Function of a column:

In a column, each pixel contains an individual 8-bit register, (called the "pixel register"). When a pixel is hit, a control block, integrated in the pixel, checks the status of the pixel register. If it is empty, the control block loads the address (8 bit) of the pixel into the pixel register. This value will be shifted, at the BCO frequency, along the column to the periphery.

In the present algorithm, if a pixel is hit at a time when there is an address in its 8 bit register from a previous hit, the new pixel address cannot be loaded since a priority is given to the descent. This new pixel hit is lost.

We have made physics simulations to evaluate the efficiency of this system, and propose some modifications to improve this efficiency.

An example of simulation with evolution of data is shown in figure 110.

- At the BCO(i) pixels 3 and 4 are hit; these pixels were empty, so pixel 4 loads its address 0100 (4) in its pixel register and pixel 3 loads 0011 (3) in its register.

- At the BCO(i+1) these values are shifted downward one step in the column and written into the pixel registers of pixels 3 and 2. Pixel 5 is hit in this BCO.
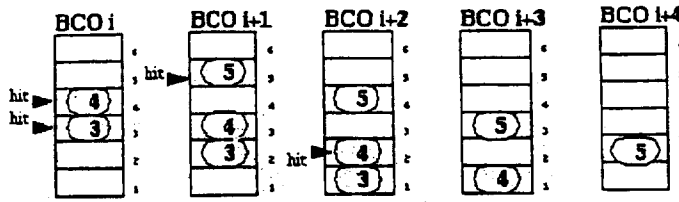
119

Figure 110: Example of evolution of data in a column.

- At the BCO(i+2) address 0100 (4) is in the pixel register 2, address 0011 (3) is in the pixel register 1; this last value is now present on the output of the column.
- At this BCO(i+2), pixel 2 is also hit; but its pixel register is full (address 0100) in memory. So this hit is lost.
- At next BCO, value 0100 is present to the output.

Here we note an important property of this organisation of columnar shift registers: the latency time of a pixel in the column is equal to its address in the column (see fig. 110. E.g., the address of pixel 1 will take 1 BCO to reach the bottom of the column, the pixel 2 will take 2 BCOs ... and the pixel N will take N BCOs. This feature helps us to time-track any pixel.

Now we examine the function of the peripheral memory which processes non-zero pixel address values presented at the outputs of columns.

(b) Function of the peripheral memory:

This part of the chip has the following functions:

- to record all non zero pixel addresses presented at the outputs of the columns during the trigger latency ($\approx 2 \ \mu s$);
- timing tracking of all hit pixels;
- output of the addresses of pixels hit in a BCO with a valid level 1 trigger. These are the "good hits".

When an address of a hit pixel is presented at the bottom of a column, its value is recorded in an 8 bit register. In parallel, a counter, synchronised to the BCO clock, totalises the latency time in the memory. With this we have all information necessary to decode "good hits" corresponding to BCOs with a valid level 1 trigger.

This decoding algorithm is very simple:

- value of address is equal to the latency time in column,
- value of counter indicates the latency time in memory.

The sum of these two values gives the time interval since the pixel was hit. The algorithm compares this result to the value of trigger latency; if they are equal, this address corresponds to a "good hit".

In order to be able to record pixel hits arriving during the formation of a level 1 trigger, it is necessary to have depth in the peripheral memory. From physics

120

simulations we will determine the optimum depth. as a function of the expected column occupancy, in order to lose minimal data.
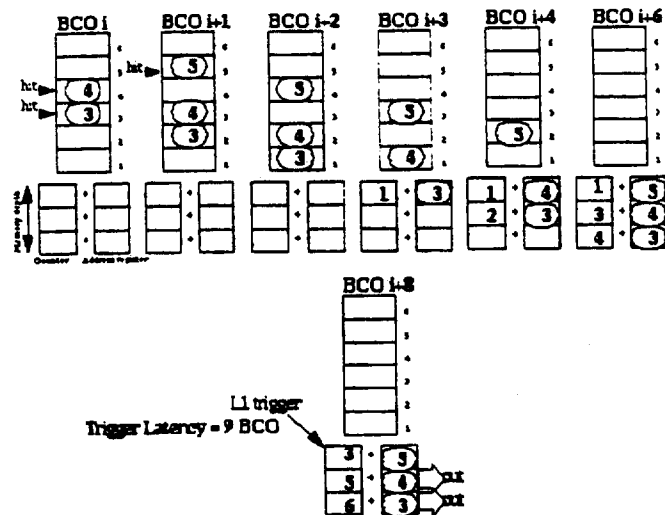

Figure 111: Data evolution in a complete system.

Following the same example, pixels 3 and 4 are hit at the BCO(i) and pixel 5 is hit at the BCO(i+1).

At BCO(i+3), the address of Pix3 enters the memory, and, in parallel, a counter is started; at BCO(i+4) the address of Pix4 enters the memory and a new counter is started; the counter assigned to address 3 is incremented.

At the BCO(i+8) the chip receives a L1 trigger. In this example, the trigger latency has been fixed to 9 BCOs; in reality it will be $\approx$ 80 BCOs (2 μs).

The algorithm adds the value of the pixel addresses with their associated counter values. In our example $4 + 5 = 9$ and $3 + 6 = 9$ but $5 + 3 \neq 9$ so only Pix4 and Pix3 have been hit at the right time.

We note an interesting property of the algorithm: it can accommodate changes in the level 1 trigger latency in a wide band around the nominal 2 μs.

We have implemented these algorithms electronically.

3. Schematic electrical and layout implementation

Each pixel contains (fig. 112):

- 8 × 1-bit registers in parallel driven by a single clock synchronised with the BCO,

- an 8 bit address, wire-implemented, to identify each pixel in a column,

- and a control block.

We now discuss the electronic implementations made or planned:

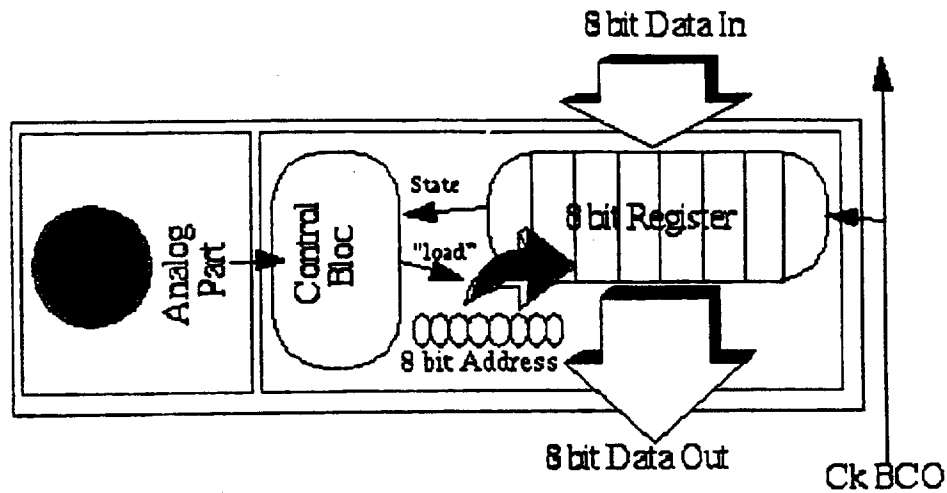- test results of the CPPM-AMS1 chip,

121

Figure 112: Composition of a pixel.

- evaluation of power dissipation,
- analogue-digital interface,
- shielding techniques.

A first test chip was implemented in AMS 1.2 μm technology (fig. 113). It includes two columns of 4 pixels with digital inputs and outputs. This chip was successfully tested on a Tektronik DAS9200 environment up to 50 MHz with a 3.7 V power supply. DAS9200 chronograms are presented in figure 114.

Simulations have been made to evaluate power dissipation of this system: in this version, each pixel dissipates 82 μW at 3 V at 40 MHz.

We have developed a second design (fig. 115) with reduced power consumption. This has been implemented in the DMILL and HSOI3HD technologies. In the AMS design, each pixel contained an inverter (the pixel driver), and power dissipation was shared between this driver and the column clock driver. Electrical simulations showed a 72 μW power dissipation in each pixel driver and 10 μW power dissipation in the column driver. Due to the very low occupancy in column (in our physics application), the pixel driver wasted a lot of power by often replacing a nil value by a new nil value.

In the new design, we send a clock only to pixels which need it (the ones which must shift data or are hit). In this case, physics simulations show that a pixel is active only 1 time over 250 (when the chip is hit). So the system proposed will divide power dissipation of the pixel driver by 250 and the pixel drivers will dissipate an average of only 0.3 μW per pixel. With this new design, we hope to achieve an acceptable digital power dissipation of 10 μW per pixel.

The interface between the analogue and digital parts of the pixel presents another interesting characteristic: the digital part is able to automatically "kill" "stuck-at-1 pixels" (fig. 116). When a pixel is hit, the analogue part generates an output signal
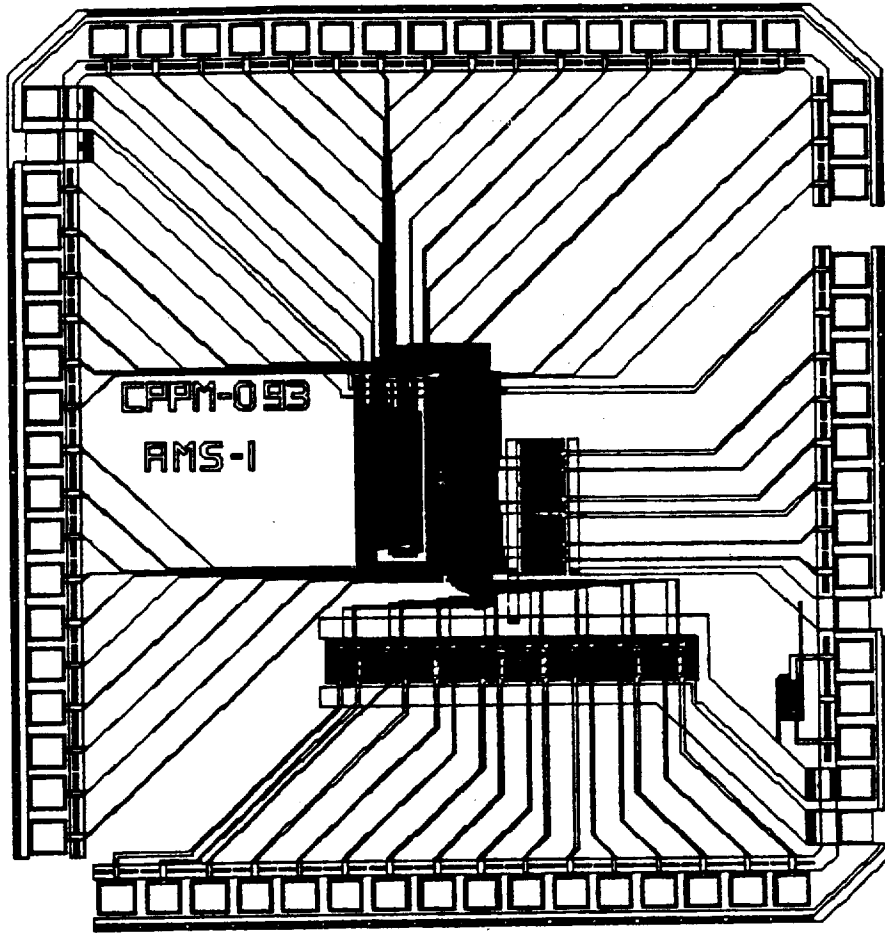
Figure 113: Layout of the AMS chip.

