EXPERIMENTAL TECHNIQUES

D. Fournier and L. Serin
Laboratoire de l'Accélérateur Linéaire,
IN2P3-CNRS and Université Paris-Sud
91405 Orsay CEDEX, France.

Abstract

Experimental techniques to be used in the new generation of high-energy physics experiments are presented. The emphasis is put on the new ATLAS and CMS detectors for the CERN LHC. For the most important elements of these detectors, a description of the underlying physics processes is given, sometimes with reference to comparable detectors used in the past. Some comparative global performances of the two detectors are also given, with reference to benchmark physics processes (detection of the Higgs boson in various mass regions, etc.).

1 INTRODUCTION

Experimental techniques, used to exploit in the best possible way the accelerators available to the high-energy particle physics community, cover a broad and rapidly evolving field. This evolution has been particularly rich in the last few years, mostly because of the very demanding performances for experiments at the new high-energy, high-luminosity accelerators, the former Superconducting Super Collider (SSC), and the now-approved Large Hadron Collider (LHC) at CERN.

The approach taken for the lectures summarized below was to focus the presentation on the new techniques used by ATLAS [1] and CMS [2], the two large, multipurpose experiments at the LHC recently approved by CERN. Some global aspects of these experiments (choice of magnets, size, etc.) are also analysed.

Needless to say there are many other important experiments, planned or ongoing, which use new and clever techniques. Some examples are taken from KTeV, NA48, and BaBar. Having made such a selection, it is more than obvious that our presentation is going to be incomplete. Many good reviews exist already, including, for example, those of Ref. [3]. The material presented is divided into four main sections:

1. Magnets and muon detectors
2. Calorimeters
3. Tracking and vertexing detectors

4. Trigger, signal processing, and data transmission

Before the main presentation, it may be useful to recall a few parameters of the LHC machine [4]:

1. It uses the LEP tunnel which is 27 km in circumference.
2. The energy available in proton–proton collisions is 2×7 TeV. Ions up to lead can also be used, with an energy up to 3 TeV/nucleon.
3. Successive bunches collide every 25 ns. The size of the collision region is 5.6 cm (r.m.s.) in length and less than 20 μm transversally.
4. At the design luminosity of 10^{34} cm^{-2}s^{-1} there are about 20 inelastic interactions per bunch crossing.
5. The pseudorapidity distribution, $\eta = -\ln(\mathrm{tg}(\theta/2))$, extends up to ± 6 units, with a mean anticipated multiplicity of 7 (3.5) charged (neutral) particles per rapidity unit.
6. The irradiation levels corresponding to one year of high-luminosity operation are displayed in Figure 1 taking the ATLAS detector as example.

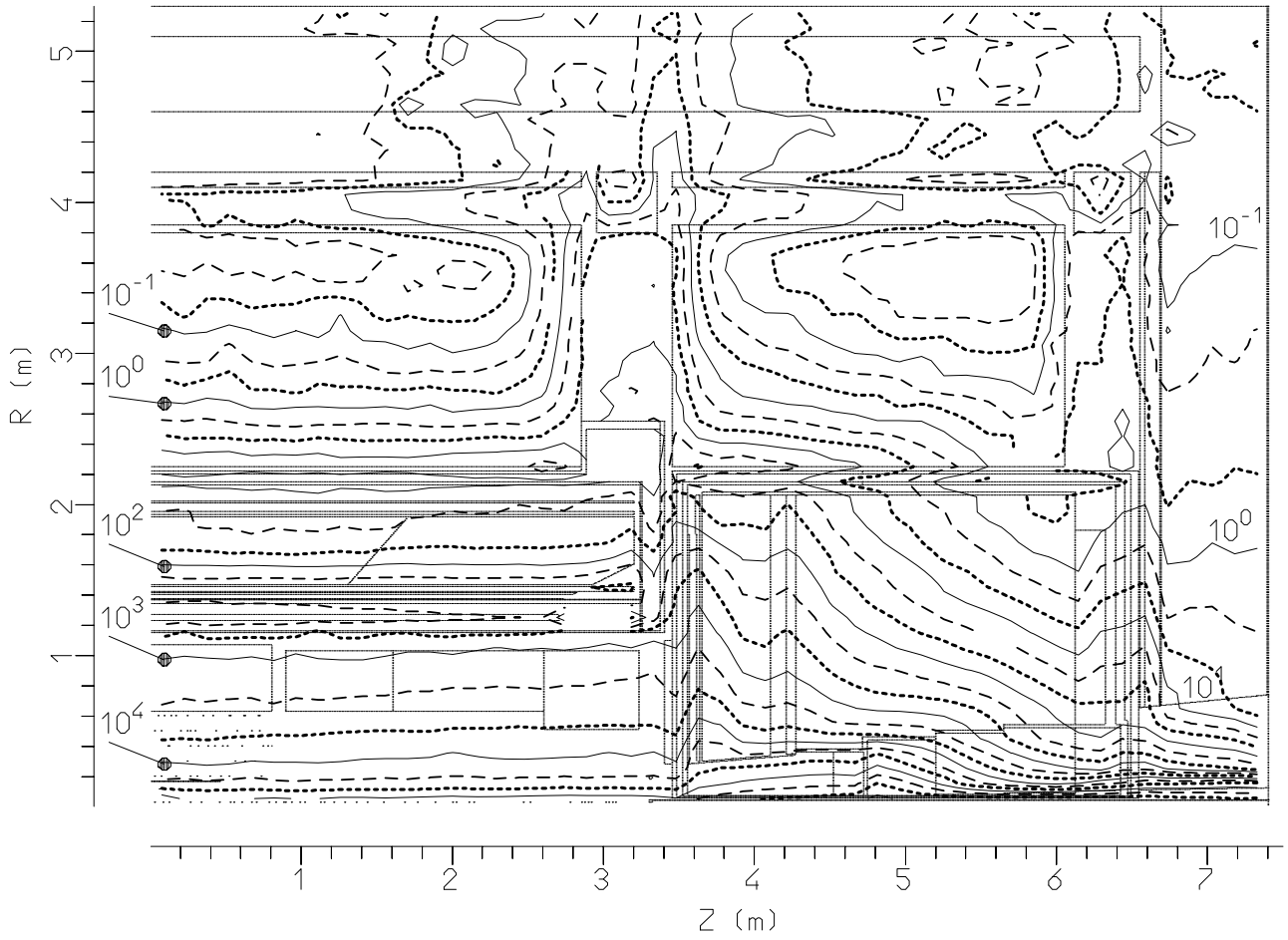

Figure 1: *Yearly integrated dose (Gy/year) calculated in the ATLAS detector for one year at high luminosity. The neutron fluence (over 100 keV) has a rather similar shape, with the isoline at 10^{13} n/cm^2/year more or less superimposed on the 1 kGy line (equivalent to 0.1 Mrad).*

2 MAGNETS AND MUON DETECTORS

2.1 Muon layouts at collider experiments

The layouts used so far, or planned, for collider experiments can be divided into three broad classes:

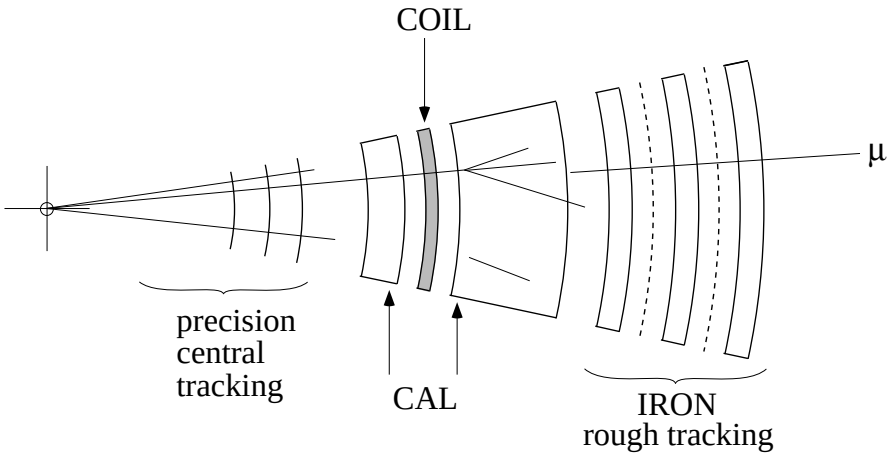

Figure 2: *Measurement in the tracking volume of muons identified in the calorimeter and absorber.*

– In the first one, the tracking system is used to analyse the momentum of the muon tracks, which are identified in the calorimeter and a backing filter consisting of iron slabs interspaced with coarse gaseous chambers (see Figure 2).
Such an approach has been used for example by the ALEPH, DELPHI and OPAL experiments at LEP. This way of proceeding leads to comparatively cheap detectors, but is limited to low or medium multiplicity. In a dense environment it becomes harder and harder to correlate the signals in the muon filter with one particular track in the tracking system, and measure its momentum.

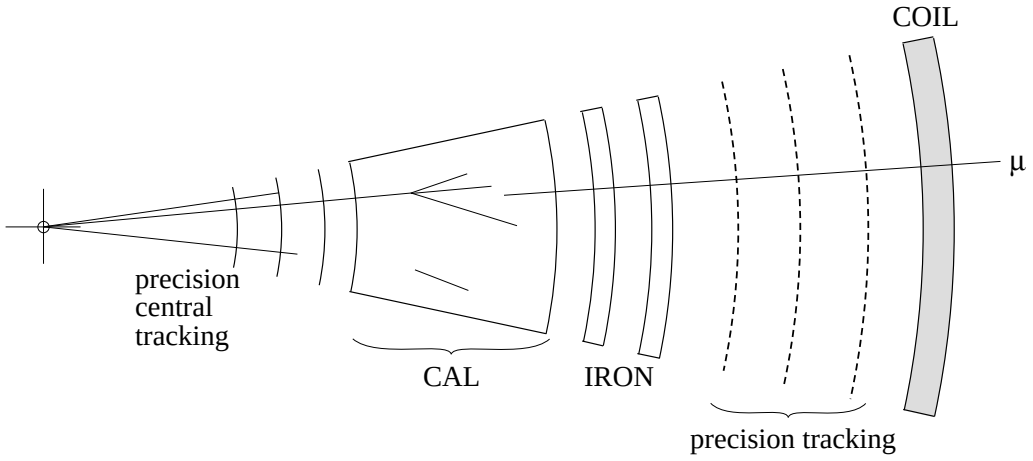

Figure 3: *Identification and measurement of muons after full absorption of hadrons.*

– The approach at the other extreme consists in absorbing as completely as possible all hadrons produced in the collision, and measuring, in a cleaned-up environment

the leftover particles, which should predominantly be prompt muons, i.e. produced at the vertex (see Figure 3). This solution was adopted by the L3 experiment at LEP, and was also retained by the former GEM experiment at SSC, both with a solenoidal field. The ATLAS experiment is also going this way, but using a toroidal field. This layout is clearly the safest one for high multiplicities, provided the absorber is thick enough (see below) since it can provide 'stand-alone' measurements in the harshest possible environment. It is, on the other hand, a rather expensive approach since one needs to equip a very large magnetic volume with precision tracking.

– As usual, an intermediate way of doing things is also possible. This is the way chosen by the CMS experiment. The magnetic field in the iron absorber is large enough to give, using precision chambers, a stand-alone measurement, (less precise than in case 2), and in many cases an extrapolation to the precision tracking volume is possible, which greatly improves the performance and brings it to the same level as, or possibly better than for scheme 2. A sketch of this approach, using a solenoidal field like in CMS, is shown in Figure 4.

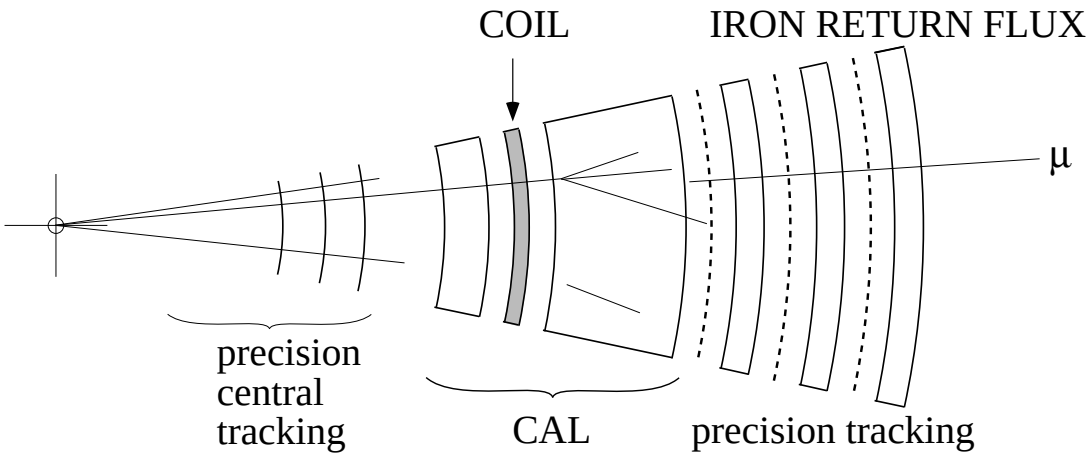

Figure 4: *Combined measurement of muons in the return flux and in the central tracking.*

2.2 Size of the muon detection system

2.2.1 Absorber thickness

As will be discussed in the Calorimeter section, high-energy particles entering a material slab start first to 'multiply' by the showering process, and it is only after the low-energy fragments have been slowed down, and captured or absorbed, that the mean number of particles in a section transverse to the initial particle direction starts to decrease.

In order to judge the 'cleanliness' of the space behind an absorber slab, where a muon detector would be positioned, the main quantity of interest is the 'punch-through' probability, i.e. the probability that at least one particle be observable, behind the absorber, in a certain area of sufficient radius around the initial particle direction.

The main scale parameter in this problem is the 'interaction length' of the absorber, namely, the mean distance between successive, inelastic collisions of the initial high-energy particle with the medium. Typical numbers are 17 cm for iron and lead, 15 cm for copper, 10 cm for tungsten.

In Figure 5 are shown recent results obtained by the RD5 Collaboration [5]. In order to reduce the punchthrough probability to 1% or lower, at least 2 m of iron are needed for an initial 30 GeV pion, and about 3 m at 100 GeV. It is also observed that the presence of a magnetic field transverse to the initial particle direction does not change the results significantly. The histogram shown is the result of a simulation of the cascade process using the 'FLUKA' package [6].

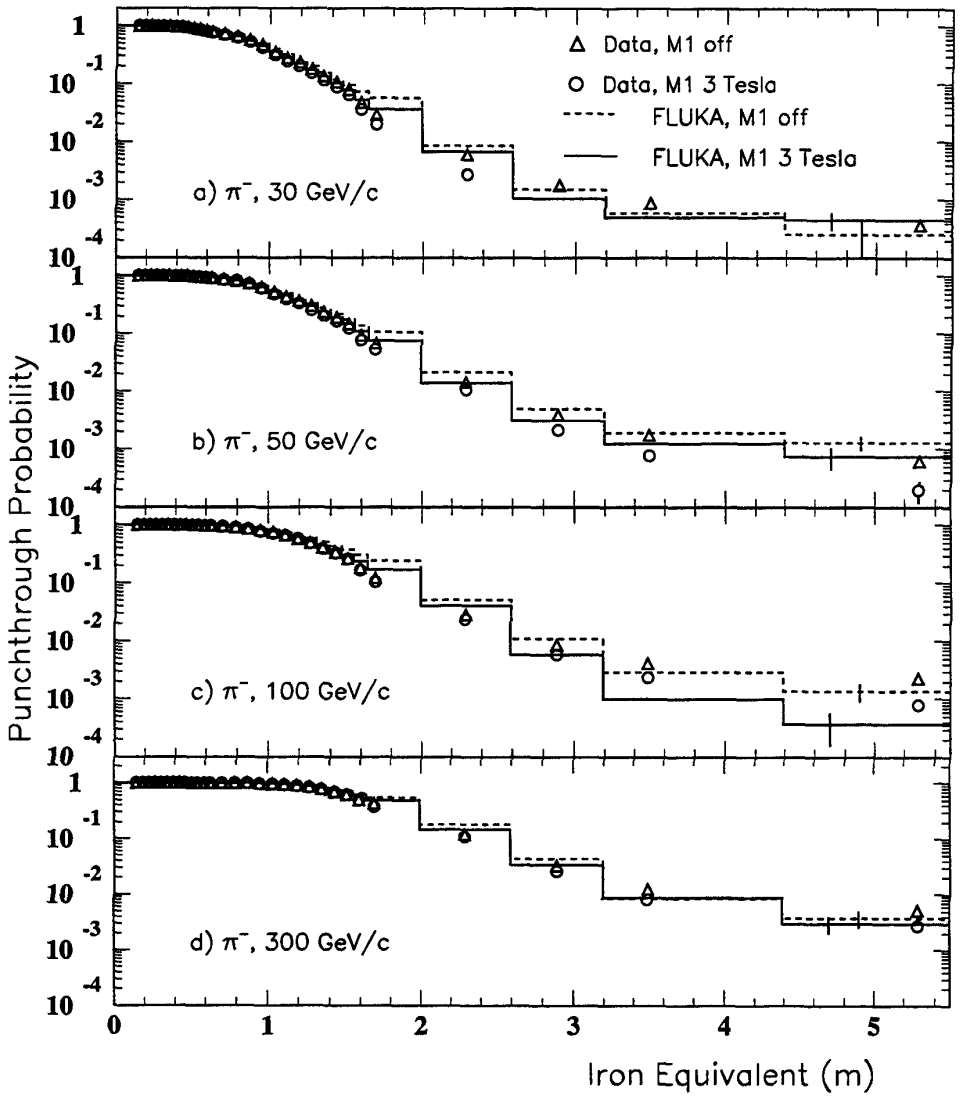

Figure 5: *Punchthrough probability of pions as measured by RD5 [5].*

In practice, taking into account the particle content of jets, the effect of the Lorentz boost etc., an equivalent of 10 (14) interaction lengths (λ) is required in ATLAS in front of the muon chambers for central (forward) rapidities. Most of this material is in fact contained in the calorimeter, and only a few λ of shielding are added to satisfy the requirement. Taking into account also the size of the inner detector cavity (1.2 m radius, 3.4 m half-length), the muon detection in ATLAS starts at 4.2 m radius and 7 m along the beam.