(1) active for 200 or 300 ns. The digital control block generates a pulse, synchronised with the leading edge of this output pulse (2), which drives the loading command of the pixel register with the pixel address. A system integrated in the digital control block checks that a new "loading pulse" can be generated only after a trailing edge of the analogue signal (3). In this way, a "stuck-at-1" pixel will send its address the first time and will be automatically "killed" after.

The digital readout system requires shielding to eliminate cross-talk between the electronic chip and the detector chip. To resolve this problem, we have designed the digital functionality using only poly-silicon and the aluminium 1 layer. We use the aluminium 2 layer, as an overall shield to cover the digital circuitry.

Figure 117 shows the layout of the digital part of the pixel; light and bold traces are the metal 1 and poly-silicon lines. They implement the digital gates. Grey hatching represented the metal 2 shielding.

The present best pixel (analogue + digital) size is 260 μm $\times$ 50 μm in DMILL technology with a 4 bit pixel's register. Presently the digital cell alone with an 8 bit pixel register is 180 μm $\times$ 50 μm; in future full pixel designs we hope to reach 150 μm $\times$ 50
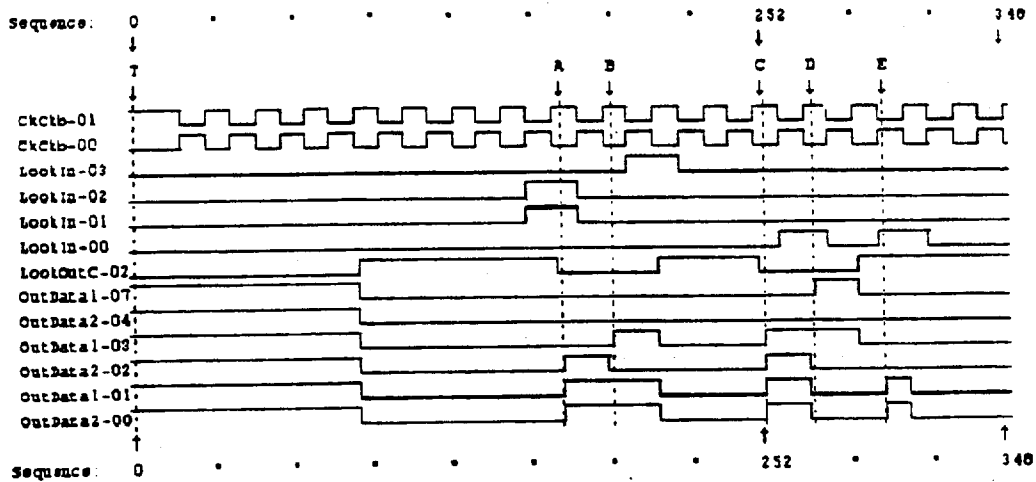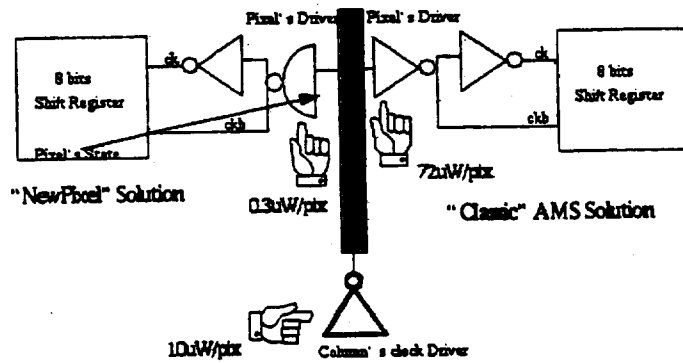
Figure 114: DAS9200 test results.

Figure 115: Power dissipation places for a pixel; old and new design.

µm for this part.

A readout in the HSOI3HD 1.2 µm technology is under design. We eventually wish to achieve 400 µm × 50 µm for the total cell (bump bond pad + analogue front-end + digital cell).

We have presented the decoding algorithm which is implemented in the peripheral memory of the chip. Now, we define the electrical structures necessary to implement this algorithm. Each hit recorded needs:

- an 8 bit address register,
- a counter function,
- an adder function,
- a comparator function.

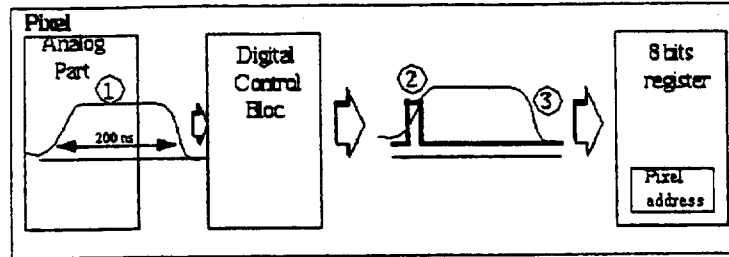This decoding structure is shown in fig 118.
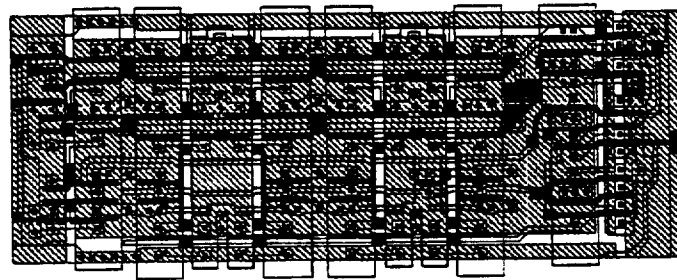
124

Figure 116: Analogue-digital interface.

Figure 117: Shielding technique over pixel's design.

A special counter has been designed which starts from the value of the pixel address. So, when an event arrives to the bottom of the column and enters the decoding zone, addition is directly made in this counter. When a level 1 trigger is sent to the chip, this value is directly compared to the trigger latency. If they are equal, value recorded in the address register is sent to the output of the chip. With this method, we have not needed to implement adders and have economised on the silicon surface.

In order to be able to record all events during the trigger latency this structure is organised as a FIFO: the pixel address and running time counter move down the FIFO by one step on the arrival of the address of the next hit pixel. This FIFO is long enough to lose minimal data at the expected column occupancy.

All these functionalities previously described have been implemented in DMILL and HSOI3HD technologies. They will be tested and measured in the DAS9200 environment with digital I/O, and in a real environment with the analogue part. Special structures have been designed to characterise analogue-digital and digital-detector cross-talk.

4. Physics simulations and technical improvements

In the present design, if a pixel is hit while its shift register contains a valid address, the value of the new address is not recorded. We have done physics simulations to evaluate the loss of data in the system proposed.

To improve the efficiency, we have proposed modifications to the design of the present system. Two solutions are under investigation:
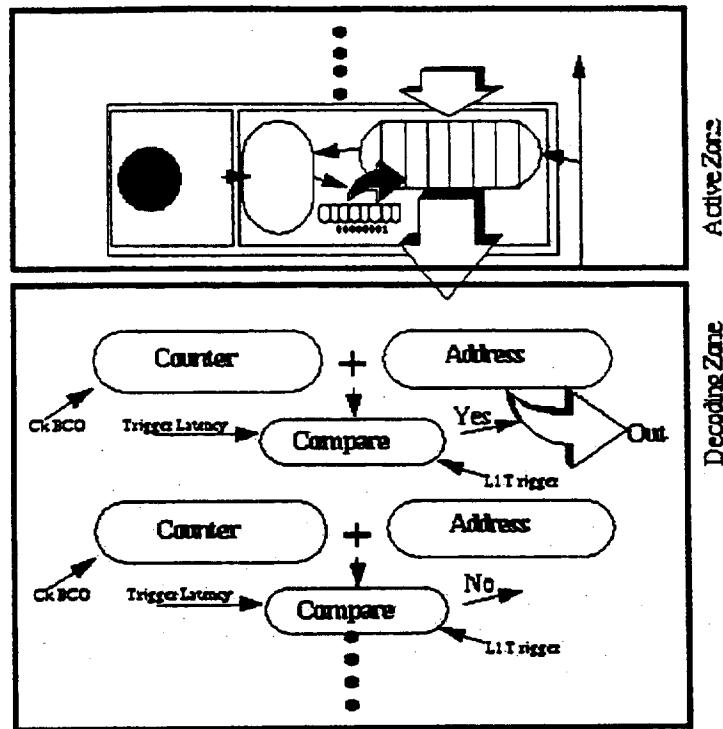
125

Figure 118: Architecture of decoding part.

The first exploits the possible use of the register of a pixel located downstream in the column. Since a charged particle can hit 2 or 3 pixels, it must skip at least two pixels and store in the third (fig. 119). When this loading possibility is used, an alarm bit is activated. This alarm bit says that we must add +3 to restore the good address or to start the counter with (-3) value. An example of operation is shown in the fig 119.

The second solution has 4 shorter columns clocked in parallel, with pixel (i) chained to pixel (i-4) rather than pixel (i-1) (fig. 120).

In the example, pixels 16, 12, 8, 4 can be considered one of four independent columns.

The efficiency of this system would be naturally increased by a factor 4; in reality we can expect better results because this solution reduces the multiple hit rate in a same column.

These two solutions impose an interlacing of pixels, and so require the use of the metal 2 layers to drive signals over pixels. To add shielding over the pixels, a three metal technology would probably be necessary.

5. Conclusion

We have proposed an algorithm for a read out system. Modularity, robustness, and interesting electrical characteristics are principal features of this system. The two matrices of pixels with analogue front-end and digital read-out which have been implemented in DMILL (CEA/LETI) 109 and HSOI3HD (THOMSON) technologies will guide us for the development of a real scale demonstrator. Parallel efforts will be made
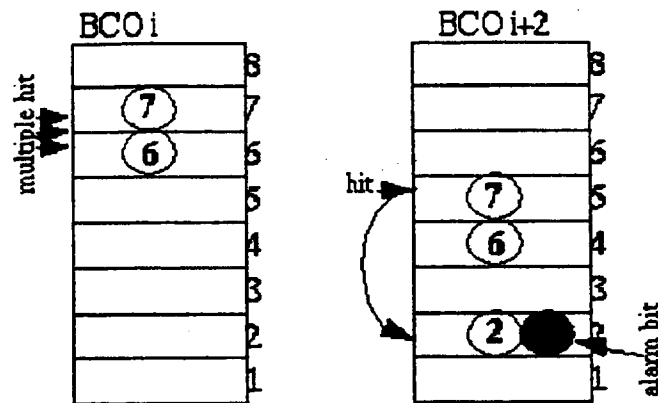
126

Figure 119: (-3) escape loading technique.

to improve efficiency, electrical characteristics and layout size of this system. Finally, future efforts will be also directed to define an acceptable data format for chip output (channel number, speed output ...), in conformity with the ATLAS DAQ system.

3.8.3 Readout architectures under study at CERN and Genoa.

1. **Abstract**

The proposed CERN-Genoa readout architecture, like all the other under study, is a column architecture. The basic idea is the correlation between a precise spatial information of the hit stored inside a pixel with a precise time information at the column level. The two informations together allow, with rather small ambiguity, the precise assignment of each hit of the event to the correct beam cross over (BCO). The main requirements we have considered for the proposed readout architecture are:

- very simple pixel cell for low power consumption, small cell dimension and large yield;

- no clock and minimum possible digital activity in the pixel cells when their analog part is active;

- data driven design as far as possible: data rate for a pixel is much lower than a 40 MHz clock, i.e. BCO rate;

- regular structure: it helps the testability of a large system and reduces the consequences of the propagation of faults from a single pixel, column or chip to larger regions;

- evolution from a proven architecture: the OMEGA2 design.

In the next few pages we will describe the principle of the architecture, we will discuss about events/hits lost because of dead time and finally we will illustrate the steps and tests we want to do in the years 1995-96 to arrive at a working prototype, suitable for Atlas.
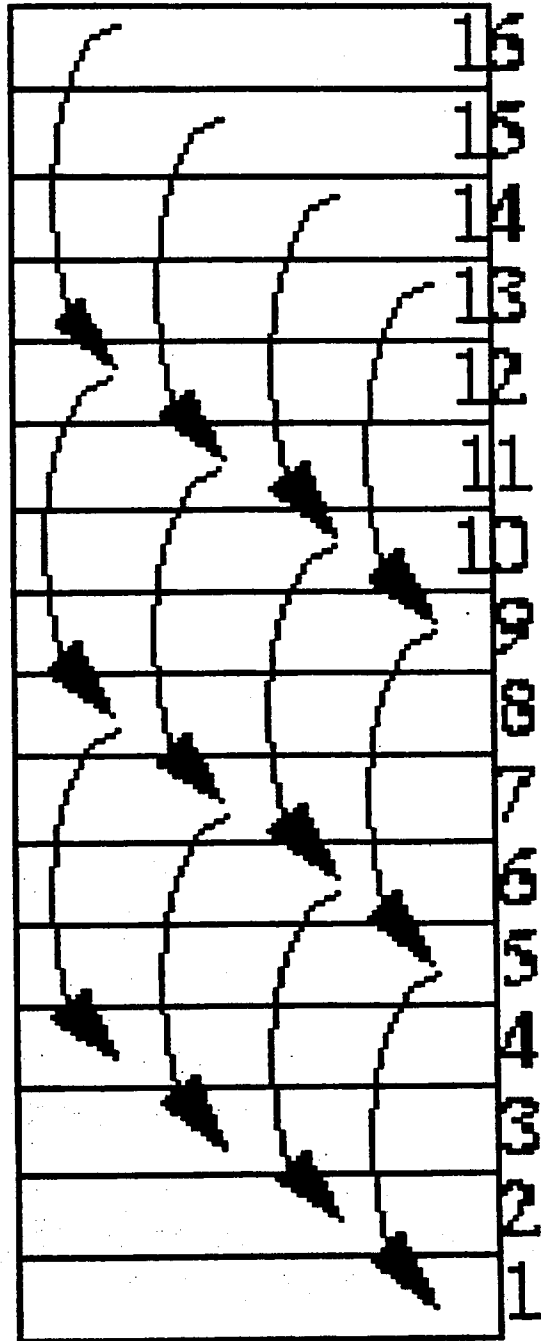
127

Figure 120: Parallelism technique.

128

2. Principle of the readout architecture.

The three main requirements for a front-end and readout architecture suitable to AT-LAS are:

- Time resolution: the event hits in the selected BCO must be unambiguosly resolved:

- L1 trigger latency: a record of the last $2\mu s$ (i.e. 80 BCO's) must be kept in each cell front-end;

- low dead-time: no more than 50 ns (2 BCO's) must be lost once an event is selected by L1.

This readout architecture follows from the design of the OMEGA2 [29, 30. 32] which has been successfully implemented in a Omega experiment and is shown schematically in fig. 121. The pixel cell contains a fast charge amplifier, a comparator with an adjustable threshold followed by a precision delay and coincidence logic and storage. To adapt this architecture to the needs of Atlas requires the introduction of a fast column time stamp.
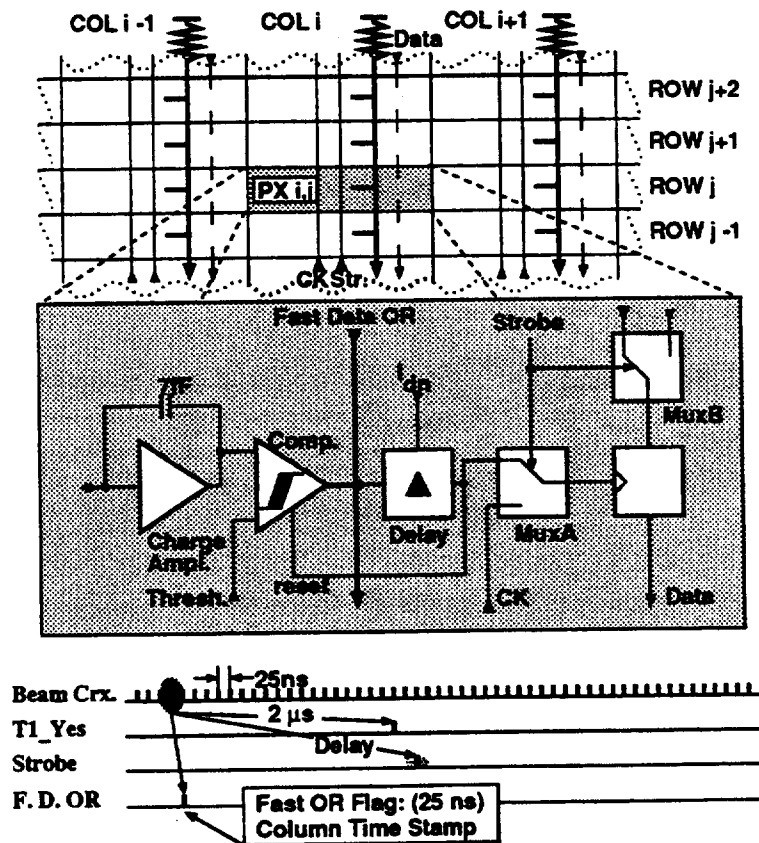

Figure 121: The Omega pixel front-end cell with the addition of a "Fast-OR" line to implement the proposed Atlas read-out architecture.

The pulse from the comparator is delayed by an adjustable time set equal to the L1 latency. The comparator is latched and is reset by an intermediate tap in the delay logic. A L1 trigger generates "strobe" pulses to all the pixels of the selected columns (see later) and, if a hit is present at the output of the delay, it is latched for later redout.

The fast column time stamp is formed from the Fast-OR (FO) of the signals from the comparator outputs in a column. This pulse is sent to a shift register at the end of the column fig. 122, which is clocked at the beam crossing frequency ($\sim 40MHz$) and whose depth (80x25 ns) is determined by the L1 latency ($2\mu s$).
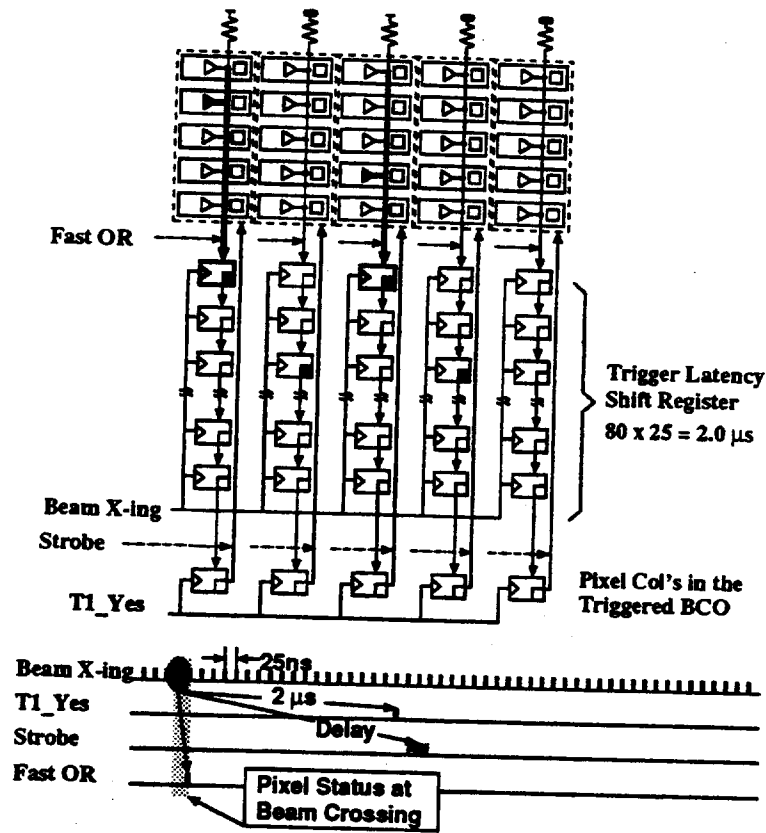

Figure 122: Schematic diagram of the proposed Atlas read-out architecture. The signal from two pixel cell discriminators in row 1 and 3 propagate to the periphery in less than 25 ns and enter the "trigger latency shift registers".