In CMS, the muon detection starts somewhat earlier (9 and 13 λ in the central and forward regions, respectively) but, because of the approach chosen (magnetized iron, see above), the last layers are shielded by 18 (26) λ.

2.2.2 Size of the magnetic system

This is most easily seen in the case of an air magnet (ATLAS toroid or L3 solenoid). The barrel toroid system of ATLAS is sketched in Figure 6. The field (decreasing as 1/R) is on average 0.8 T, over 4 metres. This results in a sagitta of about 0.5 mm for 1 TeV P_{T}. If the goal is set to measure it at the 10% level (for example for unambiguous sign determination), one immediately deduces that the sagitta should be measured with a 50 μm accuracy, indeed not a simple task for such huge volumes.

To obtain the field considered above is not trivial either: Ampère's theorem immediately gives the current needed:

$$2\pi RB = \mu_0 nI$$

$$nI \simeq 20 \times 10^6 \mathrm{At}$$

With eight coils, this amounts still to 2.5×10^6 ampere turns per coil, and definitely requires the use of the superconducting technique.

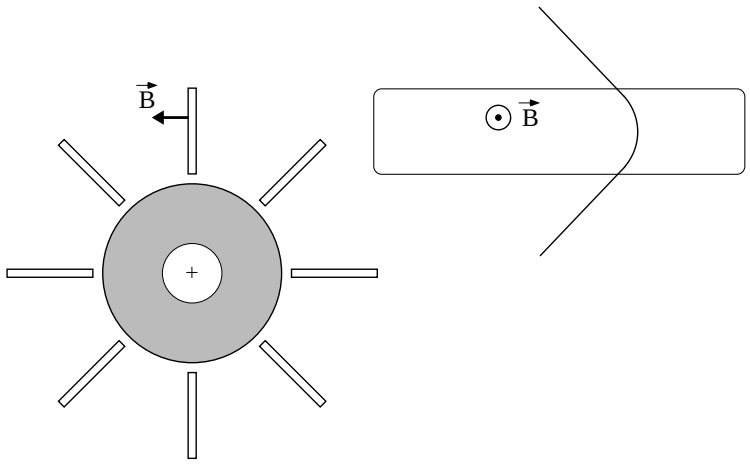

Figure 6: *Sketch of the ATLAS barrel toroid with a bent track.*

2.3 The ATLAS toroid

The layout of conductors in the ATLAS toroid is sketched in Figure 7, and the magnetic field map in Figure 8. With the parameters given above, it is straightforward to calculate the stored energy: about 1.5 GJ. How to dissipate this energy in case of a quench (i.e. loss of superconductivity) in one of the coils is one of the numerous technical problems of this magnet. Another challenge is to design a structure capable of holding the magnetic forces, both in the normal regime, and, more difficultly, in case of a quench [7]. As an illustration it is instructive to give the magnetic force per metre of coil: about 200 tons! This is mostly withstood by the 'in-plane' rods of each coil, but there remains a net inward force on each coil of about 1.5 ktons. This is to be supported by the octagonally shaped 'voussoirs' (bars linking each coil to the next in the transverse plane).

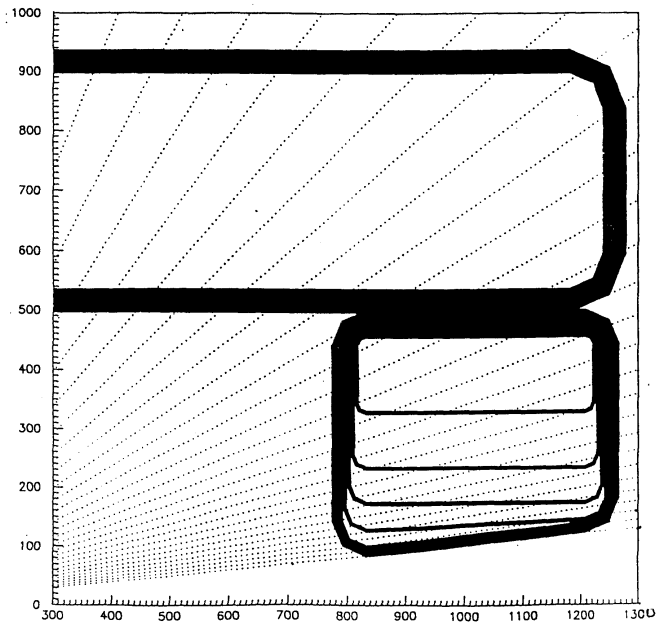

Figure 7: *Layout of ATLAS toroid conductors in the (z,r) plane. The lines are drawn every 0.1 pseudorapidity unit up to* $\eta = 3$.

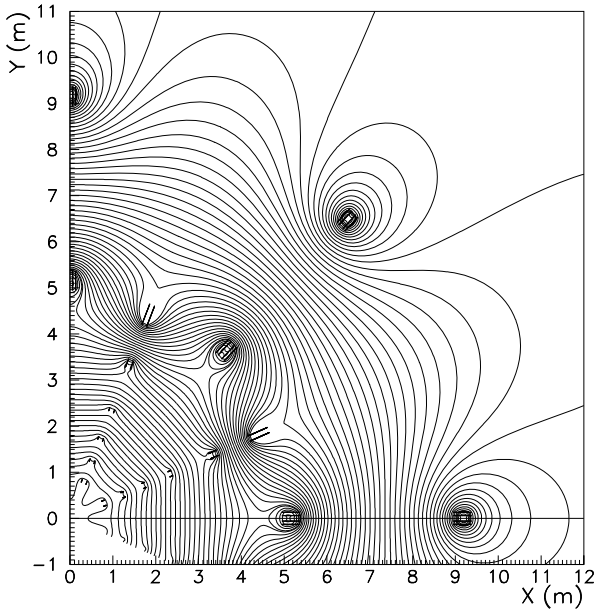

Figure 8: *Magnetic field map at* $z = 10$ *m.*

As shown in Figure 9 the field strength, represented by $\int B_{\perp} dl$, is maintained and even increases significantly up to large pseudorapidities ($\eta = 3$). This is in contrast with the solenoid case where this quantity starts to drop significantly as soon as trajectories no longer traverse the coil windings ($\eta = 1.5$ in CMS).

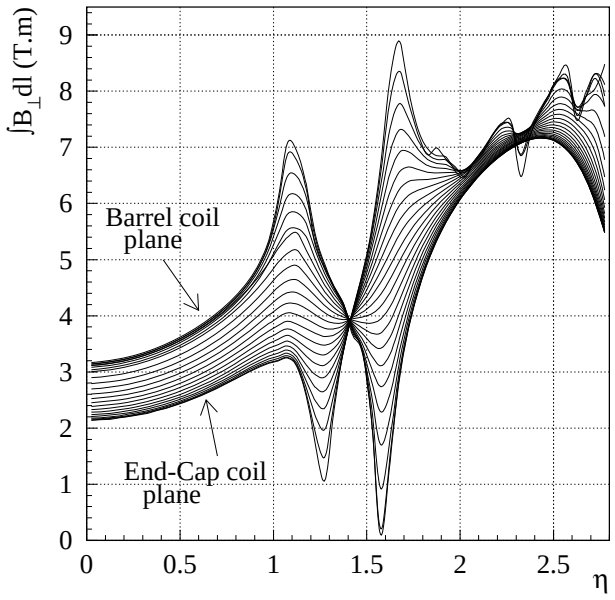

Figure 9: *Transverse field integral in the ATLAS toroid as a function of the pseudorapidity.*

A significant drawback of the toroid, however, is that the field for the central tracker has to be provided by an extra magnet (a small solenoid), thus requiring in total four sets of magnets for the experiment, while a single large solenoid does the job in CMS.

Before leaving this section, it may be relevant to indicate that toroids, even though they have been much less used than solenoids in the past, are, however, part of recent experiments, like CHORUS at CERN (neutrino oscillations, [8]), and at CEBAF [9].

2.4 Some considerations about solenoids

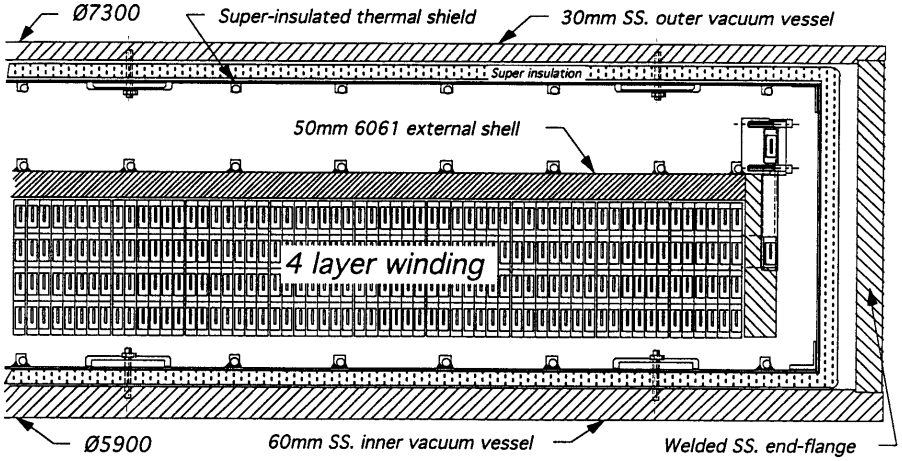

Figure 10: *Cut through the CMS solenoid end, showing the 4-layer winding.*

2.4.1 The CMS solenoid

The CMS solenoid is a very large magnet, with a 5.9 m (7.3 m) inner (outer) diameter and half-length of 6.5 m. A field of 4 T is obtained by four overlaid windings, each carrying 20 kA (see Figure 10). The magnet contains the tracker and calorimetry. It

represents by itself about one interaction length (at 90 degrees). The energy stored is about 2.5 GJ. As for the ATLAS toroid and thin solenoid (see below), the thickness is dictated by the necessity of holding the magnetic forces (and weight), and by the requirements of energy dissipation in case of quench.

Muon detection starts right after the magnet, in the 'return field' region. The behaviour of tracks in the transverse plane is sketched in Figure 11 for two values of η. The effect of the two bendings, 'direct' (in the solenoid) and 'reverse' (in the 'return field' region) is such that, in principle, outside the magnet, tracks point back to the vertex. This is only approximately true because of energy loss and multiple scattering. However, it does indicate that no useful information is to be obtained from chambers outside the iron return yoke. In terms of accuracy, the last layer of the tracking chambers is particularly important since it is close to the maximum sagitta point (about 1 mm at 1 TeV), and likely to be quiet enough to allow unambiguous assignment to a muon road extrapolated from the chambers behind the calorimeter.

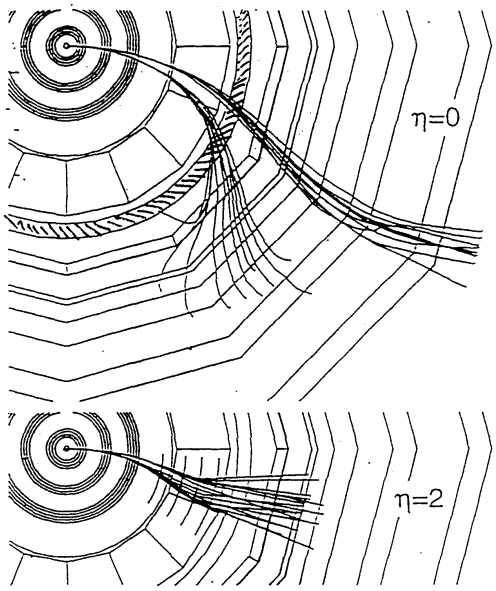

Figure 11: *Simulation of muon trajectories in the CMS transverse plane. The dispersion of low-momenta tracks is due to multiple scattering.*

It is also to be noted that the interaction vertex, known to very good accuracy in the transverse plane (10 μm or so at the LHC), represents an extremely important point on the muon track for a solenoid, while, because of vertex spread in the longitudinal direction (about 5 cm r.m.s. at the LHC), it is almost useless for a 'stand-alone' toroid measurement.

2.4.2 Thickness limit of thin solenoids

In optimizing the overall detector layout, it may be advantageous to limit the solenoid radius in such a way that it contains the tracking, which needs the field, but not the surrounding calorimeter for which the field is usually a 'nuisance'.

In this case the solenoid represents 'dead material' in the path of particles to be measured accurately by the calorimeter, and its thickness has to be reduced to a minimum. The problem is particularly acute for electrons and photons, for which this material may compromise an intrinsic resolution which is otherwise excellent (see next section). In this case the relevant parameter is the thickness of the solenoid expressed in radiation lengths (X_0).

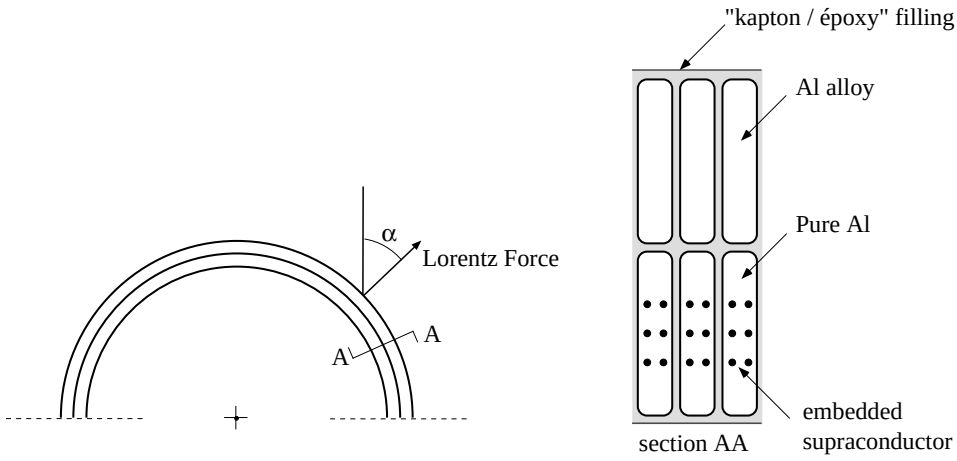

Figure 12: *Sketch of the two layers of a thin superconducting solenoid.*

The state of the art of solenoid construction has reached a point where the thickness is limited by 'fundamental' considerations, in particular the elastic limit of aluminium alloys. The calculations are rather straightforward [10].

The solenoid can be considered as being made of two layers, see Figure 12. The inner one (thickness ϵ_1) is the conductor: superconducting wires embedded in pure aluminium which is selected because of its very low resistivity at 4 K. (It carries the current in case of quench without excessive temperature rise, but it has very poor mechanical properties.)

The outer one is made of aluminium alloy; its role is to take the magnetic force. In an 'infinite solenoid model', the field is uniform with value B in the solenoid volume, and decreases to zero linearly as it traverses the conductor. It is then straightforward to calculate the tension force acting on a current loop:

$$F = \int_{-\frac{\pi}{2}}^{0} IB\cos(\alpha)d\alpha = I\langle B\rangle R = IBR/2$$

from which one derives that the thickness ϵ_2 of the surrounding cylinder, the maximum admissible stress θ and the field are related by:

$$B^2R = 2\mu_0\theta\epsilon_2 \ .$$

For example, in the SDC project, one had $B = 2$ T, $R = 2$ m, $\epsilon_2 = 31$ mm, which corresponded to a stress of 100 MPa, already at the limit when neglecting the contribution of the 'pure aluminium' layer.

The total thickness $\epsilon = \epsilon_1 + \epsilon_2$ should also be such that if all the energy W is dumped in the coil after a quench, the temperature does not rise by more than about 90 K:

$$\epsilon = W/((2\pi Rl\rho)(W/M)) \ ,$$

$$\epsilon = (B^2 R)/(2\mu_0 \rho c \Delta T)$$

in which $c\Delta T = W/M = 7$ kJ/kg corresponds to the ΔT of 90 K for aluminium, and ρ is the aluminium mass per unit volume.

In the example considered above this gives a total thickness of 70 mm, (i.e. about 0.8 X_0) with 31 mm for the support cylinder and 39 mm for the conductor.

It is finally interesting to remember that, from both effects, the thickness of a solenoid should increase like B^2 and like its radius R.

2.5 Muon chambers

The muon detectors should provide momentum measurement and allow for triggering. In both ATLAS and CMS these requirements are fulfilled by two sets of interleaved chambers: trigger chambers, and precision chambers.

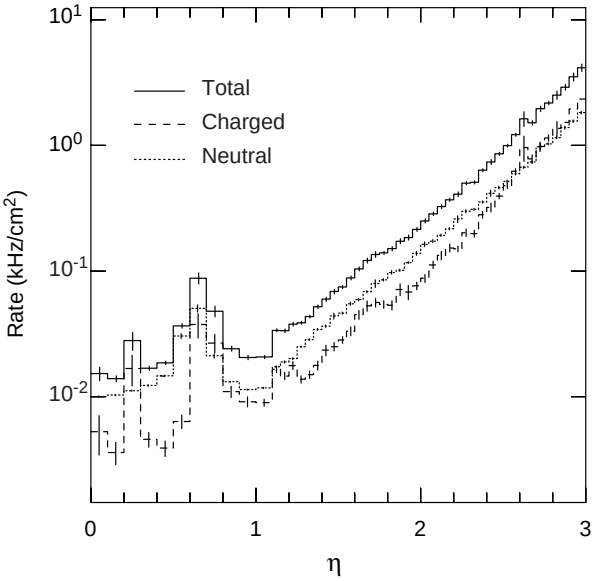

Figure 13: *Calculated background rates in the inner layer of ATLAS muon chambers.*

2.5.1 Background rates

The requirement of covering a large detection area (about 20 000 m^2 in ATLAS) with precision devices (better than 100 μm accuracy) points naturally to the drift chamber technique. While in the central barrel region the counting rate (entirely dominated by background) is well compatible with the drift chamber capabilities, it becomes so high in the forward region that proportional chambers with thinner gaps and close wire spacing have to be used. As an illustration Figure 13 shows the expected rate calculated in ATLAS for the inner muon layer, indicating a factor 100 more flux at rapidity close to 3 than in the central barrel part.

The dominant background is due to low-energy neutrons, photons, and hadrons originating from primary hadrons interacting with the forward calorimeter, the beam pipe, and other machine elements.

It is interesting to note that the background due to the neutral component is actually dominant over charged particles, except at large rapidity where punchthrough comes

in. A precise calculation of the rate therefore necessitates a good understanding of the production of low-energy neutral particles in shower tails and of the wire chamber signal formation from photons and neutrons.

Photons interact primarily via the Compton effect and conversions (see Section 3). Neutrons of thermal energy appear through radiative capture. At higher energies (up to keV) they are much less visible because of lower capture cross-sections. They become visible again when recoil nuclei from elastic scattering have enough energy to ionize the chamber gas. In this respect the fraction of free protons (hydrogen atoms) in the chamber gas is an important element.

Knowing that ageing problems in chambers appear after an accumulated charge of about 1 C/cm, it is possible from Figure 13, using a typical gas gain of 10^4 and 'reasonable' chamber geometry parameters, to place a rapidity limit to where drift chambers are usable (about 200 Hz/cm^2). Beyond this limit, proportional chambers with cathode strip readout (CSCs) have been proposed by the two experiments [11]. In CMS, for practical reasons the change of technology coincides with the barrel/endcap transition. A gap and wire spacing of about 3 mm allows high-rate capability, and cathode interpolation can lead to accuracies down to 50 μm.

2.5.2 The CMS rectangular tubes

The barrel muon detector of CMS has four concentric sections, each containing eight layers of precision tubes with wires parallel to the beam axis, one RPC layer for triggering (see below), and further tube layers for measuring the other (z) coordinate. The eight layers are clustered in two groups of four, in order to provide an averaged precision point and a local direction. This is necessary to disentangle hits from the genuine muon track from those produced by associated δ rays or converted photons exiting from the iron.

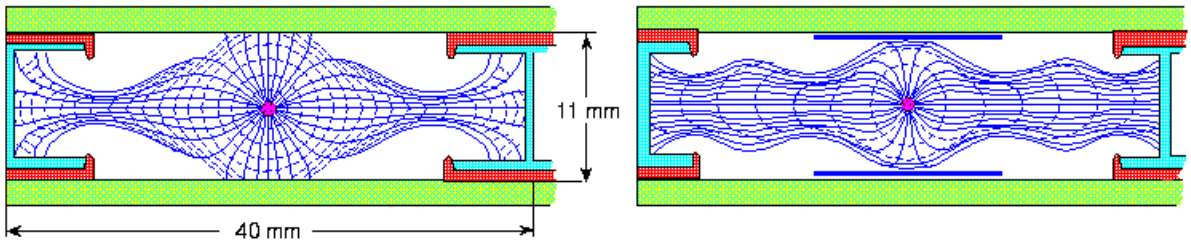

Figure 14: *Electric field lines in the CMS rectangular drift tubes.*

The field in the rectangular tubes (40 mm wide, 11 mm thick) is shaped in such a way as to give an almost uniform drift region for better accuracy (Figure 14). This requires insulating shims between the ground planes and the I-beams which separate the cells and define their geometry. The wires are 50 μm stainless steel (2.5 m long), and the gas used is Ar-CO_2. The maximum drift time is 400 ns. This will be digitized with a least count of 0.8 ns (40 μm). Multihit capability allows the separation of tracks distant by about 2 mm. Since most of the magnetic flux is returned through the iron, the field in between the iron slabs, where the chambers are located, should be weak, avoiding $E \times B$ effects. The same is, however, not true in the endcap part where the CSCs are located.

Test beam data have given a resolution per tube of 150 μm. The construction process of a layer of four chambers should insure an overall geometrical precision of better than 100 μm within a stack of typically 2.5×2.5 m^2. The precision assumed in CMS simulations, for a stack of four chambers, is 200 μm. The overall resolution of the CMS muon spectrometer is given in Figure 15.

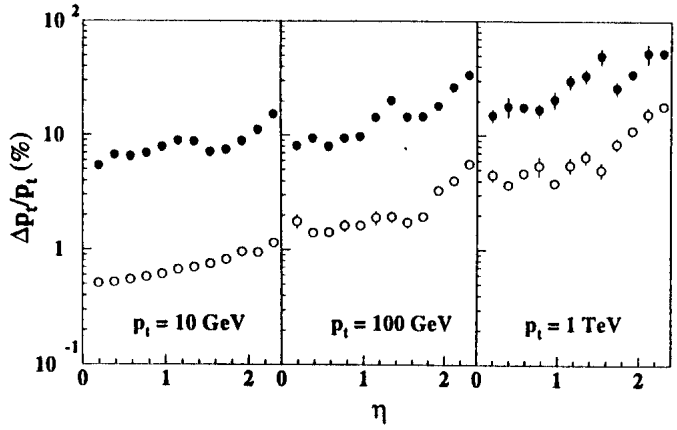

Figure 15: *Simulated resolution of muon momentum measurement in CMS. The open points correspond to the combined use of the central tracking.*

2.5.3 The ATLAS round tubes

The ATLAS muon detection system is organized in three stations. Each station contains two stacks of three layers of precision chambers with additional layers of RPCs (or TGCs) for triggering and second coordinate measurements in the two outer stations. Beyond a rapidity of 2 (inner station) or 2.6 (middle station), the drift tubes are replaced by CSCs

Given the magnet symmetry and the toroidal nature of the field, the chambers are arranged in an eight-fold symmetric pattern, with tubes (and wires) perpendicular to the colliding beam direction. As a consequence, the longer tubes are 6 m long. In the barrel/end cap transition (Figure 7) the tubes are located in a rather non-uniform field region. These aspects, among others, have led to the 'Monitored Drift Tubes' concept: tubes which are round, and arranged in stacks of a rather floppy nature, but carefully monitored in shape and position. In particular, in order to avoid instability of long wires, which sag by up to 0.6 mm under gravity, the chamber stacks are deformed in such a way as to give a comparable sagging to the tubes.

The tubes are 3 cm in diameter. They are made of extruded aluminium alloy. The wires are 50 μm in diameter. The round shape of the tubes allows over-pressure operation which leads to better accuracy. The chosen value is 3 bars.

Out of the—typically—300 ion clusters produced per track in a tube, only about 20 are used to trigger the discriminator which commands the TDC. This is equivalent to saying that only a slice of 2 mm thickness or so, perpendicular to the track, is 'useful' in the position measurement. All of the ionization, however, is amplified and participates in ageing.

A careful study of effects limiting the accuracy of round tubes can be found in

Ref. [12]. In test beams, resolutions as good as 60 μm have been obtained over most of the drift path.

In order to exploit fully this intrinsic resolution, the drift-time-to-distance relation has to be known to 0.1%. As opposed to the case of the CMS rectangular tubes, the E-field is non-homogeneous ($1/r$) and the $E \times B$ effect is sizeable (under combined E and B, the electrons do not drift along E, but at a certain angle to it, the Lorentz angle, making the drift path curved). It is proposed in ATLAS to calibrate the time-to-distance relation in an iterative way, using the muon tracks themselves ('autocalibration').

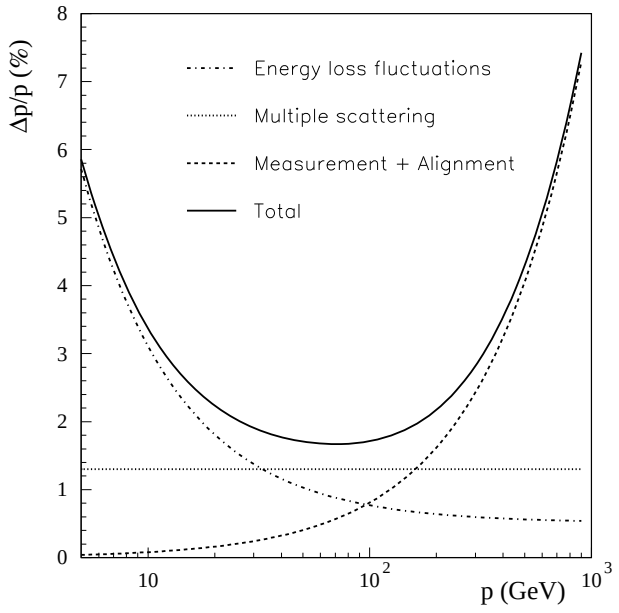

Figure 16: *Simulated resolution of the ATLAS 'stand-alone' muon spectrometer as a function of momentum.*

Another critical point when exploiting high resolution is the requirement to know where each tube is with respect to the others and with respect to the magnet system (and the tracker when combined information is used).

Within a stack of a certain size, the tubes are accurately positioned by construction. Global deformation of the stacks and the positioning of a stack with respect to a reference frame need to be monitored. One way to do this was pioneered by the NIKHEF group in L3 [13]. This uses light rays emitted by light-emitting diodes, and a CCD camera for detection. These two elements are supposed to be known in position. A lens is attached to the object to be monitored which images a mask in front of the LED onto the CCD. Any displacement of the lens is thus continuously monitored, with an accuracy of 1 μm transversely, and 30 μm longitudinally (from the image size).

The accuracy of the ATLAS muon spectrometer is shown in Figure 16 which illustrates the contributions from various effects at $\eta = 0$: at low energy the fluctuations of energy loss (4 GeV on average) in the 10 λ in front of the spectrometer are the dominant effects. This can be improved by measuring the energy deposited in the corresponding calorimeter tower. The spectrometer intrinsic performance (resolution and alignment) limits the high energy part. Figure 17 shows the resolution as a function of pseudorapidity

for a selected numbers of P_T values. In these plots the inner tracker is not used.

2.5.4 Trigger chambers

Owing to the rather large size of both the ATLAS and CMS drift tubes, the drift duration of about 400 ns is much larger than the time distance between bunch collisions at the LHC (25 ns). This excludes the use of these detectors, at least in a simple and fast way, for triggering.

It has been proposed by both experiments to add dedicated layers of 'Resistive Plate Chambers' for this purpose [14]. These chambers are wireless. The amplification in the d.c. field (typically 10 kV for a 4 mm gap) is limited by the resistive nature of the electrodes. A position measurement is obtained from strips capacitatively coupled to the HV electrodes. These chambers are easy to build, and their time accuracy is excellent (1 ns). They are however limited in rate to about 100 Hz/cm^2. This is insufficient to cover the forward region of ATLAS for which 'regular' thin-gap MWPCs are proposed.

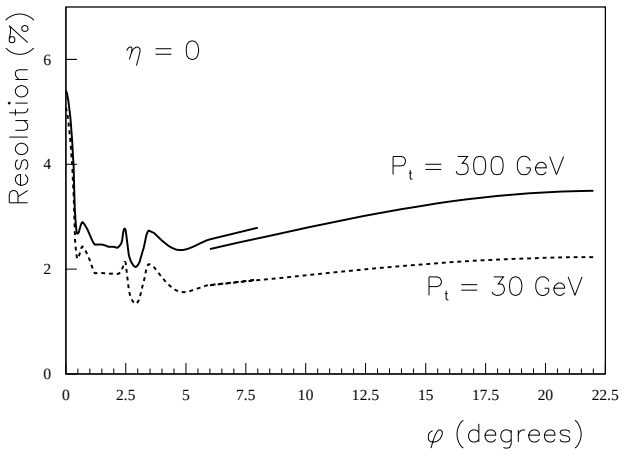

Figure 17: *Calculated momentum resolution of the ATLAS 'stand-alone' spectrometer as a function of the azimuthal angle. In the coil plane near $\varphi = 0$ the resolution is limited by multiple scattering.*

2.5.5 Illustrative performance

A particularly interesting channel for gauging the performance of the muon spectrometer is a Higgs boson of about 130 GeV mass, decaying into four muons. Mass resolution for this channel is given at the end of Section 4, where consideration is given to the performance of the tracker which also contributes significantly to this measurement.

3 CALORIMETRY

The principle of this technique is to perform the energy measurement of an incident particle by total absorption, where a fraction of the total energy is transformed into a measurable quantity (charge, light, heat...). The basic processes for electromagnetic ($\gamma, e^{\pm}$) and hadronic showering are described and some basic calorimeter concepts and limitations reviewed. The state of development of the homogeneous and sampling calorimeters is illustrated by various examples, with emphasis on the CMS and ATLAS calorimeters.

3.1 Electromagnetic (e.m.) interaction

The interaction of electrons or photons with matter is a multistep phenomena in which various energy loss mechanisms take place, depending on the energy, as displayed schematically in Figure 18. In Figure 19 are shown the fractional energy loss of electrons and the interaction cross-section of photons in lead as a function of energy [15]. Above 1 GeV, bremsstrahlung for electrons (interaction with a nucleus) and pair production for photons are the dominant mechanisms leading to a cascade of electrons and photons until the energy of these secondary particles falls into a region where the energy loss is dominated by ionization. It should be noticed that the clear dip in the photon cross-section implies that the longitudinal leakage will be dominated by photons of a few MeV. These photons are also responsible for the tails observed in lateral profiles (see below).

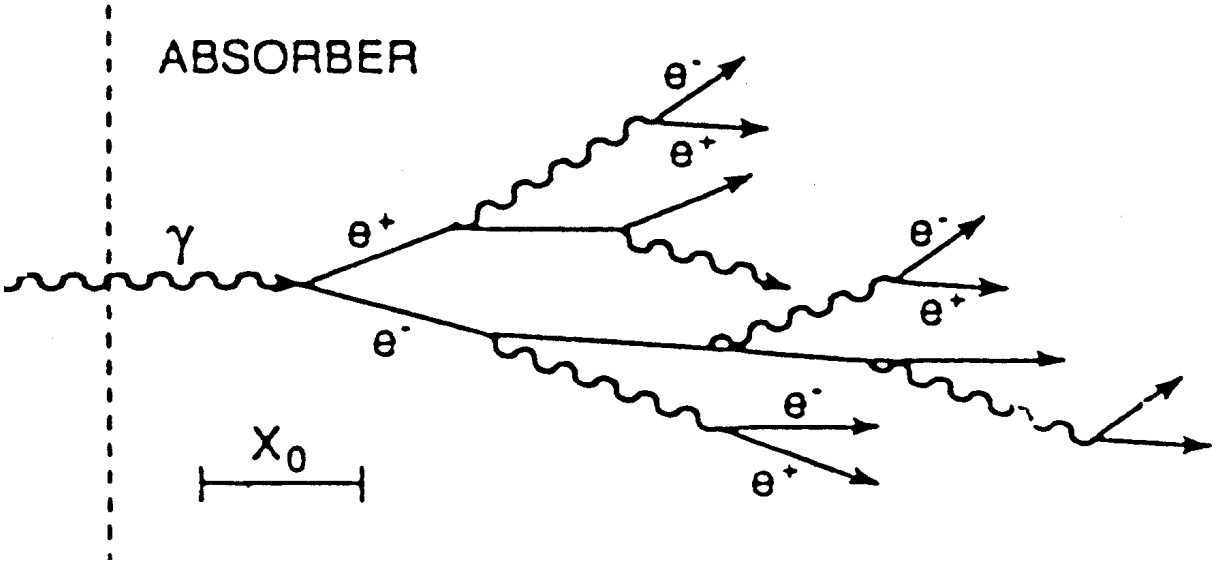

Figure 18: *Simplified development of an electromagnetic shower.*

In order to discuss electromagnetic calorimetry independently of the material used, it is convenient to introduce two quantities:

- The critical energy, E_C, is defined as the energy at which the loss by radiation is equal to the loss by ionization. This can be parametrized as a function of the atomic number by $E_C = 800/(Z + 1.2)$ MeV.
- The radiation length X_0 is defined as the distance over which the energy of the electron is reduced by a factor $1/e$ by bremsstrahlung only. The radiation length can be parametrized with a 5% accuracy as $X_0 = 716.4A/[Z(Z+1)\ln(287/\sqrt{Z})]$ g.cm^{-2} where Z is the atomic number and A the atomic weight.

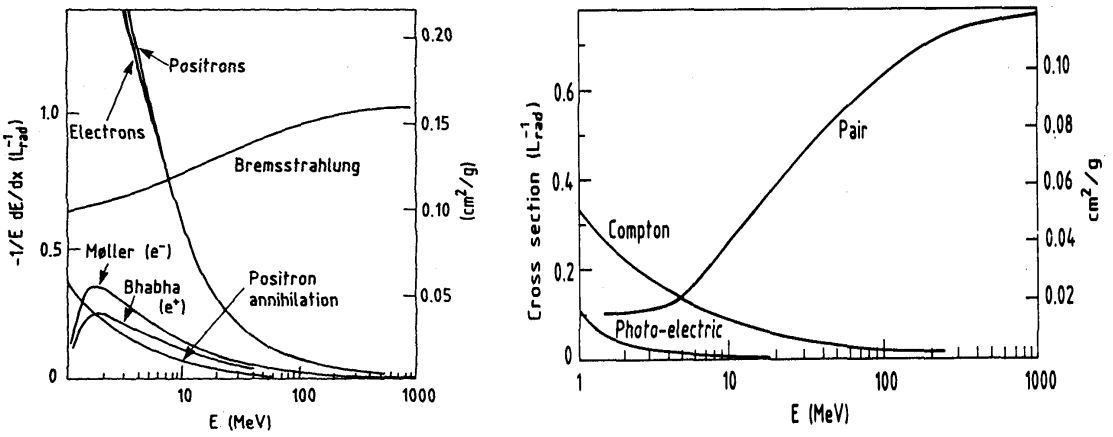

Figure 19: *Energy loss of electrons (left), cross-section of photons in lead (right) [15].*

The mean longitudinal shower development can be described by [16]:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}$$

where $t = x/X_0$ is the depth in radiation length, E_0 is the incident energy, a and b being parameters depending on the nature of the incident particle, $\mathrm{e}^{\pm}$ or γ.