The outputs of the shift registers are transferred to a register that is synchronized to the L1 latency (see fig. 123).

A hit in this register means that one or more pixels in the column were hit in time with the BCO of interest. The strobe signal is sent back to such columns in time to latch the delay outputs of those pixels fired by a particle in time with the BCO of interest. The L1's initiates also the readout of the columns with hits in time with the triggered BCO. Hit pixels are clocked out at 40 MHz (ie. $3\mu s$ for 128 chained pixels
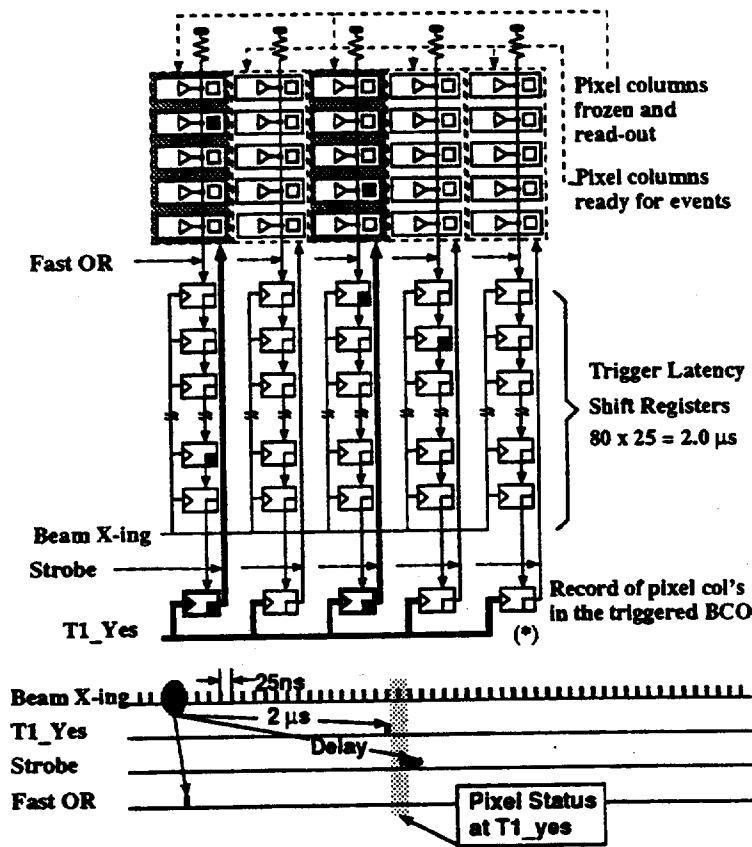
Figure 123: Schematic diagram of the proposed Atlas read-out architecture. The first level trigger (L1_Yes) strobes "trigger latency shift registers" outputs.

in a column), while all other columns stay active waiting for new events.

While clocking out, each time a non zero word is detected, its hit pattern together with the row number in the pixel matrix are stored in a buffer at the periphery of the pixel chip as is shown in fig. 124.

The above circuitry of the chip periphery is reproduced several times to permit the readout of the new event, while other events are waiting for being transferred out of the chip. Finally the data once structured in an event buffer at the chip periphery is transmitted asynchronously with a serial protocol to one or more chip controllers on the detector ladder (see fig. 124 down).

3. Architectural implications on performances.

A critical aspect of this design is the circuitry that must produce a uniform delay for all pixels in an array. It is expected that a precision of about 5% can be achieved in the delay circuitry, which implies a strobe width of $100 \div 125$ ns. Since the column occupancy is $\simeq 1\%$, the additional noise hits, due to $\simeq 5\%$ superimposed events only in the strobed column, are expected to be acceptable. Assuming a uniform spatial distribution of the hits, from the binomial distribution one obtains :
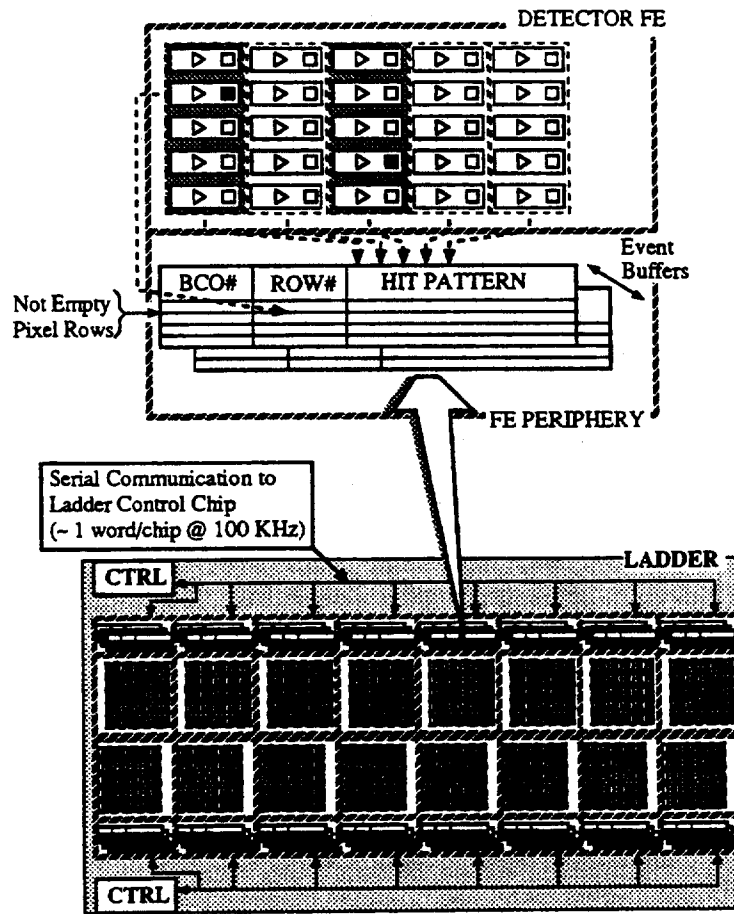
131

Figure 124: Structure of the front-end pixel chips and their layout to form a Atlas detector layer.

$$N_n \sim N_B \times O \sim 5 \times 0.01 \sim 0.05$$

where N_n is the number of noisy hits in a readout column, N_B is the number of integrated BCO's and O is the column occupancy. The sources of inefficiencies of the system come from the dead time of critical circuits in the front-end. The two most critical sources of the dead time are

- The lost or recovery time

- The clocking out of the columns when they are readout;

The first source of inefficiency does not seem to be critical, since with a 100 kHz L1 gate, $3/mus$ of column readout and 1% of column occupancy, a very safe value of 3% of total inefficiency is well acceptable. More critical is the second source which causes, with a column occupancy of 1% , a (n-1)% inefficiency, where n is the length of the FO pulse in units of 25 ns. This implies that either a very fast recovery of the FO is achievable or several FO's in a column are used, each serving a lower number of pixels.

The sizes of the event buffers on the chip periphery, and their total number must be better studied by simulation. From simple considerations on the chip occupancy (less than 1 bit/chip on average) and data rates in input and output 4 event buffer with $10 \div 20$ words each are fully satisfactory for the system. Serial transmission of data out of chip is preferred due to the low occupancy: an average of 400 bits can be transferred per event between two L1 triggers ($10\mu s$)

4. **Planning for prototype development.**

A test version (Omega 3) of this architecture will be built in Faselec [29] $1\mu m$ CMOS technology. This chip will implement the whole functionality needed for Atlas at the level of the pixel cell, while only limited functionality foreseen for Atlas will be included in the chip periphery.

Omega3 coupled with a control chip that will be built in ES2 $1\mu m$ standard cell, will permit to test the critical analog elements (charge amplifier and comparator) togheter with the interface to the digital circuitry (delay, FO) and the influence of the clock on columns next to the active ones. For 1996 and 1997 the full functionality needed in Atlas will be included in the new versions of the front-end chip: multievent buffer and serial communications. During the same time the ladder control chip will be better defined and new prototypes designed.

3.9 Dataflow studies

A detailed study of the data flow from the pixel read-out system has been performed. The aim of this study is to establish the amount of data coming, not only from the pixel detector, but also from the various subset (the column, the chip, the module and the read-out unit) which are relevant for the pixel read-out architecture.The read-out unit is defined as half ladder (also called a stave) in case of the barrel detectors and as four wedges in case of the disk detectors. The rationale behind this definition is that any read-out unit should deal with the same data flow. These studies have been done in the context of the LBL readout design, but are applicable to other designs and readout chip sizes.

The details of the detector model used in the simulations described in this section are summarized in Table 18 so that extrapolations can be made to other layouts.

3.9.1 The Physics Models

The following simulation is performed. First, the charged tracks produced by PYTHIA are tracked out through a homogeneous magnetic field of 2.0 T. The tracking proceeds only through the pixel system defined above. No attempt has been made to simulate looping particles, etc. as such a simulation requires a realistic model for energy loss and material in the detector (at which point the proper step is to carry out the full GEANT simulation in DICE and study the resulting occupancies – which is the next logical step for the present study). In order to convert from the number of charged tracks traversing some part of the pixel detector, to the number of hit pixels expected, the following additional "corrections" will be applied:

133

Table 18: A summary of the relevant geometric parameters of the AT-LAS pixel detector used in the present simulation. The two barrel layers are referred to as SL0 and SL1. The four disk layers are referred to as EC0, EC1, EC2, and EC3. The optional b-layer is referred to as BL0. The active area of each pixel chip is taken to be roughly 1.0 cm (32 by 300 μ) by 1.0 cm (192 by 50 μ).

Parameter	Values		
Radius/Half-Length BL0	4.90	36.25	16 Ladders
Radius/Half-Length SL0	11.03	36.25	36 Ladders
Radius/Half-Length SL1	15.93	42.30	52 Ladders
Radii/Z Location EC0	11.0 – 20.9	52.2	144 Wedges
Radii/Z Location EC1	11.0 – 20.9	59.2	144 Wedges
Radii/Z Location EC2	11.0 – 20.9	77.2	144 Wedges
Radii/Z Location EC3	11.0 – 20.9	85.0	144 Wedges

1. In order to correct for charge sharing (that is the charge from a single charged track producing several hit pixels), a factor of 2.5 has been used. This factor has been taken from the CDF SVX detector experience, based on a 30 cm long detector consisting of 60 μ pitch strips on a 300 μ thick substrate, and hence should be a relatively good estimate of what we would expect from a pixel system (further sophistication is perhaps warranted here).