Figure 20 [17] demonstrates that the longitudinal shower development of 10 GeV electrons has a similar shape in different absorbers to first order. However, a shower starts earlier in aluminium and is longer in lead because of a different $\mathrm{d}E/\mathrm{d}x$ energy loss in the material considered: 38.8 (7.2) MeV in one X_0 of Al (Pb). A 100 GeV electromagnetic shower deposits practically all of its energy in a 25 X_0 depth calorimeter. The longitudinal shower energy containment scales logarithmically as $\ln(E/E_{\mathrm{C}})$ which means one can keep the calorimeter rather compact even at LHC energies (a few TeV electrons).

An equivalent distance for the photon interaction is the mean free path, X_γ, i.e. the average distance after which the number of identical high-energy photons has decreased by $1/e$ through pair creation: $X_\gamma = \frac{9}{7} X_0$. The relative importance of the two processes (bremsstrahlung and pair creation) as the function of the depth is almost equal. However, the fraction of cascade energy carried by photons increases with depth.

The lateral spread of an electromagnetic shower is mainly due to the multiple scattering of electrons which do not radiate and have enough energy to travel far away from the shower axis. The transverse dimension is most conveniently measured in terms of the Moliere radius defined by $R_{\mathrm{M}} = X_0 E_{\mathrm{s}}/E_{\mathrm{c}}$ ($E_{\mathrm{s}} = 21$ MeV). To a good approximation R_{M} is inversely proportional to the density and gives a correct description of the lateral shower containment. Typically an infinite cylinder of radius $2 \times R_{\mathrm{M}}$ contains 95% of the shower energy. Figure 20 presents the lateral shower spread in a lead fibre sampling calorimeter [18] whose Moliere radius is about 2 cm: it is composed of a core very well explained by multiple scattering (over a few R_{M} distance) and a low-energy tail up to 50 cm. This peripherical part is due to low-energy photons (10–20 MeV) which are less attenuated in matter (see Figure 19): they can travel a relatively long transverse distance when they come from electrons which have scattered to a large angle from the shower axis.

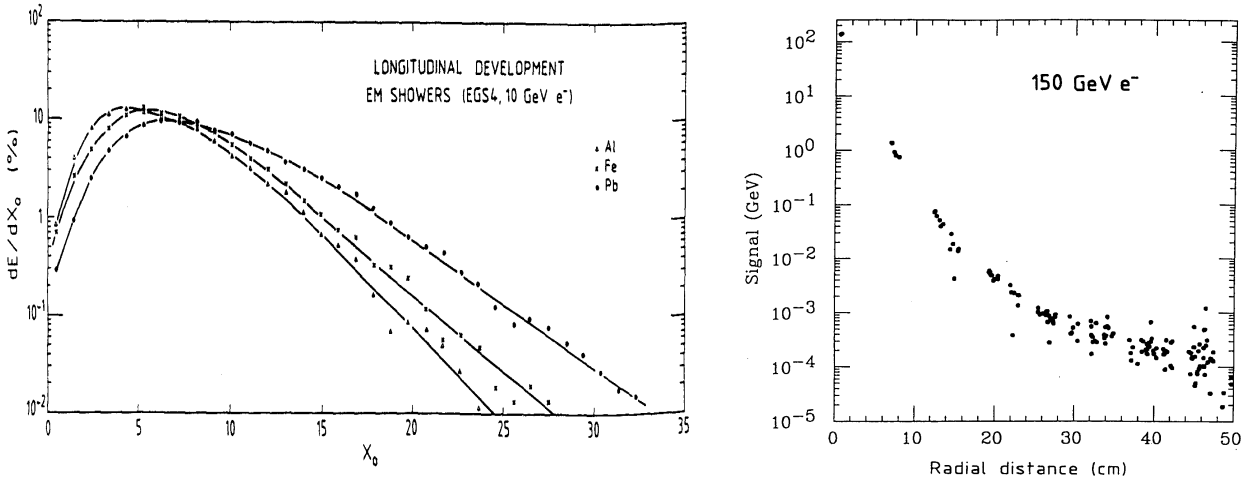

Figure 20: *Simulation of longitudinal shower development of e^- (left) [17], experimental lateral profile [18].*

3.2 Experimental requirements and limitations for electromagnetic calorimetry

Even if the primary goal is the energy measurement, the role of electromagnetic calorimeters is also to provide position and angular measurements, and some particle identification. Each of these points is considered below.

3.2.1 Energy measurement

The most important aspects of energy measurement are linearity and resolution. Non-linear response could be a major concern in experiments with a large energy range. For instance, in LHC experiments, one has to measure electrons with a transverse energy from a few GeV ($H \rightarrow 4e^{\pm}$) up to a few TeV ($Z' \rightarrow e^+e^-$). An example of response linearity over a 300 GeV range [19] is displayed in Figure 21 where there is a residual non-linearity smaller than $\pm 1\%$. Non-linearity is often due to the electronics chain and can in principle be corrected offline. Operating in a high magnetic field can also be a source of poor linearity at small energy (see Figure 21) [20]: for example low-energy charged particles can be swept out of the collected region.

The energy resolution is commonly parametrized by the following quadratic sum:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

where a is the stochastic term (intrinsic resolution), b the noise term, c the constant term and E is expressed in GeV. The parametrization by these three terms is not always the best for all calorimeters ($E^{-1/4}$ instead of $E^{-1/2}$ is sometimes used for crystal calorimeters because of rear leakage). Depending on the energy range involved in an experiment, it is obvious that the optimization of these three parameters can be very different.

The intrinsic resolution is strongly linked to the technology used. The best performance is reached by homogeneous calorimeters in which the particles release all their

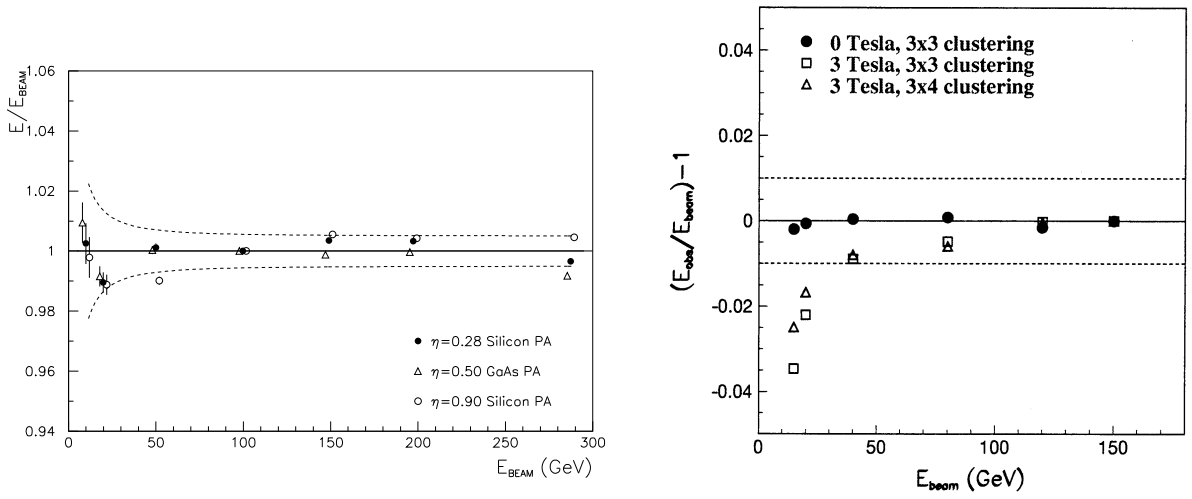

Figure 21: *Linearity response of the RD3 lead/liquid argon sampling calorimeter [19] (left), linearity of the lead/scintillator shashlik sampling calorimeter in a magnetic field environment [20] (right).*

energy in an active medium which provides the measurable signal (crystals, semiconductors, noble liquids, etc.). The intrinsic resolution is given by the statistical fluctuation of the number of detected primary processes. The threshold of detection is usually small enough that the number of primary processes, which could be detected, is large and gives good intrinsic resolution. This limit is reached only in semiconductor detectors used at low energy. In general this fluctuation has to be convoluted with instrumental effects which dominate the resolution: efficiency of collection and photoelectron statistics in crystals, lateral and longitudinal leakage in any calorimeter, etc. A particular case is lead glass in which the fluctuation of the observed primary process can not be neglected. The cut-off energy for Cherenkov radiation of electrons is 0.7 MeV in lead glass, and at most 1400 tracks are produced per GeV leading to a stochastic term of about $\sigma_{\rm p} \geq 2.6\%$. However, the number of produced photoelectrons per GeV (what is indeed measured) is about 1000 giving $\sigma_n = 3.2\%$ at the level of the photon statistics term. The resulting fractional energy resolution cannot thus be better than about $4\%/\sqrt{E}$. Some results achieved in large lead glass detectors are not far from this limit [21]. More detailed examples of homogeneous calorimeters are discussed in Section 3.4.

In contrast with homogeneous calorimeters, sampling calorimeters have only a fraction of the energy released in the active medium which is interleaved between passive layers called absorbers. The resolution is dominated by the sampling fluctuations in the total number of $e^{\pm}$ tracks crossing the sensitive layers. Using some approximations[1)], one can derive the behaviour of the energy resolution. The cascade development is thus modeled by the total track length $T = EX_0/E_{\rm C}$. Taking a sampling calorimeter with absorbers spaced regularly with a pitch x, the number of tracks intercepted in the gap is $N = T/x = E/\delta E$. δE is the energy loss of a minimum-ionizing particle in one cell of

[1)] The energy loss dE/dx is assumed energy-independent ($= E_{\rm C}/X_0$), the processes at low energy (Compton, multiple scattering) are neglected (Approximation B of Rossi [22]), and the cutoff energy is neglected.

thickness equal to x. The energy resolution is given by the fluctuation in the number of tracks N:

$$\frac{\sigma(E)}{E} = \frac{1}{\sqrt{N}} \propto \sqrt{\frac{t}{E(\mathrm{GeV})}}$$

where $t = x/X_0$ is the absorber thickness. The approximation used is enough to demonstrate the energy dependence of the resolution but more detailed explanations can be found in Refs. [22, 23].

As a very large fraction of the energy is deposited by low-energy electrons (MeV) in the high-Z material (absorber), the energy resolution is improved when decreasing the absorber thickness (or increasing the sampling frequency). This scaling law is valid as long as the absorber thickness is not so small that the crossings between consecutive layers are not correlated. It is useful to introduce the mip sampling fraction f_{mip} which is the fraction of a minimum-ionizing particle energy deposited in the active medium:

$$f_{\mathrm{mip}} = \frac{d(\frac{\mathrm{d}E}{\mathrm{d}x})_{\mathrm{act}}}{d(\frac{\mathrm{d}E}{\mathrm{d}x})_{\mathrm{act}} + t(\frac{\mathrm{d}E}{\mathrm{d}x})_{\mathrm{abs}}}$$

where $\mathrm{d}E/\mathrm{d}x$ is the energy loss, d and t the active and passive medium thicknesses. Since the electromagnetic showers contain low-energy photons at the end of the shower development which are absorbed by photoelectric effect ($\sigma \propto Z^5$), the fraction of measurable signal, f_{S}, is smaller than f_{mip} by the e/mip ratio, around 0.6–0.7. The resolution is improved with increasing sampling fraction. The performance of various sampling calorimeters with solid or active medium is summarized in Figure 22 [18] (d/f_{S} is to first order proportional to the absorber thickness in a sampling calorimeter with interleaved layers perpendicular to the particle direction).

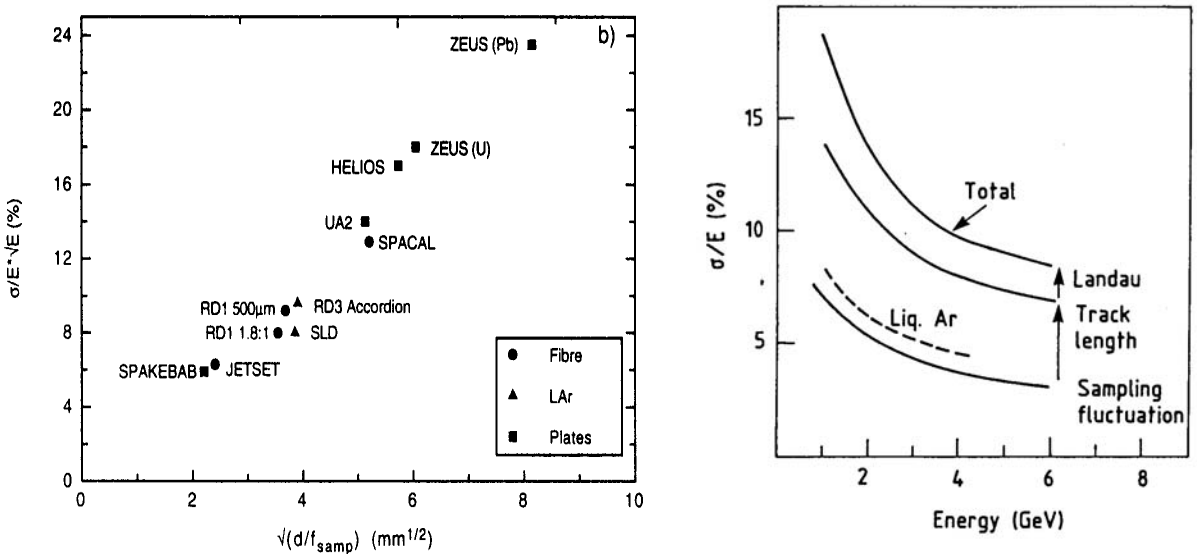

Figure 22: *Energy resolution of some sampling calorimeters (a %) as a function of $(d/f_S)^{-1/2}$, where d is the active medium thickness (left) [18], energy resolution of a gas sampling calorimeter (right) [24].*

In the case of a gaseous active medium, the intrinsic energy resolution is no longer dominated by the sampling fluctuations but by two other contributions:

- The Landau tails of the energy loss distribution in the gas which does not scale as $1/\sqrt{E}$.
- The large fluctuation on the path length coming from the low-density medium (gas) and the wide-angle spread of low-energy electrons. These electrons can easily move away from the shower axis and induce large signals.

As an example in Figure 22 the energy resolution of a lead/MWPC sampling calorimeter is plotted with the various contributions and compared with what could be obtained by filling the gap with liquid argon.

The (b) term of the energy resolution describes the noise and is dominant at low energy. Besides the electronics noise, a second contribution is important in LHC calorimeters: the pile-up noise due to multiple interactions. This reflects the high multiplicity of charged and neutral particles with a low average momentum (500 MeV/c) which impact in the calorimeter during each bunch crossing. The mean value of this energy can be either subtracted by hardware (filtering in ATLAS) or software. Figure 23 shows the probability of containing a transverse energy E_{T} in a calorimeter cell when 20 minimum-bias events are superimposed. Although the mean pile-up energy can be subtracted, its fluctuation remains and manifests itself as an energy-independent term in the energy resolution.

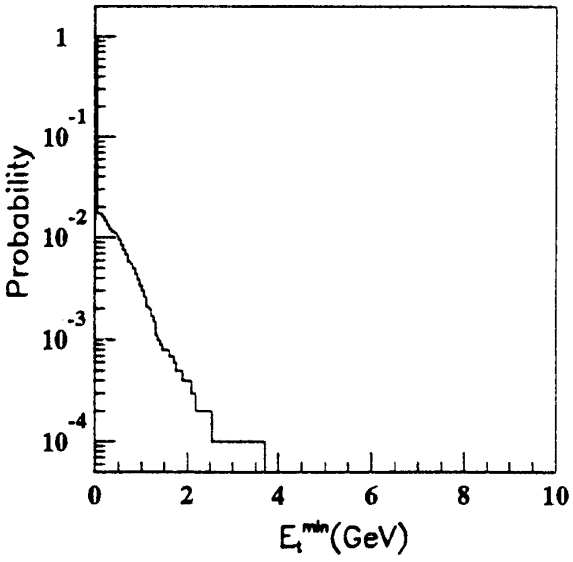

Figure 23: *Probability P(E) to have a transverse energy greater then E_T in a calorimeter cell when 20 minimum-bias events are added [25].*

The constant term (c), dominant at high energy, includes many contributions:

- Electronic non-uniformity, mainly due to the calibration of the cell-to-cell response.
- Mechanical non-uniformity, for example variation of the active and passive thicknesses as a function of position in sampling calorimeters.
- Signal variation with temperature or pollution which may vary across the calorimeter or as a function of time.

– Lateral and longitudinal energy leakage. The first contribution is imposed by cluster size considerations (noise, shower separation) and is in general small as the transverse shape differs little from one shower to another and is not very sensitive to the energy. The second term can give larger effects as the longitudinal shower shape presents more fluctuation and the fraction of energy escaping from the back is energy dependent. To illustrate this effect, the degradation of the fractional energy resolution as a function of the longitudinal leakage is plotted in Figure 24 with a homogeneous liquid xenon prototype [26].

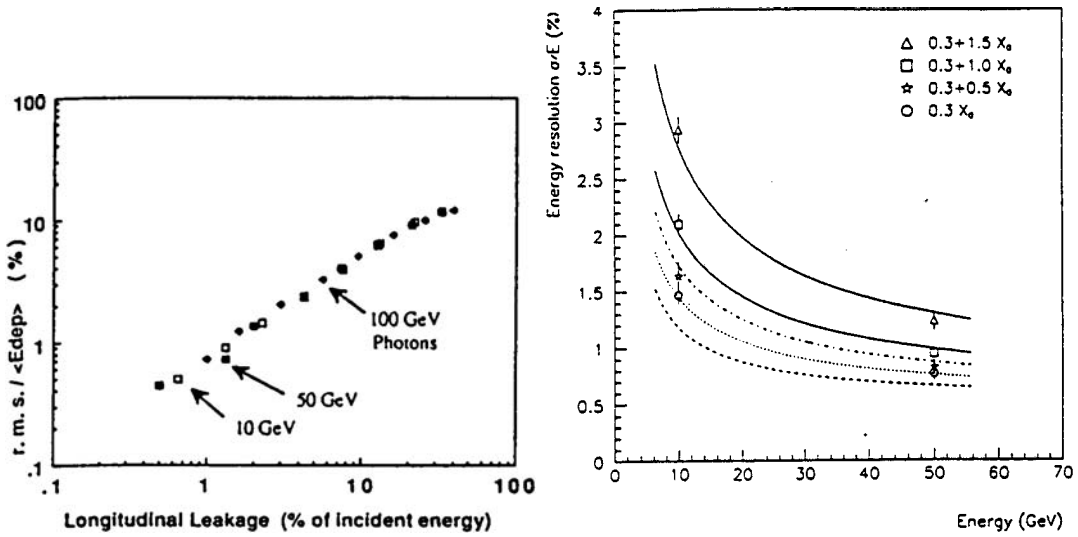

Figure 24: *Degradation of energy resolution with longitudinal leakage (left) [26] and material in front (right) [27].*

– Degradation of the energy resolution due to the material in front of the calorimeter as shown in Figure 24 from experimental results obtained with the NA48 liquid krypton calorimeter prototype [27]. An amount of more than 2–3 X_0 of inactive material in front of the calorimeter is considered a hard limit. This contribution can be reduced by using a presampler, in front of the calorimeter just behind the inactive material, whose signal gives an estimate of the energy lost.

An example of the different sources of non-uniformity identified in the ATLAS liquid argon prototype is presented at the end of this Section.

3.2.2 Position and angular measurement

The partition of calorimeter readout elements in strips or towers has been widely used to find shower impact positions by centre-of-gravity methods. The energy resolution in each of the cells reflects in a position resolution which also scales as $1/\sqrt{E}$. For readout elements of relatively large size (as compared to the shower width at the depth where it is sampled), it is well known that the method is biased, and requires correction factors. After those have been applied, resolution as good as a few mm $/\sqrt{E}$ is currently reached.

There is much less experience in obtaining from the calorimeter itself the shower direction. The need for an angular measurement provided by the calorimeter was especially stressed by the LHC physics requirement of the H $\rightarrow \gamma\gamma$ channel. The $\gamma\gamma$ angle enters in the mass resolution, and since at high luminosity the vertex is not known to better

than the 'diamond' size, which has an r.m.s. of about 5 cm along the beam line, it is thus necessary to measure this angle with the calorimeter. Such a measurement requires at least two segments in depth and a sufficiently long lever arm.

In ATLAS detailed simulations as well as prototype studies have shown that the required accuracy (60 mrad/$\sqrt{E}$) can be obtained combining tower readout (of size $\Delta\eta = 0.025$) for the shower core, with thin strip readout for the shower upstream part (first 5 X_0). At 50 GeV E_{T} and $\eta = 1.3$ (end of the barrel) the resolution of (60 mrad/$\sqrt{E}$) corresponds typically to 0.5 mm accuracy in the strips, 1.2 mm in the shower core, for a lever arm of about 18 cm. In order not to increase too much the number of readout channels, the strips have a coarser ϕ granularity (0.1 compared to 0.025 in the second compartment), implying a loss of ϕ accuracy. However, in this direction (azimuth) there is no need to measure the angle since the vertex transversal size is very narrow, giving an excellent direction from the vertex and impact point in the second segment of the calorimeter.

The solution retained by CMS is to use either an independent device in front of the calorimeter (preshower) as the crystal calorimeter is not segmented in depth, or to use the high transverse momentum charged particles in the event to find the vertex position.

A further use of a good impact point accuracy in ϕ is for the so-called 'bremsstrahlung fit'. Using this point in the trajectory fit improves significantly the amount of tails in the impulsion measurement (see Section 4 on tracking).

3.2.3 Particle identification

The particle identification is performed using both the electromagnetic and hadronic calorimeters on the basis of lateral and longitudinal shower profiles. The demand is quite different depending on the desired physics process. This can be illustrated by two examples taken from LHC requirements:

- At the LHC the performance in electron identification is best gauged by comparing the 'prompt' electron production rate to the jet rate. For transverse momenta of 20 GeV/c or above, the jet-to-electron ratio is about 40 000. Calorimeter cuts based on the longitudinal and transverse segmentation of ATLAS (including towers of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the hadronic calorimeter) allow the reduction of the jet rate by a factor about 1500. Further rejection requires use of the central tracker to allow for track-shower match (E/p and position), lower energy companion finding (against Dalitz pairs and conversions), and possibly Transition Radiation identification. The ultimate performance expected in ATLAS is a combined rejection of more than 3×10^5, allowing in principle to filter out single-electron inclusive samples.
- The most difficult problem is actually the γ-jet identification. The required performance is driven by the search for the Higgs boson in the $\gamma\gamma$ mode, for masses between 80 and 160 GeV. Backgrounds from misidentified jet–jet and γ–jet events add up to the irreducible $\gamma\gamma$ background and therefore reduce the sensitivity. Rejections of 2×10^7 and 8×10^3, respectively, are desirable (to maintain the reducible/irreducible background ratio below 10%), where the last figure is known as the rejection per 'leg'. At 40 GeV E_{T} the calorimeter cuts provide a rejection of about 3000 (against 1500 at 20 GeV, see above). The remaining background is dominated by jets fragmenting to a leading π^0. The rapidity gap between the two γ's from such a π^0 decay peaks around 0.006 (about 9 mm at $\eta = 0$), calling for a correspondingly high granularity. This

is achieved in ATLAS with the narrow strips already described. The performance obtained from full shower simulation is plotted in Figure 25. This indicates that a rejection factor close to 8000 is within reach.

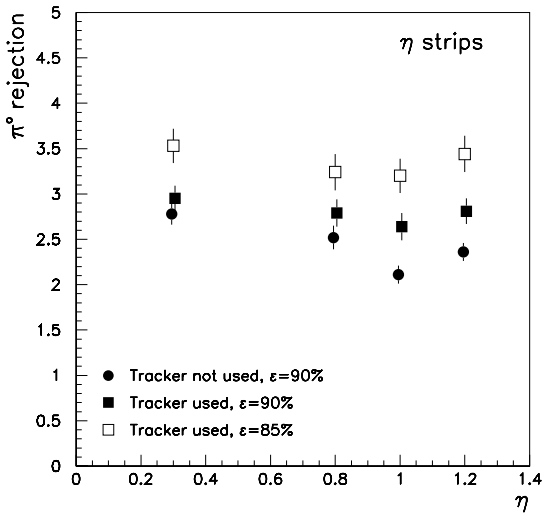

Figure 25: *Rejection of π^0 using the thin strip section of the ATLAS e.m. calorimeter [1].*

3.3 Hadronic interactions

In contrast with electromagnetic showers described by a cascade of $e^{\pm}$ and γ interactions, the hadronic shower development is more complex resulting from inelastic hadronic interactions which give rise to a cascade of various particles as illustrated by Figure 26. A few hadronic shower characteristics can be distinguished:

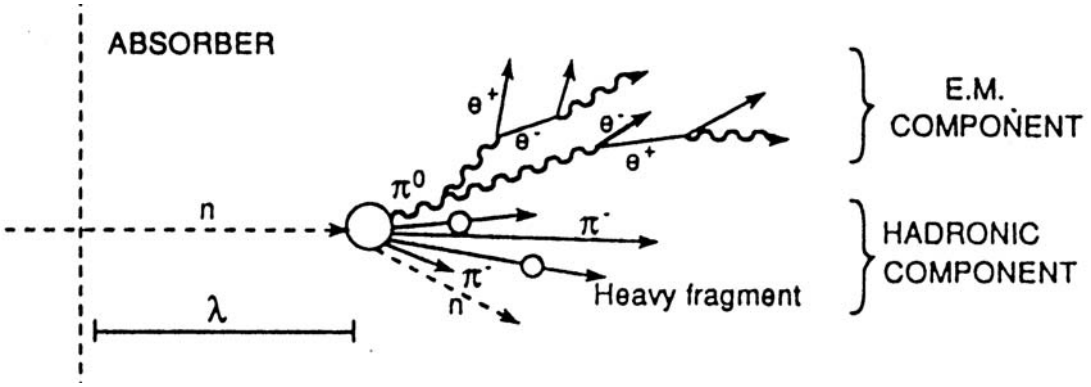

Figure 26: *Hadronic shower development.*

– On average 30% of the incident energy is released as electromagnetic energy at 10 GeV, resulting mostly from $\pi^0, \eta \rightarrow \gamma\gamma$. This fraction increases with energy up to 60% for 150 GeV charged pions [18] and has large fluctuations from one shower to another.

- Muons and neutrinos, produced mainly from pion decays, escape the detection (1% of the incident energy at 40 GeV).
- A large fraction of the remaining energy, whose origin is from nuclear break-up (strong interaction), is either not or poorly detected, and is dissipated in a variety of mechanisms.

  First the binding energy used to break up the nuclei is lost and amounts to a few per cent of the incoming particle energy (Figure 27) [28]. A part of the energy loss may be recovered when neutrons are captured by other nuclei and give rise to a visible signal.

  Secondly the non-relativistic protons, α's, and heavy fragments produced suffer from energy loss saturation in the scintillator or recombination effects in liquid argon.

  Finally a large number of neutrons are produced in the break-up and spallation process as displayed in Figure 27. As they can lose their kinetic energy only by strong interaction, they can travel far from the particle axis before being absorbed with a decay time as long as 0.5 μs. Some of them thus escape either the time or the space window used for the measurement.

  The fraction of non-detected energy can be as large as 40% of the non e.m. released energy, with large fluctuations which dominate the energy resolution.

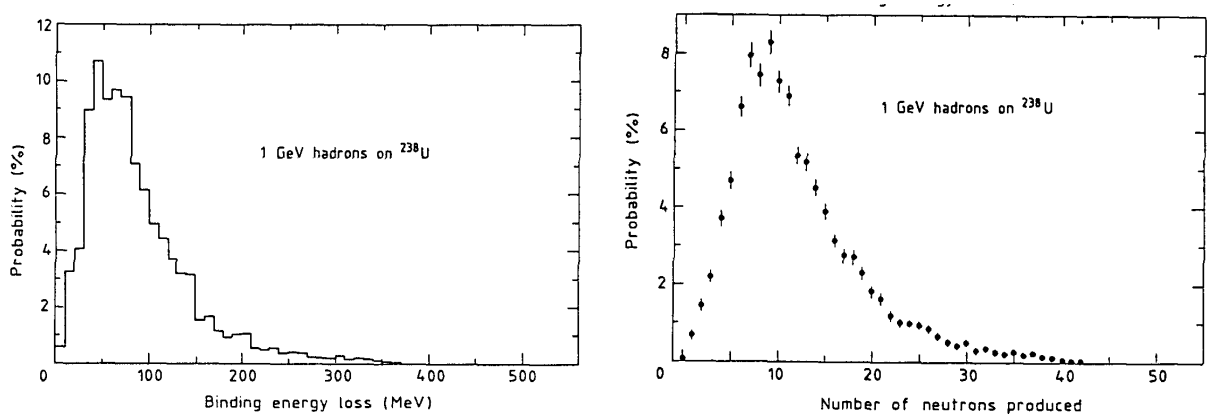

Figure 27: *Binding energy loss and number of neutrons produced in 1 GeV hadronic interaction in uranium [28].*

In order to describe hadronic shower shapes independently of the calorimeter material, it is convenient to introduce the interaction length[2], λ, which is the mean free path between two inelastic nuclear interactions:

- The longitudinal shower size increases slowly with energy with about 10 λ being required in depth for a 200 GeV hadron (Figure 28) [17].
- 95% of the energy is contained in approximately a 1 λ radius. The lateral shape presents a narrow core due to the e.m. fraction of the shower development (about two Moliere radii) and a broad hadronic tail (Figure 28) [18].

[2] $\lambda = A/(\sigma N_0 \rho) \propto A^{1/3}$. Typical values are λ = 15, 17, 17 cm for Cu, Fe, and Pb, respectively.

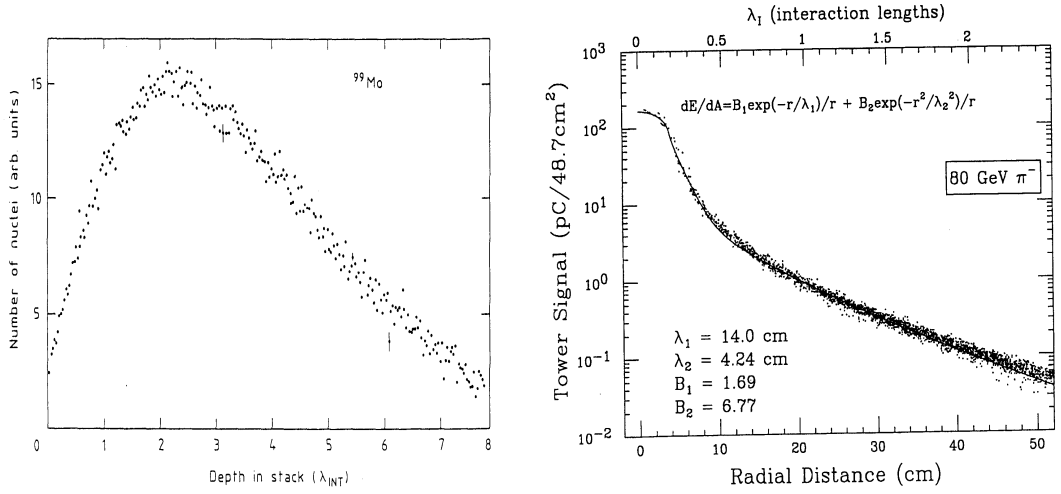

Figure 28: *Longitudinal (300 GeV) [17] and lateral [18] hadronic shower profiles.*

Hadronic calorimeters should clearly be much more massive than their e.m. counterpart (typically 2 m of iron as opposed to 15 cm of lead equivalent), imposing sampling calorimetry as the only affordable solution.

3.4 Energy resolution and limitations in hadronic calorimeters

As for electromagnetic calorimeters, one can parametrize the resolution by the relation defined in Section 2.2.1. The first term, a, is the quadratic sum of sampling fluctuations and intrinsic resolution. The sampling fluctuation of hadronic showers is about twice the one obtained for e.m. showers with the same calorimeter. However, this contribution is small and the dominant one is the intrinsic resolution with its large fluctuation of the energy loss and the different behaviour of response of the various components of the shower. This is related to the deviation from unity of the e/h ratio defined as the ratio of the e.m. and non e.m. components of an hadronic showers.

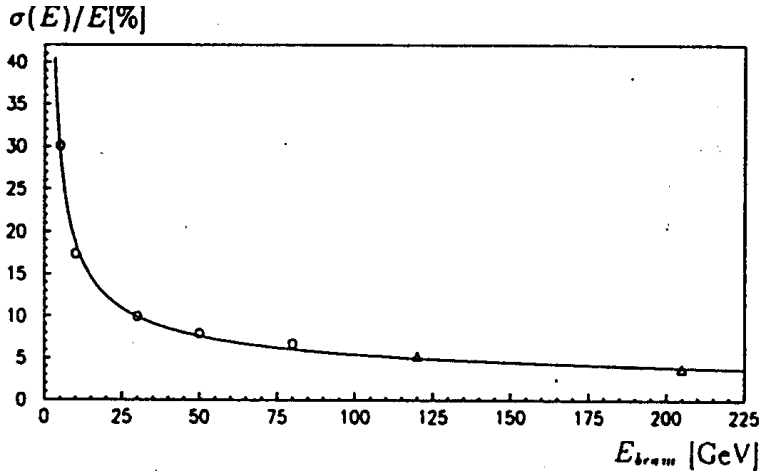

Figure 29: *Energy resolution of the H1 calorimeter [29].*

The best energy resolution observed is about $40\%/\sqrt{E}$ but this deviates from the $1/\sqrt{E}$ law, becoming worse at high energy. Two ways have been considered to improve the hadronic calorimeter resolution. By software one can try to estimate event by event the e.m. content of the shower and apply weights. This allows the restoration of the $1/\sqrt{E}$ behaviour but the energy response linearity is difficult to maintain. As an example, the energy resolution of the sampling Fe/LAr calorimeter of H1 is displayed in Figure 29 with charged pions [29].

By the use of hardware one could try to reduce the fluctuations by being more sensitive to the nuclear binding energy loss and to make e/h $\simeq$ 1. Various possibilities have been investigated:

- Use of U^{238} fissile material as absorber in order that a part of the energy lost in nuclear break-up induces neutrons by fission and is recovered. Indeed the kinetic energy of the neutrons is strongly correlated to the binding energy loss (see Figure 27).
- Use of plastic scintillators as active layer because of their high sensitivity to neutrons by elastic scattering on protons. This is illustrated by Figure 30 in which the e/h ratio is presented as a function of the ratio of the absorber thickness to the active thickness [17]. In liquid argon or with silicium, even with a 1 μs gate, one cannot achieve e/h = 1 with a reasonable uranium thickness. On the contrary, with scintillators (PMMA, SCSN) or warm liquid (TMP) rich in hydrogen, one can make the electromagnetic and hadronic response equal. Such a scheme was adopted in the ZEUS sampling calorimeter [30] with uranium/scintillator: the e/h ratio is about 1 as displayed in Figure 30 and the resolution for hadrons is good, about $35\%/\sqrt{E}$. The drawback is a modest electromagnetic energy resolution limited to $18\%/\sqrt{E}$ because of the too small sampling fraction and large absorber thickness dictated by the e/h requirement.