2. A factor of two is used to correct for additional hits NOT associated with charged tracks, but not due to electronic noise, that is resulting from the presence of additional soft photons, etc. in the tracking volume. This figure is also based on CDF experience.

3. Finally, a factor of two is used to correct for the presence of looping tracks. This is a global factor which can be reasonably applied to uniform events such as MinBias, and is based on SDC simulation experience with a similar 2T magnetic field.

The second and third factor above can be more accurately determined by a GEANT simulation, but the overall factor of ten above is probably accurate to about a factor of two, and hence provides a reasonable starting point.

3.9.2 Data Samples

In the present discussion, we have chosen to study the expected dataflow using several data samples intended to capture the possible variations in both local and global event complexity. The samples which have been chosen were:

- PYTHIA MinBias events at luminosities of both 10^{33} cm^{-2}sec^{-1} and 10^{34} cm^{-2}sec^{-1}.

- PYTHIA dijet events with $P_t(jet)$ greater than 1 TeV (a sample concentrated at $\eta \leq$ 1.5 and hence covering only the barrel region).

134

- PYTHIA Dijet events with $P_t(jet)$ greater than 0.5 TeV and a dijet mass cut of 5 TeV in order to preferentially select only those jet pairs at large values of η.

Samples of 1000 events of each type were generated, and the results for average and worst-case occupancies were tabulated for all levels in the detector which are relevant for DAQ studies (presently the column, the chip, the module, and the readout unit). Further studies were carried out in which the integrated occupancy for an entire Level 1 Latency period (i.e. a period of 80 crossings) was accumulated for each detector element. Such studies are particularly relevant for understanding the required buffering, prior to Level 1 filtering, which is needed in the End of Column logic.

In the present simulation, PYTHIA itself directly produces the additional pileup events, resulting in a mean number of interactions of roughly 2.5 per crossing at 10^{33} and 17.5 at 10^{34}, giving about a factor of 7 difference in MinBias multiplicity between the two cases (the difference arises because of the truncated Poisson that is generated at the lower luminosity, since the present sample always requires at least one inelastic interaction per crossing). These values are in reasonable agreement with those used elsewhere in ATLAS simulations.

3.9.3 Occupancy Results

In order to capture average properties of the events as well as the effects of fluctuations within events, two basic quantities were calculated. The first is the average occupancy (referred to as AVG in the figures), which is just the number of tracks summed over the full detector and divided by the number of elements of the proper type. The second is the "worst-case" occupancy for each event, defined as the maximum occupancy for a given entity in an event (for example, the chip in the first barrel cylinder with the highest occupancy). It is referred to as MAX for the highest occupancy element, and MAX2 for the second highest occupancy element. Note that MAX2 is defined so that it must be less than MAX, which is a significant constraint for small integers .

As an indication of the relatively uniform distribution of occupancy throughout the detector, Fig. 125 shows the mean maximum value of the chip occupancy, plotted versus the psuedorapidity of the chip. The variations in this worst-case occupancy are relatively modest as one traverses the different cylinders and disks, with the exception of the optional b-physics layer which suffers substantially greater occupancies.

The mean values of these occupancies (averaged over the 1000 event samples of the present study) are tabulated in Table 19–20. Some representative distributions of these variables are displayed in Figs. 126–135, where the first series is for MinBias events, and the second is for high-P_t dijets which display much larger fluctuations in the local occupancy than do the MinBias events. Note that the actual plots are included only for the central 1 TeV dijet event sample in which most of the jets pass through the BL and SL detector layers. The behavior of the disk EC layers when the more forward jet sample with the dijet mass cut is used is similar to that observed in the barrel layer plots shown here.

These tables and figures contain occupancies computed using the number of charged tracks traversing detector elements. To convert these values to occupancies for the number of hit pixels, the factors discussed in Section 3.9.1 must be applied. It is plausible to apply the complete factor of 10 to the average quantities, as they are global sums over the occupancies in large areas of the detector. For the worst-case calculations, it is more reasonable to

135

apply only the first factor of 2.5, as the other factors should not significantly change local occupancy fluctuations. Finally, for the MultiEvent quantities, the situation is still more subtle, as what is most significant in the pixel architecture is the number of different beam-crossings which have hit pixels. Assuming that the probability of multiple hits in one column in one beam-crossing can be neglected, then it seems reasonable to apply a scale factor of 1 to this case – the major assumption here is that the information from a charged track is collected in one crossing time of 25 nsec.

Note that throughout this study, we have assumed that only the data from a single crossing is relevant. This requires collecting all of the relevant charge information within 25 nsec. While the present LBL-4 pixel unit cell is capable of doing this for a moderate range of pulse heights (roughly 1-8 fC, where one track should deposit 4 fC), the time-walk of the shaping circuitry will cause smaller pulse heights to appear later. If full sensitivity to any charge above threshold is required (that is charges above about 0.2 fC), in order to take advantage of the large S/N of the pixel array for centroid finding, then it is likely that hit information from two adjacent crossings will need to be read out, thereby doubling the average occupancy for MinBias events.

Finally, to gain further insight into the distribution of occupancies for typical events, histograms have been accumulated where the occupancy of each hit detector element is entered for each event. By examining normalized integrals of these distributions, it is possible to extract the fraction of the total data in an event which would be lost due to limitations of the DAQ system. Table 21–22 summarize the beahvior of these distributions, whereas Figs. 136–145 display the integral distrubutions, including values of n and p such that $P(x \geq n) = p$. Values are listed on the figure for the first case where the value of p was less than 10^{-m}, where m ranges from 1 to 5.

Figures 136–140 show the behavior for MinBias samples, showing the expected exponential falloff expected from Poisson behavior. Figures 141–145 show the large occupancy tail created by the presence of the dense high-P_t jets. The two cases shown in Fig. 140 and Fig. 145 are of special interest, as they can be used to derive the required buffer depth for the End of Column logic. If we assume that all the relevant pixels for a given charged track appear in the same beam-crossing, and that the probability for multiple hits in one beam-crossing is negligible (Fig. 136 suggests that this is accurate for the MinBias sample, Fig. 141 implies that it is breaking down for the high-P_t jet sample), then we can deduce the following requirements. In deriving these requirements, the idea would be to strive not for perfect readout of all events, but at least for limited inefficiency from this source even when in the presence of high-P_t jets.

- For the BL layer, 8–10 buffers are required.

- For the SL layers, 4–6 buffers are required.

- For the EC layers, 4–6 buffers are required.

In summary, the present LBL architecture, perhaps slightly expanded to six buffer locations, appears adequate for running up to 10^{34}. This buffering is also adequate for the BL layer provided it is not operated much above the 10^{33} luminosity for which it has been proposed. At higher luminosities, the large column occupancy would require a substantially deeper buffer requirement for this layer.
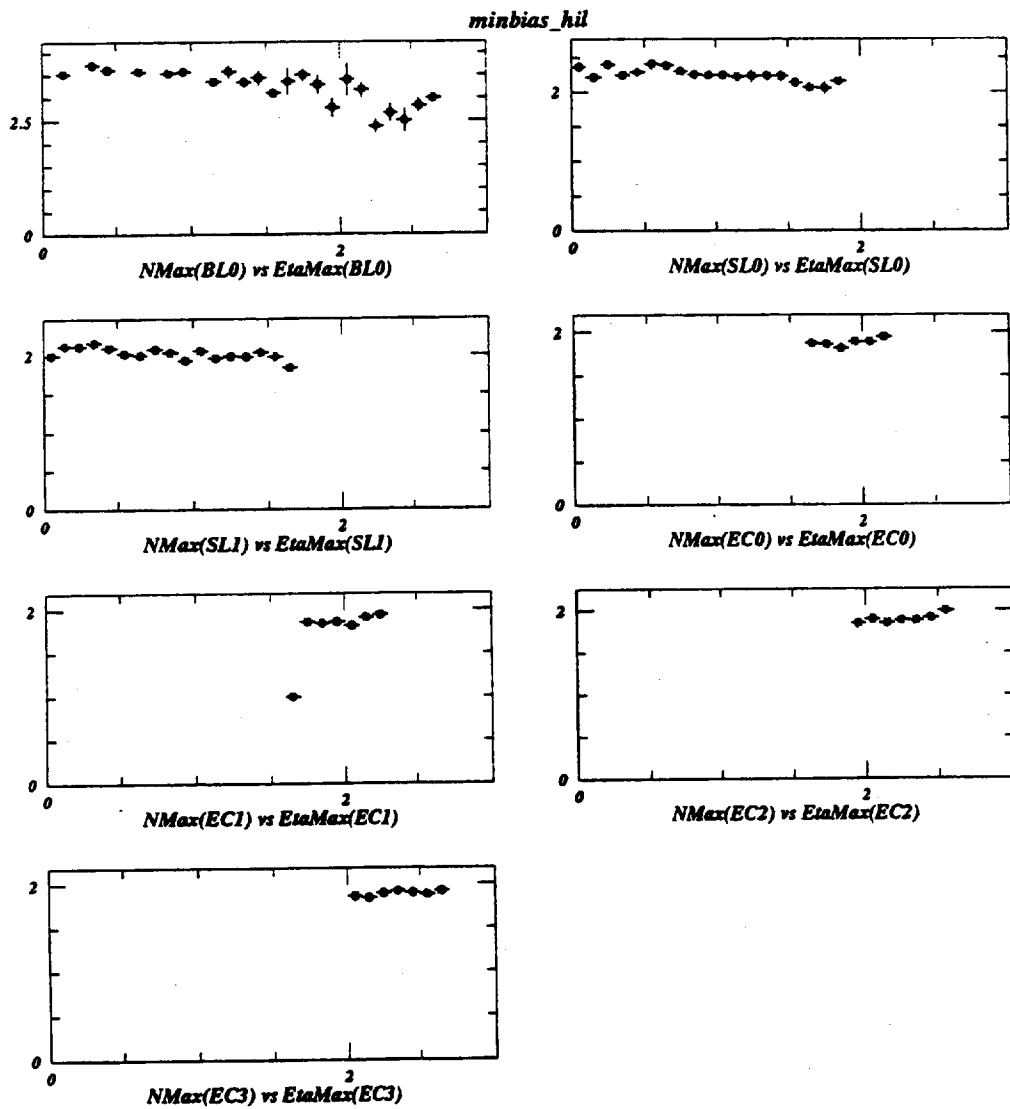
Figure 125: A series of plots for the various elements of the pixel system (three cylinders and four disks) showing the worst-case chip occupancy versus the pseudorapidity (η) of this chip for MinBias events at a luminosity of 10^{34}.