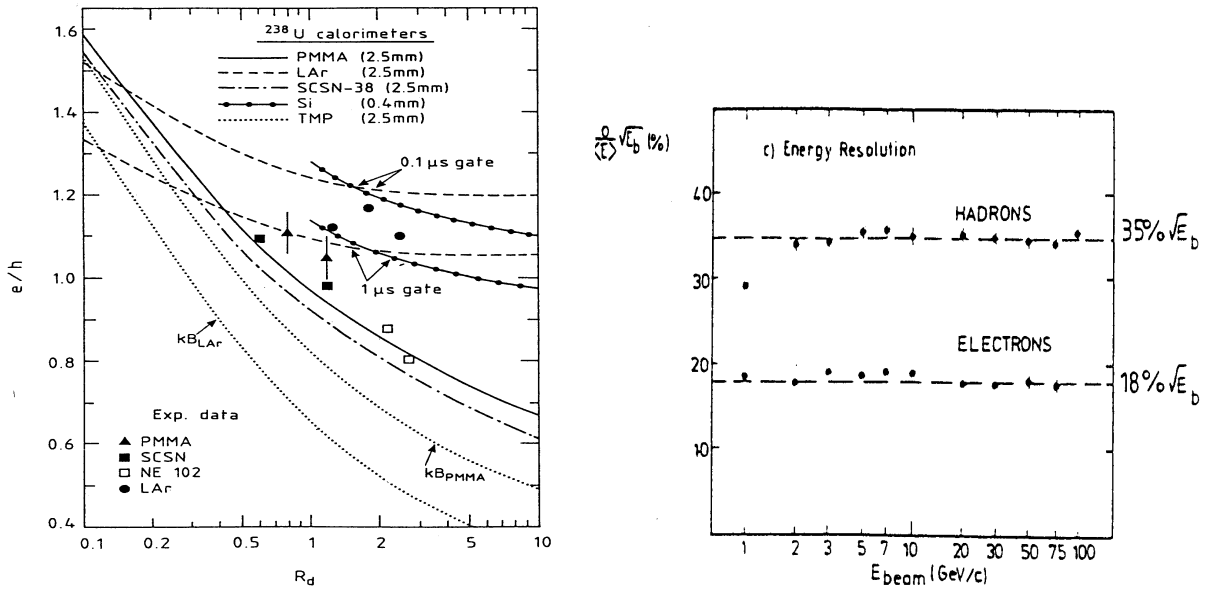

Figure 30: *e/h ratio for various active media [17] (left), and energy resolution of the ZEUS calorimeter [30] (right).*

The second term includes the electronic noise, the pile-up at the LHC (6 GeV for a cone size of $\Delta R = 0.5$ at high luminosity) and an additional cone size effect as

demonstrated in Figure 31. The constant term includes all the other effects, especially instrumental (mechanical and electronic), and is of the order of a few per cent. In general results shown are from pion beams, but in a real experiment the important feature is the response to jets for which the calibration and energy scale are different and have to be tuned in situ. For this purpose the CDF experiment uses two-jet events, for which the energy of one jet is mainly electromagnetic, and applies the transverse momentum conservation [31].

The two LHC experiments have chosen hadronic calorimeters without hardware compensation. The ATLAS central hadronic calorimeter is made of large scintillating tiles based on a sampling structure with interleaved steel absorber and scintillating plates read out by wavelength shifting fibres (WLS). This structure is sketched in Figure 32. With weights on the different compartments in depth (including the liquid argon calorimeter in front), the energy resolution with pions obtained in a test beam is $45\%/\sqrt{E} + 1.3\%$ which fulfils the LHC requirements of jet measurement. For rapidities larger than 1.2 and up to 5, the ATLAS hadronic calorimetry uses the liquid argon technique, mainly on account of its superior radiation hardness.

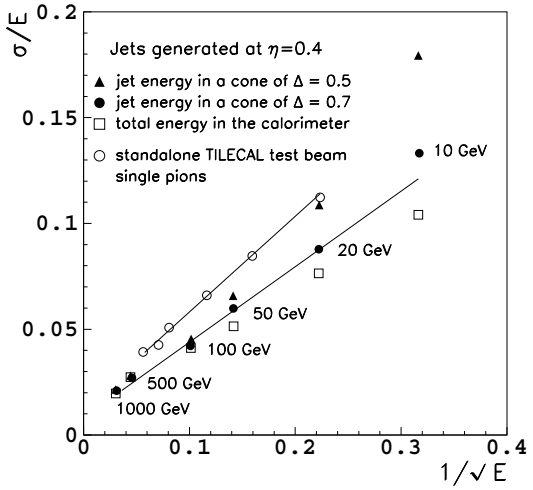

Figure 31: *Fractional energy resolution of single-jet simulation for various cone sizes with the ATLAS barrel calorimeter [1].*

The CMS calorimeter is based on the same concept, but as it is contained inside the solenoid the space is limited leading to the use of copper which is non-magnetic and has a shorter interaction length than steel. Nevertheless, as the depth is too thin in the central region, it is followed outside the solenoid by a tail catcher as sketched on Figure 32. The expected performance for jets is about $60\%/\sqrt{E}$ with a 3% constant term.

3.5 Examples of homogeneous calorimeters

3.5.1 Semiconductor detectors

On account of their high cost, these detectors are not used in large high-energy physics experiments but only as γ counters in nuclear physics. Their performance is interesting because they reach better resolution than the naive expectation. Indeed, in these

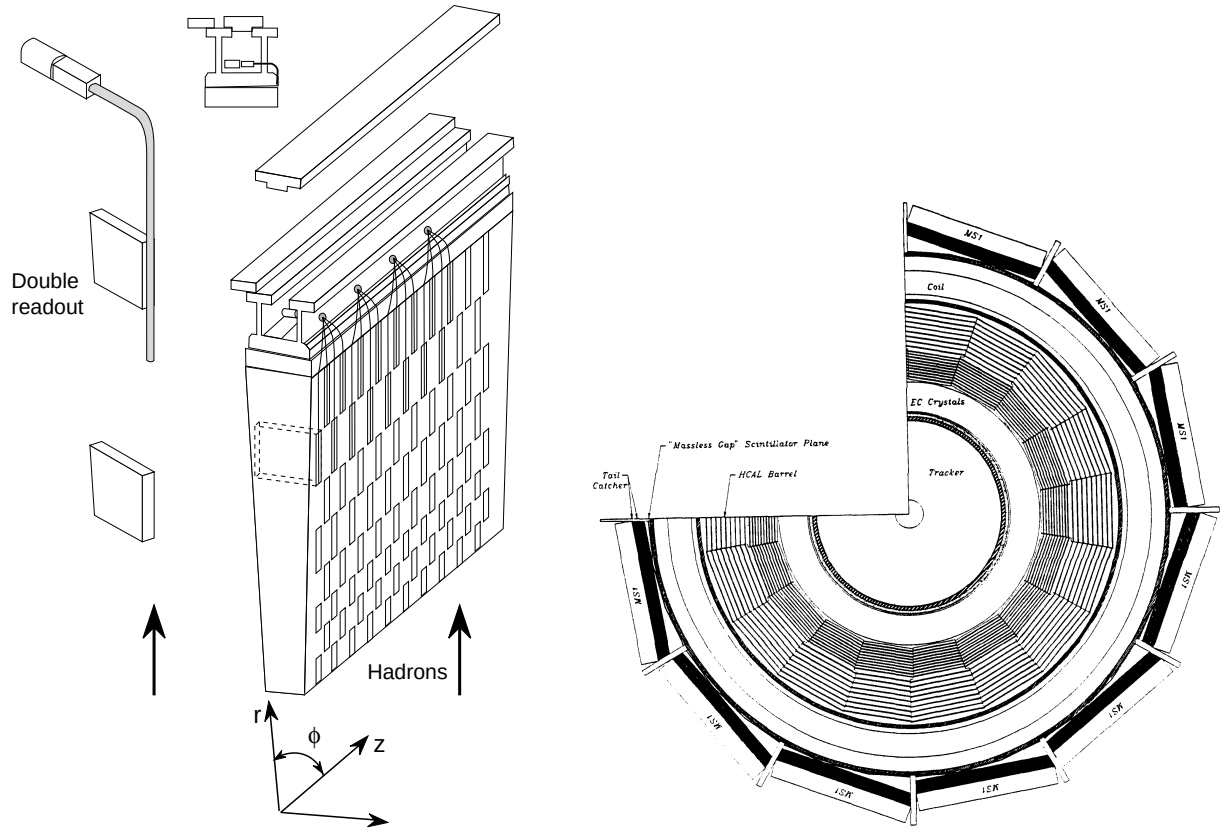

Figure 32: *Principle of the tile hadronic calorimeter of ATLAS (left), view of the barrel hadronic calorimeter in CMS inside the solenoid (right).*

devices the full energy of the radiation is absorbed and 'does not fluctuate' as opposed to sampling calorimeters for instance. The total number of ionizations which can occur leading to electron-hole pair production (the energy needed to create an electron-hole pair is 3.6 eV in Si and 4.2 eV in Ge) and the energy lost are thus constrained by the total energy. This means that these electron-hole pairs are not statistically independent and the naive assumption of Poisson statistics is incorrect. The energy resolution is induced not by the fluctuation on the total number of electron-hole pairs ($\sqrt{N}$) but by $\sqrt{FN}$ so that the energy resolution is given by:

$$\frac{\sigma(E)}{E} = \frac{\sqrt{F}\sqrt{N}}{N} = \sqrt{F}\sqrt{\frac{\varepsilon}{E}} \, .$$

The factor F, called Fano factor, is a measurement of the correlation which exists between each pair-creation. Figure 33 [32] shows the comparative pulse height spectra of an Ag γ source with a Ge detector and a NaI(Tl) scintillator. The resolution is obviously better with the Ge semiconductor which benefits from a smaller pair creation energy and from the Fano factor (0.13) (this is difficult in NaI(Tl) where the collection efficiency can not be 100%): typically 0.1% at 1 MeV against 8% with the crystal! The high resolution is, however, lost as soon as energy leakage or an instrumental bias appears and for this reason semiconductors are not used in high-energy experiments as calorimeters.

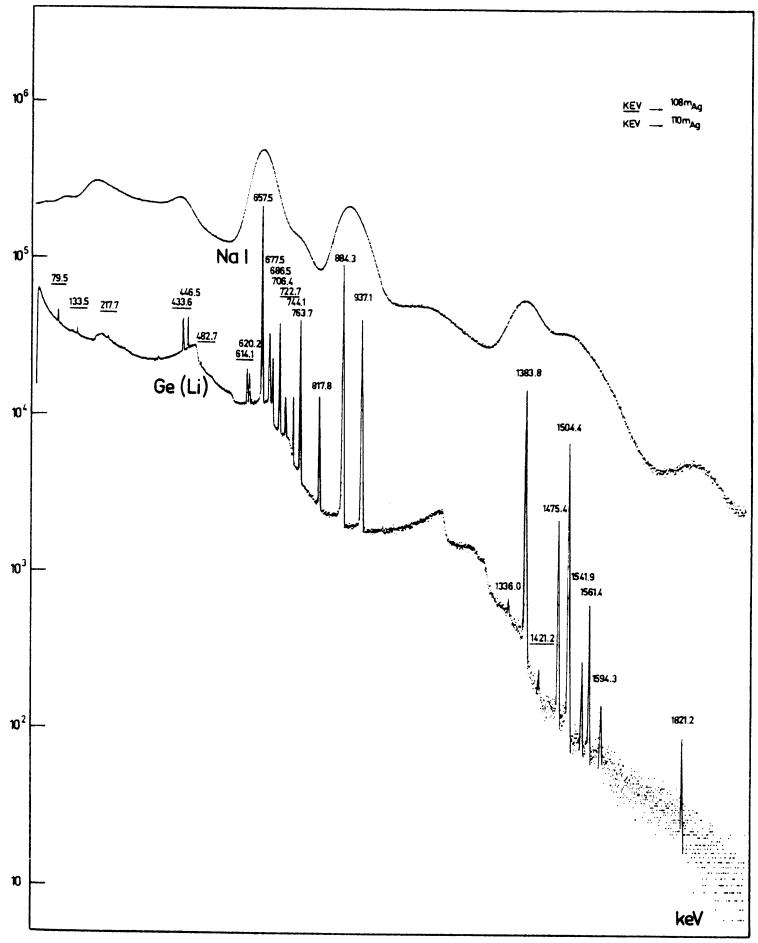

Figure 33: *Pulse height spectra of a Ag γ-source with a Ge and a NaI detector (peaks are labelled in keV) [32].*

Table 1: *Properties of noble liquids*

	Ar	Kr	Xe
Z	18	36	58
A	40	84	131
X_0 (cm)	14.0	4.7	2.77
E_C (MeV)	41.7	21.5	14.5
R_M (cm)	7.2	4.7	4.2
W (eV/pair)	23.3	20.5	15.6
dE/dx (mip) MeV/cm	2.11	3.45	3.89
Drift speed at saturation (mm/μs)	10	5	3

3.5.2 Homogeneous noble liquid calorimeter

The noble liquid detectors currently in operation are mainly based on charge collection in sensitive cryogenic liquids (Ar, Kr, Xe). The main characteristics of these liquids

are summarized in Table 1. It should be noted that the electron/ion pairs do not all contribute to the ionization signal: in the absence of an electric field about half of these pairs recombine and give scintillation light by molecule deexcitation with a signal component which can be as fast as 10 ns. The correlation which exists between these two signals in argon has already been demonstrated and is displayed in Figure 34 [33]. Collecting all of the ionization signal already gives good energy resolution owing to the Fano factor. Assuming the total energy being absorbed, $N = N_{\rm ion} + N_{\rm scin}$, the fluctuation on the total collected ionization signal $N_{\rm ion}$ is $\sqrt{N(N_{\rm ion}/N)(N_{\rm scin}/N)} = \sqrt{N_{\rm ion}(N - N_{\rm ion})/N}$. Taking for instance $N_{\rm ion}/N = 0.90$ leads to an improvement by a factor 3 of the resolution with respect to the naive Poisson expectation $\sqrt{N_{\rm ion}}$. Collecting also the scintillation light could give excellent energy resolution.

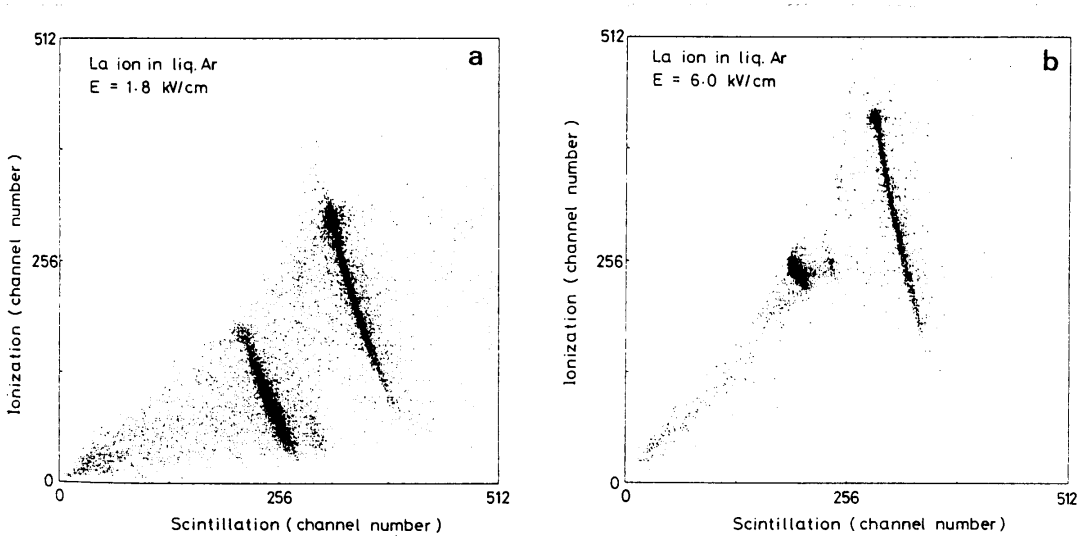

Figure 34: *Correlation between scintillation light and ionization in liquid argon for two operating high voltages [33].*

For high-energy particles, xenon is obviously the best candidate with the smallest critical energy (better resolution) and a short radiation length (2.77 cm) which would lead to a compact calorimeter (70 cm for a 25 X_0 depth). However, no large-sized liquid xenon detector has been built up to now because of the prohibitive cost and the high sensitivity to impurities.

The use of krypton was first investigated by the KEDR group [34] which required a very good energy resolution at low energy (500 MeV), and obtained 2% at 1 GeV. The NA48 experiment, designed for a high-precision measurement of CP violation in the kaon system [35], is building a large-sized (quasi) homogeneous krypton calorimeter (Figure 35) [36]. Electrodes are oriented along the incoming particle direction defining 2×2 cm^2 towers. In order to reduce the dependence of the signal of drifting charge with the distance between the core of the shower and the electrodes, especially a loss of signal very close to them, they have been given a slight accordion shape. High resolution ($a = 3.5\%$), good uniformity ($< 0.5\%$), operation at high speed (40 MHz) to minimize the accidental activity, and fine granularity to separate close-by photons from π^0 decay, are the major requirements of this calorimeter. The main limitation comes from the large

sensitivity to the geometrical imperfections in the charge collection structure (constant term). Figure 36 presents the energy resolution of the NA48 prototype measurements [36]. The stochastic term is mainly dominated by the lateral fluctuation (which contributes to $2.5\%/\sqrt{E}$), the energy being reconstructed in a 11×11 cell area (2 Moliere radii) resulting from an optimization of the noise contribution and shower separation. The use of cold Si JFET preamplifiers, well-suited for operation at liquid Kr temperature (120 K), and the small detector capacitances (80 pF) lead to a rather low noise ($\simeq$ 7 MeV/cell) even for a fast shaping (peaking time of the calorimeter signal of about 80 ns). The uniformity of ionization signals is good enough to give a timing accuracy of 300 ps at 10 GeV.

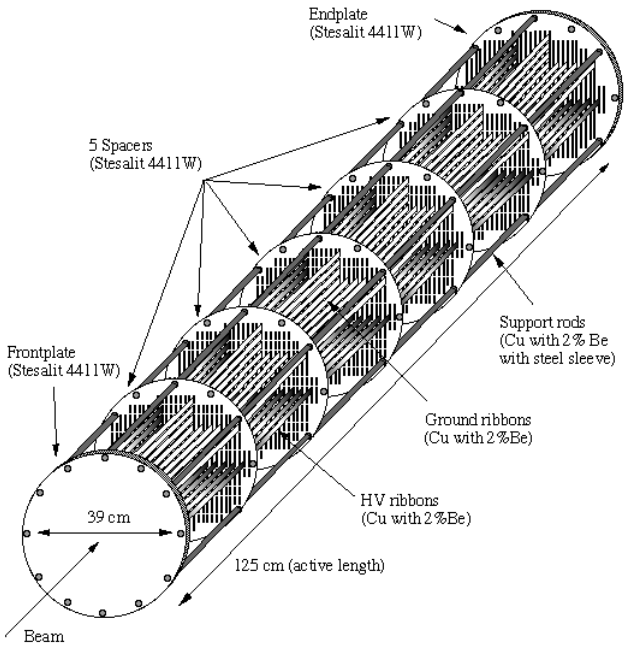

Figure 35: *Drawing of the NA48 calorimeter prototype.*

The homogeneous liquid argon calorimeters are not adapted to experiments at colliders, as a 25 X_0 calorimeter would be 3.5 metres deep. Compared with Xe and Kr, Ar is low cost and easier to handle because it needs less purification. The ICARUS Collaboration is developing a huge liquid argon homogeneous calorimeter (4400 m^3 of argon), designed for proton decay and neutrino physics. The aim is a very good energy resolution ($3\%/\sqrt{E}$) [37] and 'imaging' properties obtained by measuring time and position after charge drift over a long distance as in a TPC (see Section 4.1).

3.5.3 Crystal calorimeters

The use of crystals is motivated by the possibility to reach a stochastic term as good as $2\%/\sqrt{E}$ by collecting the scintillation light. This light results from various complex mechanisms which induce an emission spectrum with different components and structure which are very material dependent. For a large system the limitation comes from the

constant term as illustrated by the performance of the L3 BGO crystal calorimeter (10 000 crystals with a 2×2 cm^2 transverse size). The test beam resolution on prototype was $1.5\%/\sqrt{E} \oplus 0.4\%$ while the energy resolution was only 1.4% for Bhabha events at LEP ($\sim$ 45 GeV) meaning a larger constant term (Figure 37) [38].

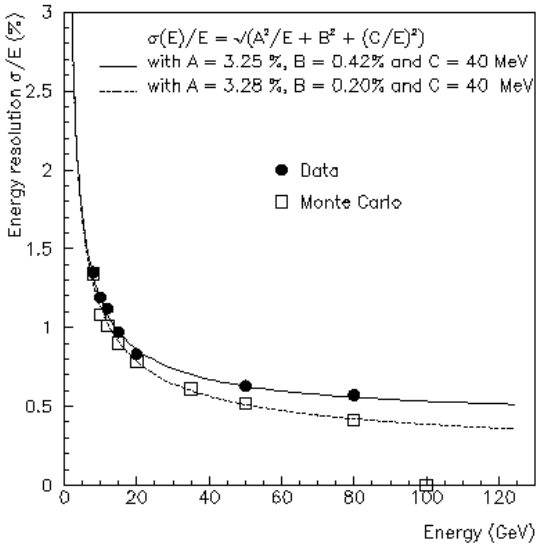

Figure 36: *Energy resolution of the NA48 calorimeter [36].*

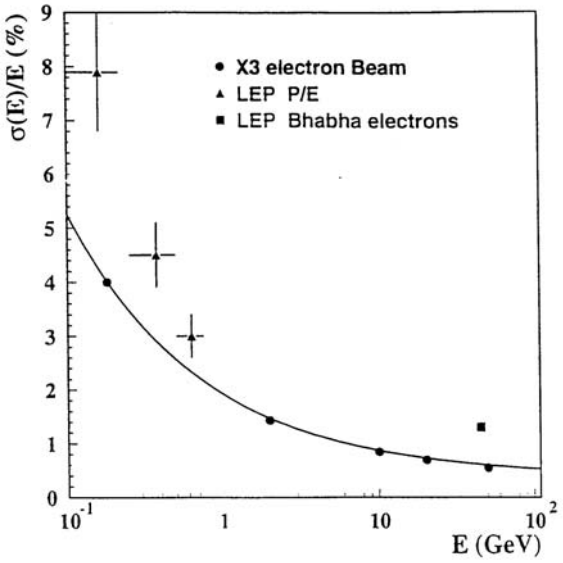

Figure 37: *Energy resolution of the L3 calorimeter in a test beam and at LEP with Bhabha events [38].*

The main characteristics of crystals already used or planned for high-energy physics experiments are listed in Table 2. The choice for an experiment is driven by practical considerations more than by intrinsic performances, among them:

– Cost.

- Mechanical stability, density, temperature dependence.
- Speed of response.
- Radiation hardness.
- Matching between emission wavelength and the photodetector device.
- Efficiency of collecting devices (photomultipliers, photodiodes etc.).

Table 2: *Properties of crystals*

	NaI(Tl)	CsI(Tl)	CsI	CeF_3	BGO	BaF2	$PbWO_4$
Density ($g.cm^{-3}$)	3.67	4.51	4.51	6.16	7.13	4.89	8.28
Radiation length (cm)	2.59	1.85	1.85	1.68	1.12	2.06	0.85
Moliere radius (cm)	4.8	3.5	3.5	2.63	2.33	3.39	2.2
Emission peak (nm)	410	560	420 310	340 300	480	300 220	450 420
Decay time (ns) fast and slow if any	230	1250	35 6	30 9	300	620 0.9	36 <10
Hygroscopic	yes	slightly	slightly	no	no	no	no
% of signal/K	0.22	0.1	0.1	0.15	−1.6	−2.0	−1.9
Light yield γ/MeV	4×10^4	5×10^4	4×10^4	2.10^3	8×10^3	10^4	1.5×10^2
Relative yield in pe	100	45	5.6 2.3	6.6 2.0	9	21 2.7	0.3 0.2
Radiation hardness (rad)	10^2	10^3	10^5	10^7	10^2	10^3	10^7

The crystal light yield quoted in Table 2 is the total intrinsic light yield. What is really measured in a detector is the convolution of this number with the geometrical, optical, time and quantum efficiencies of the collecting device which turns out to be the limitation on energy resolution. For instance in the table is also listed the signal yield of various crystals (of the same size) relative to NaI(Tl) and measured with a photomultiplier (bialkali cathode). It should be noted that this γ detector is best suited for NaI(Tl) but not for all, so that these numbers are only indicative. Another illustration of these parameters is also given by the Figure 38 where the signal pulse-height and the photoelectron yield as a function of the integration time are displayed [39].

The first crystal, NaI(Tl) (thallium doping), is one of the better know because it has been used for a long time. A resolution near the intrinsic limit was obtained around 1 GeV but the behaviour of the energy resolution was believed not to scale as $1/\sqrt{E}$ but rather as $E^{-1/4}$ until recently when this behaviour was explained by energy leakage. Its hygroscopic nature and its relatively low density do not make it a good candidate for new experiments although it is probably the cheapest one.

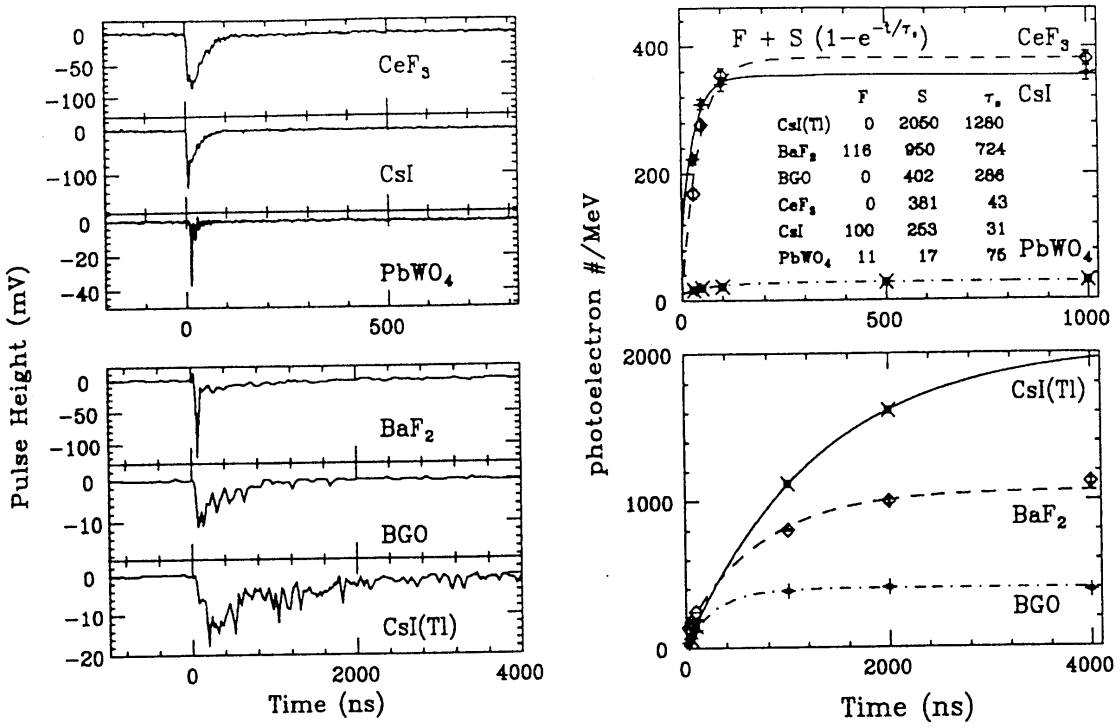

Figure 38: *Scintillation pulse of Cs γ-ray (left) and number of photoelectrons per MeV as a function of the integration time for various crystals (right) (from Ref. [39]).*

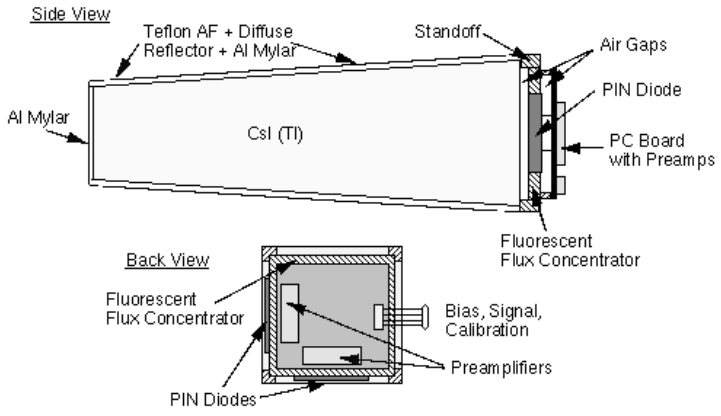

Figure 39: *A crystal from the BaBar calorimeter with its associated readout.*

The CsI(Tl) and CsI are very popular crystals which have been chosen for many experiments for spectroscopy and B physics (Crystal Barrel [40], Cleo II [41], BaBar [42], KEKB [43], etc.). The pure CsI presents a fast response and gives good energy resolution even if the signal yield of the fast components is not very large. It will equip the KTeV [44] calorimeter which has the same physics requirements as the NA48 calorimeter. The use of Tl as a dopant improves the light yield by displacing the luminescence peak but results in a much slower speed of response. This crystal is chosen by both B factory detectors with

the aim of a resolution of $1\%/\sqrt{E} \oplus 1\%$. The mechanical assembly of these detectors with their projective geometries (all crystals are not identical) is a difficult job, and it is not easy to maintain a good uniformity. Figure 39 shows a view of one crystal of the BaBar calorimeter. It is equipped with two Si PIN diodes for reasons of reliability [42].

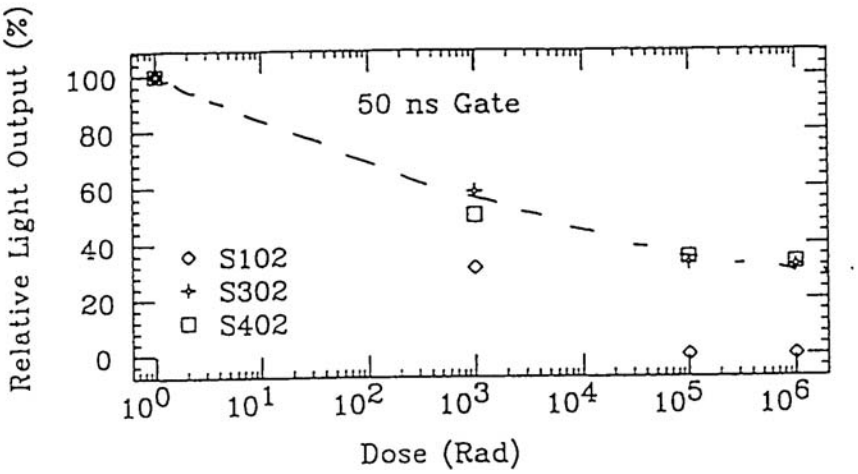

Figure 40: *Radiation hardness of BaF_2 [45].*

The BaF_2 crystal has a very fast response and a good yield. It was considered as a candidate for the GEM experimental proposal at SSC. However, as displayed in Figure 40 its radiation hardness was barely sufficient at SSC [45], and would not be suited for the LHC environment.

A complete investigation of crystals which fulfil the LHC calorimetry requirements, was performed by the Crystal Clear Collaboration [46]. CeF_3 appeared to be the best candidate by its intrinsic characteristics; however, external constraints, especially cost reasons and delay of manufacture, have led the CMS experiment to choose the $PbWO_4$ crystal. The drawbacks of this crystal are a high sensitivity to temperature variation, 1.9% per degree, and a poor light yield thus needing a very efficient light collection. PIN diodes do not produce enough signal and the interaction of charged particles in the silicon could be catastrophic (see later), photomultipliers cannot be used in the high magnetic field (4 T) and radiation environment so that avalanche photodiodes (APDs) have been considered. A schematic drawing of such a device is shown is Figure 41.

These components grown by epitaxy present a very thin high-gain region (a few μm) in which is created a strong electric field (10^5 V/m). This layer is localized behind the entrance layer where the light is converted to electron-hole pairs. The output signal can be amplified by a factor up to 50 by an avalanche process in the high-gain region. New measurements have been made recently [47]. Figure 42 shows a quantum efficiency of 65% over the range of the $PbWO_4$ emission spectrum. These devices operate in a high magnetic field environment and are not too sensitive to radiation as demonstrated in Figure 42. The noise increase results directly from a decrease of the gain and quantum efficiency and an increase of the leakage current. APDs also show a dependence on temperature: working at a temperature around 13–14 °C and applying filtering might reduce the noise by a factor of about 3.