Table 19: A summary of the average values for pixel detector occupancies at each level in the DAQ system for the different event samples under study. All occupancies are quoted as the number of charged tracks per crossing, where a crossing will include the relevant number of MinBias pileup events for the specified luminosity. The left column of a pair is the average occupancy per crossing, and the right column is the worst-case occupancy per crossing (at least one hit is required, so this number is bounded below by 1). The final column, labeled MultiEvt, is the worst-case occupancy for a column after 80 crossings of data have been accumulated. Note for the barrels, the readout unit is a stave, whereas for the disks, it is a set of four wedges.

Layer	Column		Chip		Module		Readout Unit		MultiEvt
	MinBias at 10^{33}								
BL0	0.0009	1.06	0.029	1.6	0.35	3.2	2.1	5.7	3.2
SL0	0.0003	1.01	0.009	1.3	0.11	1.9	0.6	3.1	2.4
SL1	0.0001	1.00	0.005	1.1	0.06	1.6	0.4	2.5	2.1
EC0	0.0002	1.00	0.006	1.1	0.06	1.4	0.23	1.9	2.0
EC1	0.0002	1.00	0.006	1.1	0.06	1.4	0.23	1.9	2.1
EC2	0.0002	1.00	0.006	1.1	0.06	1.4	0.24	1.9	2.1
EC3	0.0002	1.00	0.006	1.1	0.06	1.4	0.24	1.9	2.1
	MinBias at 10^{34}								
BL0	0.006	1.8	0.20	3.5	2.41	10.7	14.4	23.5	7.7
SL0	0.002	1.3	0.06	2.3	0.73	4.8	4.4	10.2	4.2
SL1	0.001	1.1	0.03	2.0	0.38	3.6	2.7	7.7	3.5
EC0	0.001	1.1	0.04	1.9	0.40	3.1	1.59	5.4	3.4
EC1	0.001	1.1	0.04	1.9	0.40	3.1	1.61	5.4	3.5
EC2	0.001	1.1	0.04	1.9	0.41	3.1	1.65	5.4	3.4
EC3	0.001	1.1	0.04	1.9	0.41	3.1	1.66	5.4	3.5

Table 20: A summary of the average values for pixel detector occupancies at each level in the DAQ system for the different event samples under study. All occupancies are quoted as the number of charged tracks per crossing, where a crossing will include the relevant number of MinBias pileup events for the specified luminosity. The left column of a pair is the average occupancy per crossing, and the right column is the worst-case occupancy per crossing (at least one hit is required, so this number is bounded below by 1). The final column, labeled MultiEvt, is the worst-case occupancy for a column after 80 crossings of data have been accumulated. Note for the barrels, the readout unit is a stave, whereas for the disks, it is a set of four wedges.

| Layer	Column		Chip		Module		Readout Unit		MultiEvt
	Dijets with $P_t(jet) \geq 1$ TeV at 10^{34}								
BL0	0.008	4.0	0.26	15.1	3.1	29.5	18.8	42.8	12.1
SL0	0.003	2.7	0.08	10.0	1.0	18.0	6.0	24.3	6.0
SL1	0.001	2.3	0.05	8.3	0.5	14.8	3.8	19.8	5.3
EC0	0.001	1.1	0.05	2.3	0.45	3.7	1.8	6.8	3.7
EC1	0.001	1.1	0.05	2.1	0.45	3.6	1.8	6.5	3.6
EC2	0.001	1.1	0.05	2.1	0.45	3.5	1.8	6.1	3.6
EC3	0.001	1.1	0.05	2.0	0.45	3.4	1.8	6.0	3.6
	Dijets with $P_t(jet) \geq 0.5$ TeV at 10^{34}								
BL0	0.008	2.5	0.25	9.4	3.00	20.3	18.0	35.9	9.3
SL0	0.002	1.9	0.08	5.5	0.91	11.1	5.4	18.1	5.0
SL1	0.001	1.5	0.04	4.1	0.47	8.1	3.3	13.1	4.3
EC0	0.002	1.6	0.05	4.5	0.51	6.5	2.05	11.8	4.6
EC1	0.002	1.6	0.05	4.5	0.51	6.5	2.07	11.6	4.3
EC2	0.002	1.5	0.05	3.9	0.51	5.8	2.04	10.5	4.1
EC3	0.002	1.5	0.05	3.8	0.50	5.7	2.02	10.1	4.0
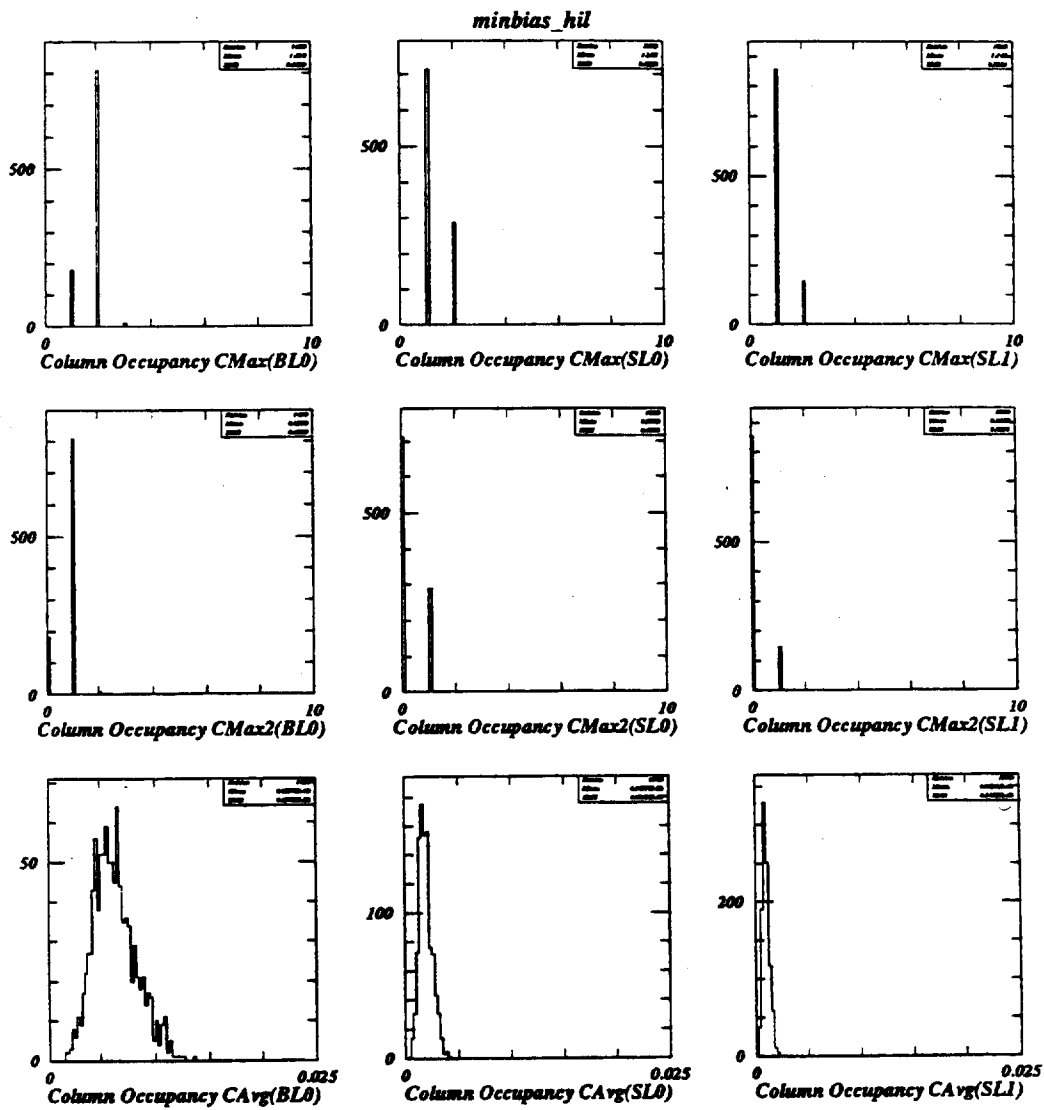
Figure 126: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) column occupancies for MinBias events at a luminosity of 10^{34}.
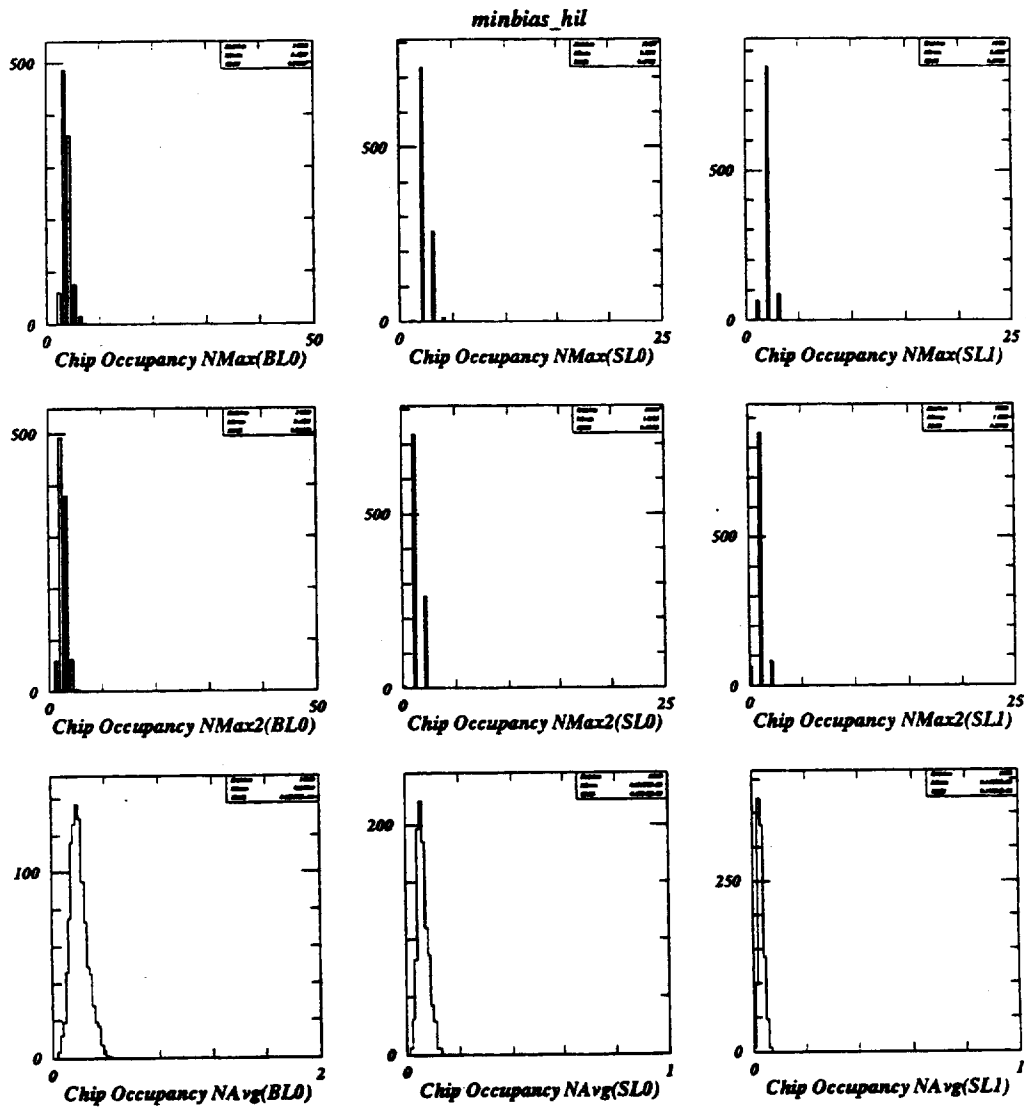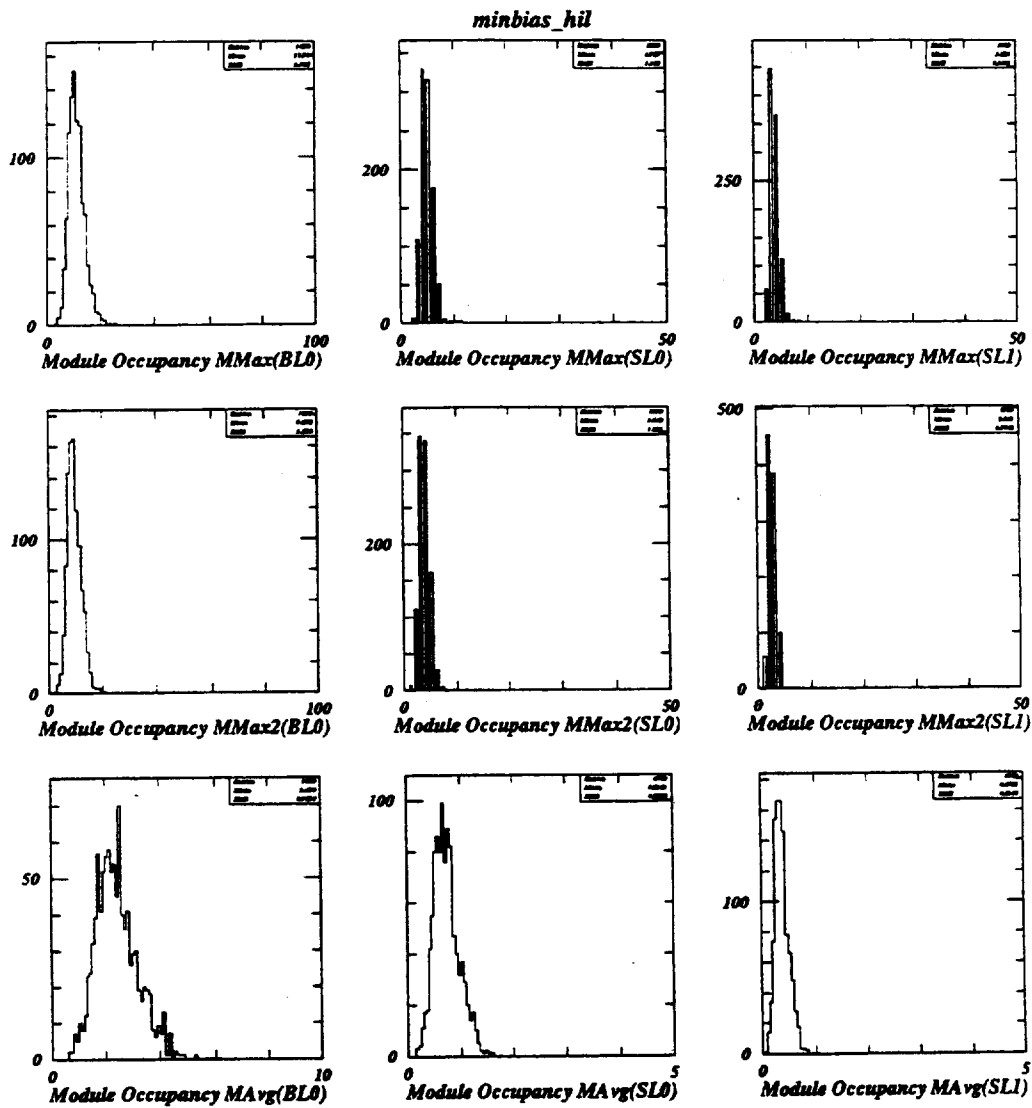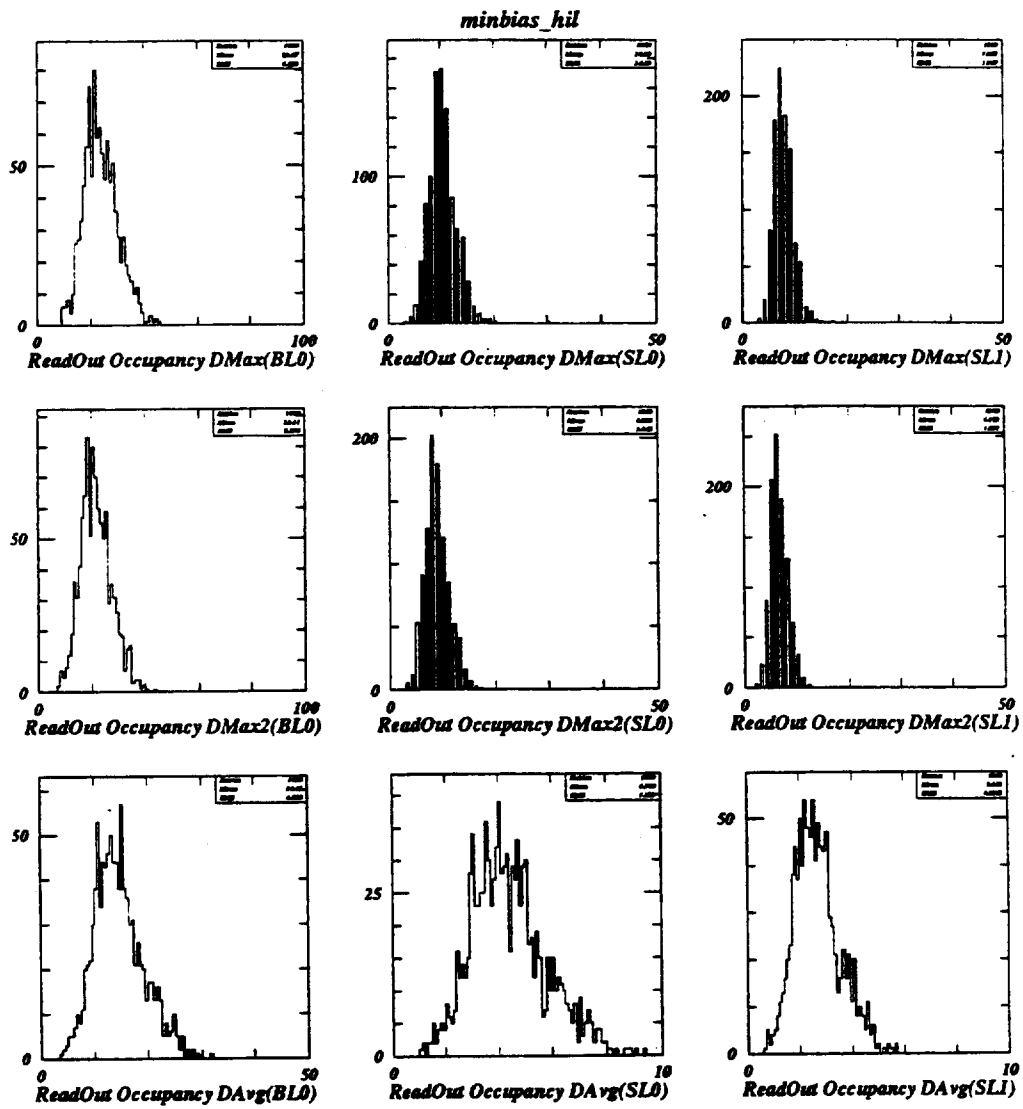
140

Figure 127: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) chip occupancies for MinBias events at a luminosity of 10^{34}.

Figure 128: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) module occupancies for MinBias events at a luminosity of 10^{34}.

Figure 129: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) readout unit occupancies for MinBias events at a luminosity of 10^{34}.
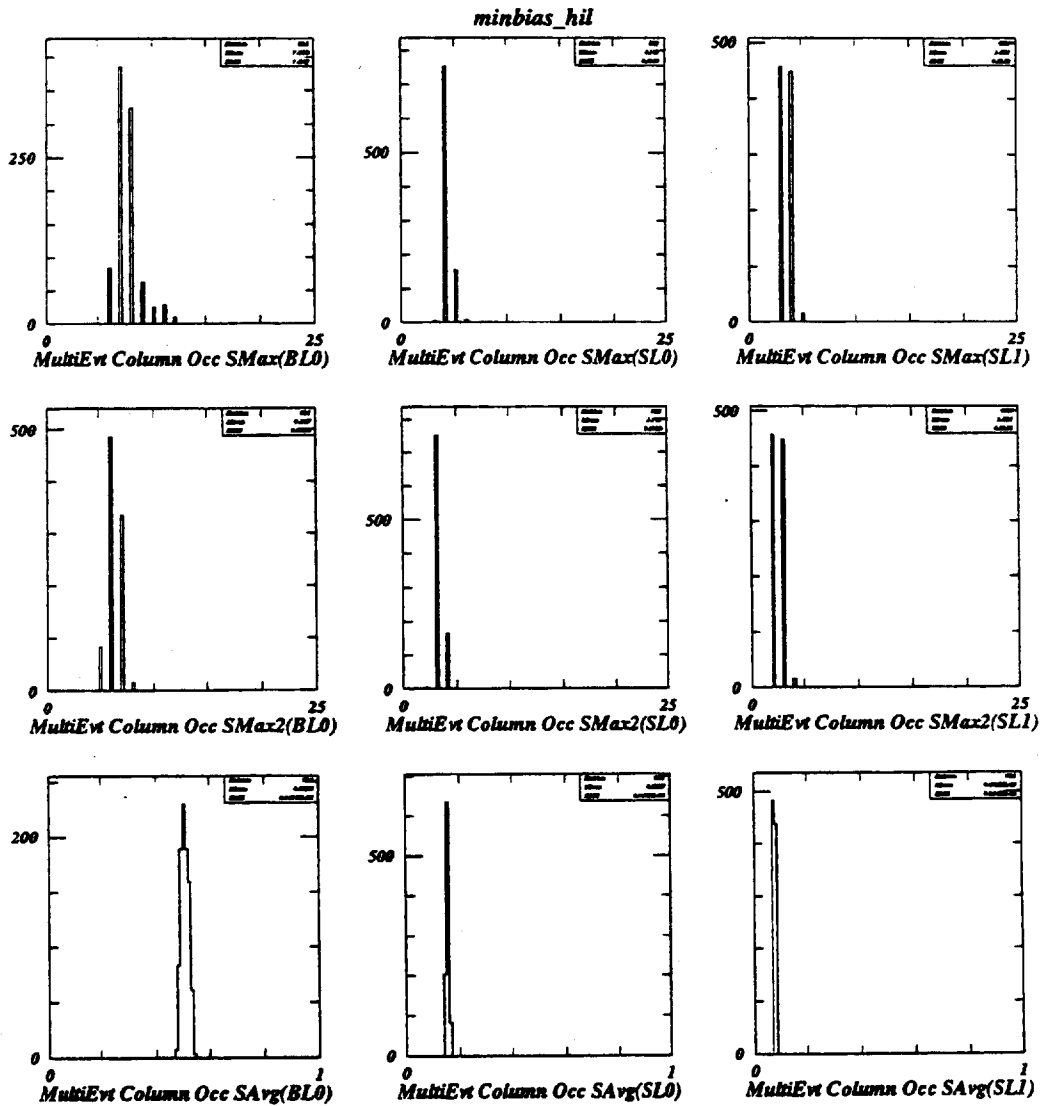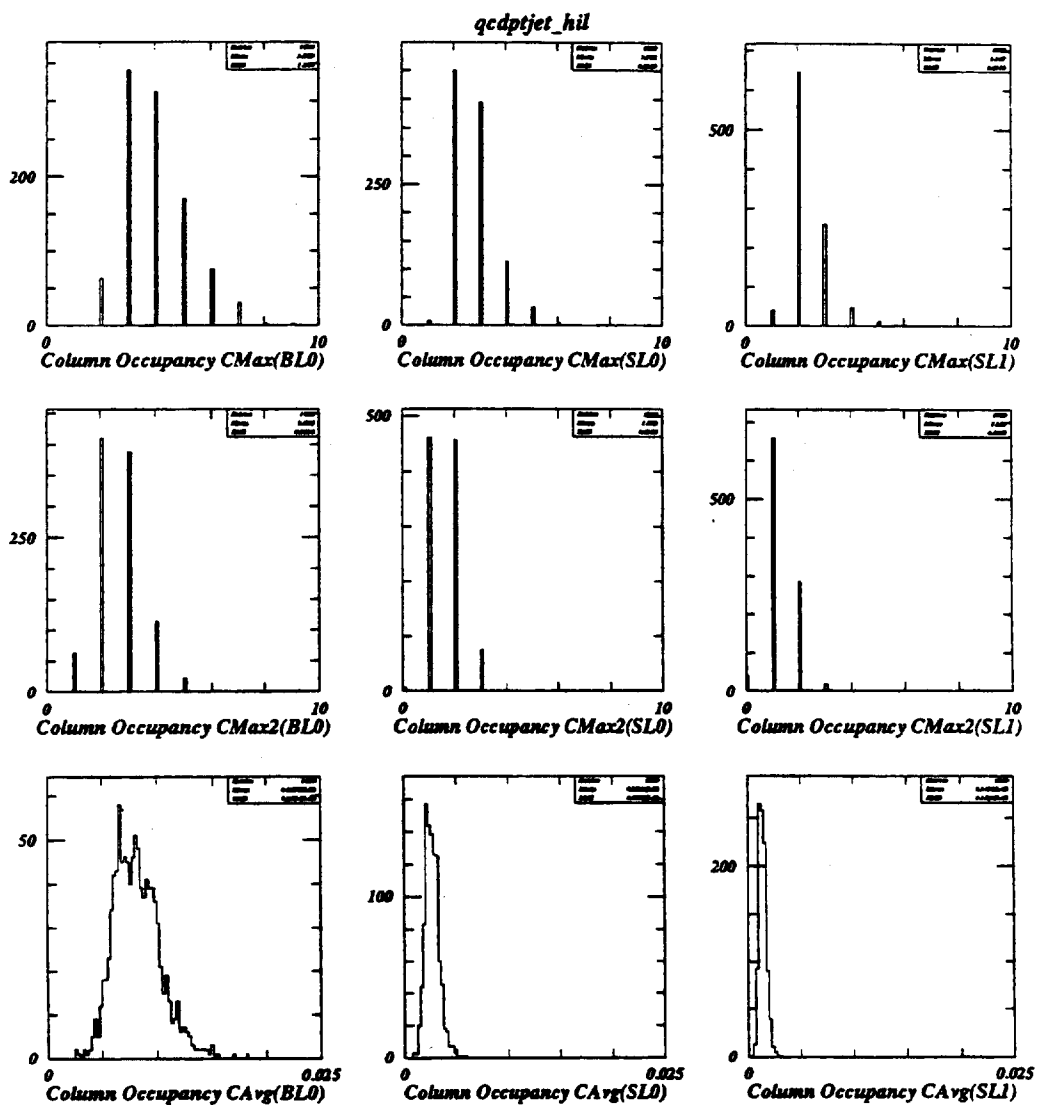
Figure 130: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) column occupancies accumulated over the Level 1 latency period of 80 beam-crossings for MinBias events at a luminosity of 10^{34}.

144

Figure 131: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) column occupancies for 1 TeV Dijet events at a luminosity of 10^{34}.
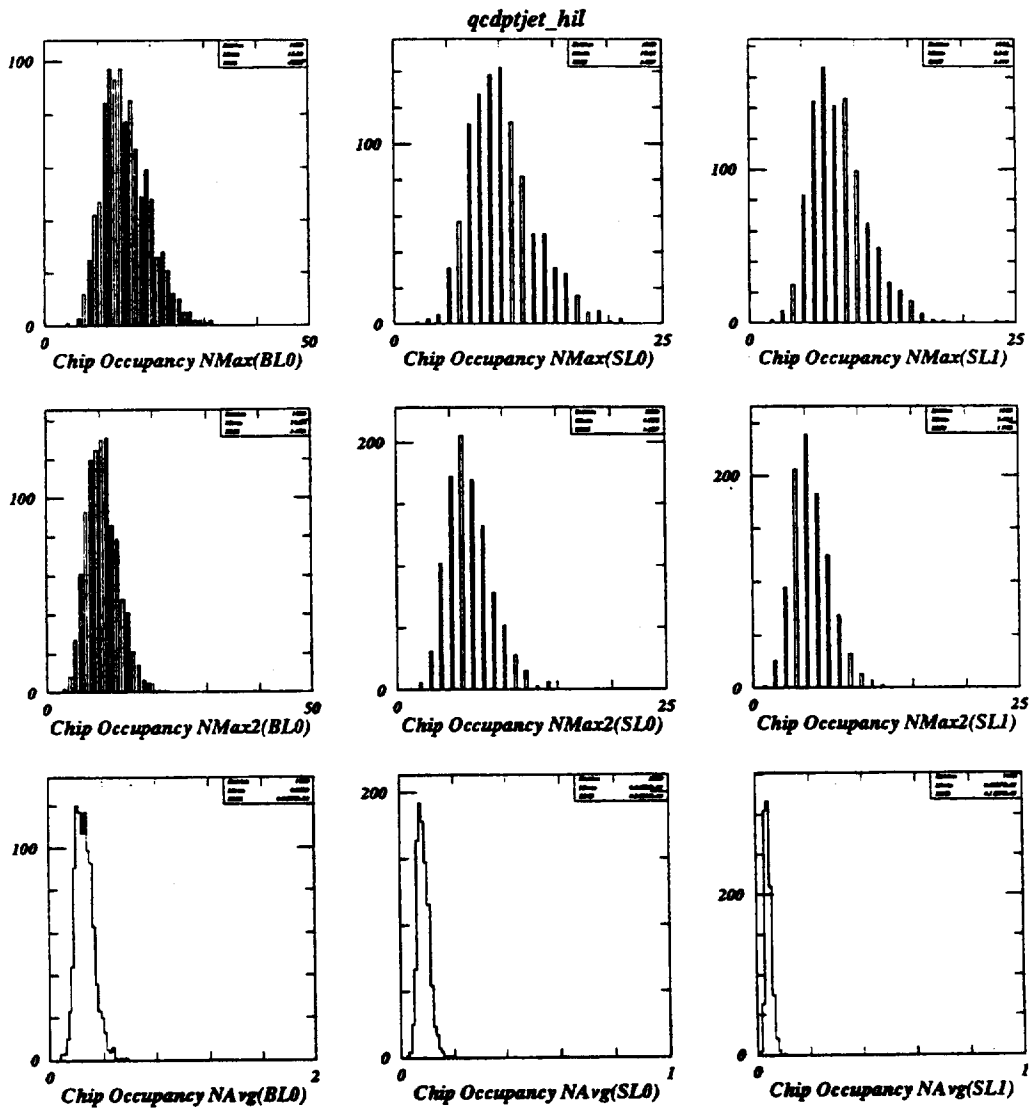
Figure 132: A series of plots for the three Barrel layers of the expected average (AVG) and worst-case (MAX = highest, MAX2 = second highest) chip occupancies for 1 TeV Dijet events at a luminosity of 10^{34}.