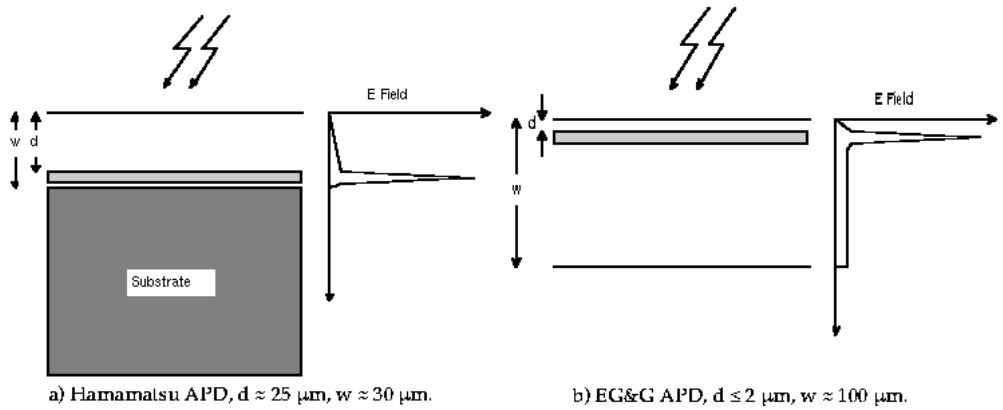

Figure 41: *Schematic of various APDs [2].*

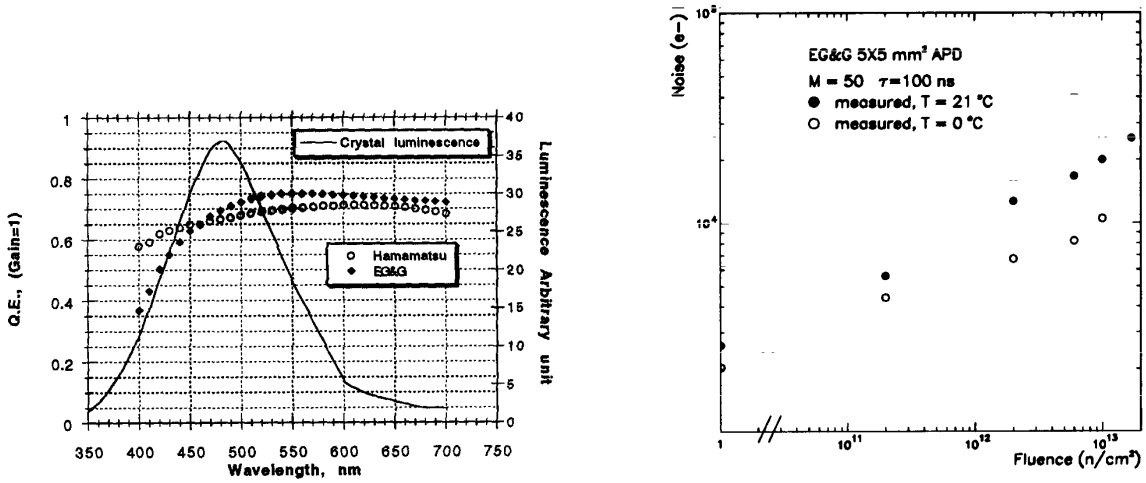

Figure 42: *Quantum efficiency of APD vs wavelength (left), noise as a function of neutron flux (right) [47].*

Progress is being made in order to reach the CMS calorimeter requirements:

- The gain sensitivity to temperature (3% per degree) and bias voltage variations (5% per volt with a 90 pF capacitance up to 16% with a 340 pF capacitance) have to be reduced.
- The excess noise factor, f, due to the statistical nature of the multiplication process which increases for high gain, has to be maintained at less than 2. This behaviour induces fluctuations on the measurable signal and degrades the sampling term as $\sqrt{f}/N$ where N is the number of photoelectrons per GeV ($N = 900$)[3].
- To minimize the nuclear counter effect, the avalanche region has to be as close as possible to the surface behind the sensitive layer. This nuclear counter effect is due to charged particles (for instance μ or rear leakage) which cross through the APD and give rise to electron-hole pairs (~ 85 for a minimum ionizing signal in 1 μm of silicon). If these charges are created in front of the avalanche region, they are amplified and produce high-energy tails in the energy spectrum.

[3] This is again an illustration of the difference between total light yield, 150 photons produced per MeV, and signal yield, only 0.9 photoelectron per MeV in this particular case!

The best performance obtained so far by a 36-crystal CMS prototype readout with APDs is presented in Figure 43, with an energy resolution of 0.55% at 120 GeV. The typical noise is about 250 MeV when nine crystals are used to reconstruct the energy.

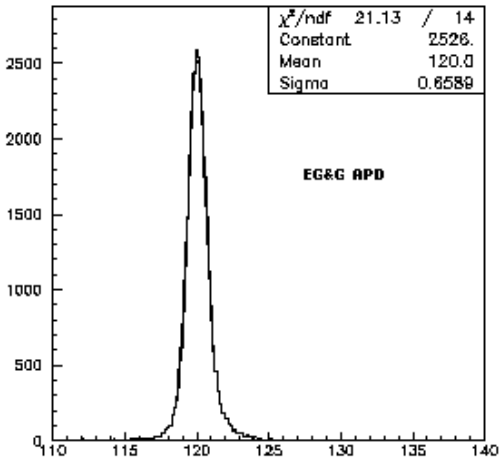

Figure 43: *CMS prototype energy resolution with APD readout [47].*

The final calorimeter will contain an order of magnitude larger number of crystals (100 000) as compared to L3 so the desired 0.5% constant term over the full calorimeter is quite challenging. A comparison of its expected overall performance with the ATLAS calorimeter on the H$\rightarrow \gamma\gamma$ benchmark is discussed at the end of this section.

The ALICE [48] experiment is also going to use $PbWO_4$ crystals for its photon spectrometer part (covering a solid angle of 0.7 steradian at a distance of 4 metres from the interaction point). The choice in that case is to operate at -25 °C, which increases the light yield by a factor 2.5, and allows the use of PIN diodes. The nuclear counter effect is avoided by a charge particle veto in front of the calorimeter. The anticipated noise is 15 MeV per crystal for a shaping of 1 μs or so, compatible with the ion-mode luminosity of the LHC.

3.5.4 Sampling e.m. calorimeters

This is a well-proven technology used in many large detectors. Charge or light is collected depending on the active medium used. Depending on the absorber thickness and active medium, energy resolutions in the range 5–10%/$\sqrt{E}$ have been obtained. Lateral granularity and longitudinal segmentation are in general easier with these calorimeters than with crystals.

A lead/scintillator electromagnetic sampling calorimeter was developed in the framework of the CMS Collaboration (shashlik calorimeter [49]). The light is read out at the back of the calorimeter with WLS fibres running through holes perpendicular to the absorbers and scintillator plates. This light is converted into electron-hole pairs using PIN diodes. A good energy resolution was obtained even in a 3 T magnetic field as displayed in Figure 44 but the linearity is a little degraded at low energy (see Figure 21) [20].

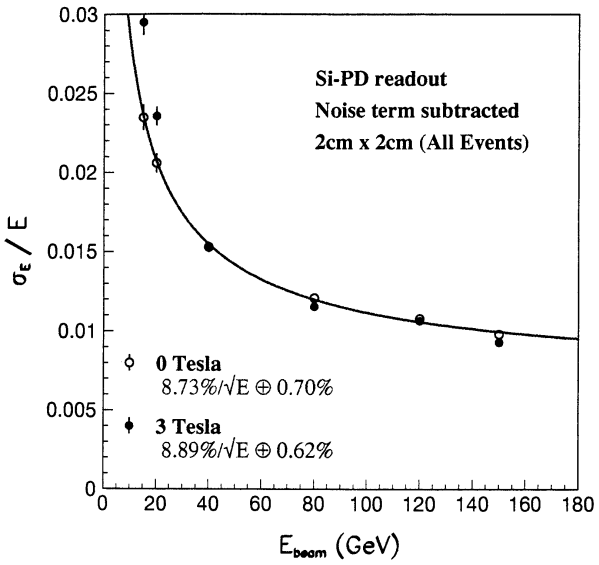

Figure 44: *Energy resolution of the shashlik calorimeter in high magnetic field [20].*

The ATLAS experiment has chosen a sampling lead/liquid argon e.m. calorimeter with an accordion-shaped geometry. The absorber plates are pleated around radial planes, as can be seen in Figure 45. This minimizes the cracks between modules and allows a fast signal propagation up to the preamplifiers [50].

Figure 45: *Picture of accordion prototype during its construction.*

As in any ionization sampling calorimeter, electrons from pairs created by charged particles drift towards the positively charged electrode. A point-like charge at a distance x from the electrode will induce a uniform current $i = q\,V_{\rm D}/g$ during a time $t = x/V_{\rm D}$ where $V_{\rm D}$ is the electron drift field in argon, (about 10 mm/μs) and g the distance between anode and cathode). Charged particles crossing the gap will lead to a uniform distribution of 'point-like charges' over the argon gap. The resulting current, $i = i_0\,(1 - t/t_{\rm D})$, is thus triangular as shown in Figure 46. The current is maximum at the origin when all charges drift and disappears when all charges created have crossed the full argon gap. This time

is about 400 ns in the ATLAS calorimeter. The calorimeter sensitivity comes out to be about 3 μA/GeV. It should be noted that in the NA48 calorimeter where the electrodes are almost parallel to the particle direction, the ionization signal is not uniform in the gap (it depends on the lateral shower spread) leading to a deformation of this triangular shape.

A small current is also induced by the ion drift but it is negligible because of the low drift velocity of the ions, typically 1000 times less than electrons. One should, however, ensure that the ion density does not modify the applied electric field and thus the signal [51].

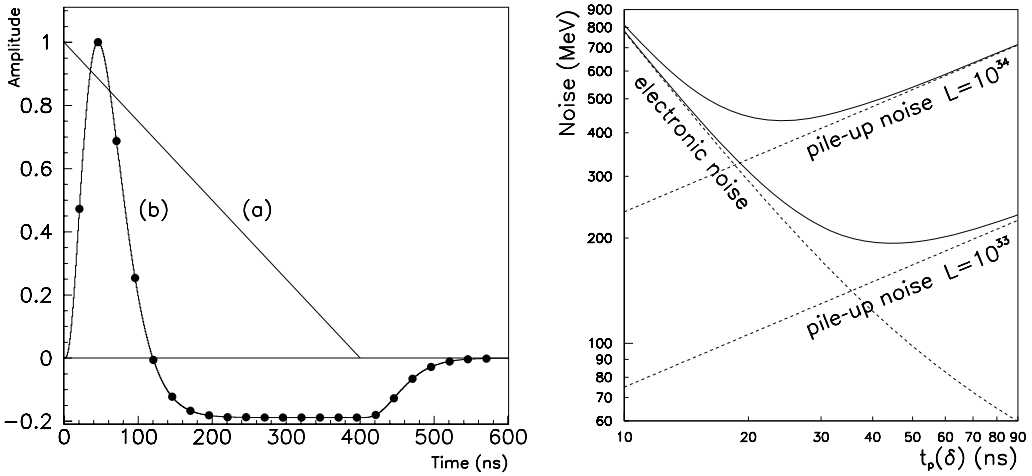

Figure 46: *Current (a) and shaper (b) output in the ATLAS calorimeter (left); electronics noise, pile-up noise and combined noise as a function of shaping time (right) [1].*

If one wanted to collect all the charge, the disadvantage of this technique would obviously be too low a speed (400 ns), not adequate for the LHC. However, the energy information is fully contained in the initial current, i_0, if it can be measured. This is solved with a signal clipping. In order to be sensitive to the initial current a $CR - RC^n$ unipolar filter, equivalent to a current derivative, is used and gives the shape of the ATLAS calorimeter signal displayed in Figure 46. It peaks at about 40 ns with a height proportional to the initial current.

The price to pay with respect to a total charge collection is an increase of the electronics noise which scales as $1/\sqrt{\tau}$ ($\tau = RC$ is the shaper time constant). Since the pile-up noise scales as $\sqrt{\tau}$, a trade-off has to be found as demonstrated by the optimization chosen in ATLAS on Figure 46 [52].

The electronics noise can be characterized by the spectral density of the preamplifier series noise generator, $e_{\rm n}$, and of the parallel noise generator, $i_{\rm n}$. The noise charge (ENC) which is the charge Q needed to produce a signal at the output of the preamplifier equal in magnitude to the r.m.s noise observed is defined by [53]:

$$ENC^2 = \frac{{e_{\rm n}}^2 C {I_{\rm a}}^2}{\tau} + {i_{\rm n}}^2 {I_{\rm b}}^2 \tau$$

- I_a and I_b are constant numbers depending only on n to first order.
- C is the capacitance at the preamplifier input dominated usually by the detector's capacitance.
- $i_n = \sqrt{4kT/R}$ where R comes from resistors on the input, T the temperature and k the Boltzmann constant. For fast shaping this noise contribution is dominated by the series noise.
- The series noise is linked to the amplification mechanism, $e_n^2 = 4kT\alpha/g_m$ where g_m is the transconductance of the input transistor and $\alpha \simeq 0.7$.

The best electronics performance so far has been realized with cold preamplifiers using GaAs MESFET technology. They feature low noise ($e_n = 0.25$ nV/$\sqrt{\text{Hz}}$) and dissipation (50 mW) and have a sufficient radiation hardness [54]. Warm electronics outside the cryostat (with a matched impedance cable) can also be used [55]. The reliability of the first one, as the preamplifiers are in the cold and could not be replaced during the experiment lifetime, ~10 years, is counterbalanced by the larger sensitivity to pick-up of the second one.

A typical electronics noise value (see Figure 46) in a cluster of 3×7 cells is 220 MeV. Including the pile-up, the total noise amounts to about 450 MeV.

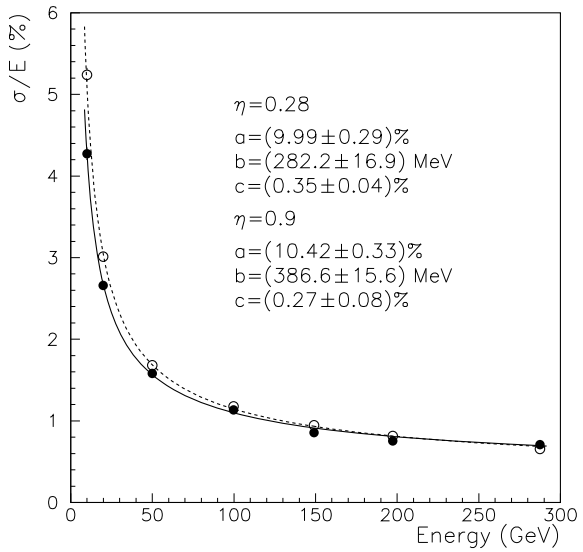

Figure 47: *Fractional energy resolution as a function of beam energy for two pseudorapidities.*

Different prototypes have been built and many tests in electron beams have been carried out [56]. To reconstruct the energy, a small correction is necessary to account for the variations of sampling fraction as a function of ϕ in the accordion direction. The fractional energy resolution obtained with a large-scale prototype (2 metres long, 27° in azimuth) [19] at two rapidities is displayed in Figure 47 and fulfils the ATLAS requirements. The study of the response of more than 150 cells over a large area has demonstrated that a 0.58% uniformity response can be reached, equivalent to a 0.7% global constant term[4]. A lot of work was done to understand in detail the various sources

[4] The cell-to-cell response has to be convoluted by the local constant term of the energy resolution to obtain the global constant term.

as listed in Table 3 with the expectation for the ATLAS calorimeter. These results and the possibility to use the large flux of $Z \rightarrow e^+e^-$ decay (mass constraint) or inclusive e^- (E/p matching) give confidence that a constant term $\leq 0.7\%$ can be reached in the full ATLAS calorimeter.

Table 3: *Contributions to the constant term (in per cent) measured in the prototype and expected in the ATLAS calorimeter*

Source	Prototype	ATLAS
Mechanics		
Residual ϕ-modulation	0.3	0.2
Gap non-uniformity	0.15	0.15
Absorber thickness	0.3	0.2
Calibration		
Amplitude accuracy	0.25	0.25
Timing precision	0.2	0
(Cross-talk effects)	(0.15)	0
Total (quadratic sum)	0.57	0.4

Up to now, the intrinsic performances of the calorimeters have been discussed separately. However, ultimately, it is important to examine the performance of the full detector for a given physics process. As an example the search for the Higgs boson by its $\gamma\gamma$ decay mode has been investigated with the CMS and ATLAS detectors. As a good illustration of the various effects previously described, Table 4 shows the size of the various contributions to the $\gamma\gamma$ mass in ATLAS. Owing to its better sampling and expected constant term, the $\gamma\gamma$ mass resolution of CMS is 1.6 times smaller which allows a narrower mass bin and a 30% improvement in statistical significance $S/\sqrt{B}$. However, as the CMS calorimeter operates in a 4 T field, only non-converted photons can be used with a good resolution, resulting in a 30% acceptance loss with respect to ATLAS, leading finally to comparable performances [57]. This shows the real difficulty in comparing the performance of a subdetector on the basis of testbeam results only, outside of its running environment.

Table 4: *Contribution to the $\gamma\gamma$ mass resolution at high and low luminosity with the ATLAS detector. At low luminosity the pointing is performed assuming that the longitudinal position of the vertex is determined using charged tracks of the underlying event.*

	High luminosity	Low luminosity
Sampling term (MeV)	900	900
Constant term (MeV)	490	490
Total noise term (MeV)	500	210
Angular term (MeV)	400	70
Total (MeV)	1210	1050

4 TRACKING AND VERTEX DETECTORS

The tracking systems have always been a major element of the physics reach of collider experiments. This is going to be true at the LHC as well. Furthermore, the ability to tag b-jets, usually referred to as 'vertexing', is now recognized as an essential tool to reach 'new physics', in particular decay modes of the top quark and, possibly, 2-jet decay modes of Higgs bosons.

After a short overview, this section is divided into three main subsections dealing with semiconductor devices, gaseous devices, and finally a summary of performances expected at the LHC.

4.1 Examples, constraints, and basic elements of a tracker

The role of a tracker is to measure the parameters of charged tracks produced in a given event, with a minimum of perturbation to them. The quantities of interest are the track momentum (with sign), the two angles defining the initial direction, and the vertex position if displaced from the primary collision point. All these quantities are affected by interactions with matter (multiple Coulomb scattering, bremsstrahlung, nuclear interactions, etc.) and it is a constant worry in designing a tracker to avoid having too much material in it. The material in the tracker, and the coil surrounding it, is also a serious limitation to the performance of the electromagnetic calorimeter (see Section 3).

Measuring the momentum requires a magnetic field, usually a solenoidal field in collider experiments (however UA1 had a dipole field). In some cases fieldless experiments have been performed (UA2, D0). This presents some advantages in terms of electron and photon identification, but is now considered insufficient in terms of the simple requirements summarized above [58].

Pattern recognition, in the dense environment of high-energy collisions, is perhaps the leading consideration in designing a tracking system. The historical example of the bubble chamber, in which one could visually see the tracks thanks to the high density of bubbles, remains in this respect a reference. The closest modern concept, adapted to interaction rates of a few kHz (as opposed to a few Hz for the bubble chamber) is the Time Projection Chamber (TPC). This detector was pioneered in Berkeley [59], and used at SLAC (PEP4 experiment) and LEP (ALEPH and DELPHI). It is also being used for heavy-ion collisions (NA49) demonstrating its superb granularity, and also planned for the ALICE heavy-ion experiment at the LHC [48].

4.1.1 The Time Projection Chamber

In the TPC the electrons and ions produced by charged tracks in the gas volume surrounding the interaction point drift under a uniform electric field towards the end plates (Figure 48). Electrons create measurable signals by avalanche amplification close to anode wires, like in any wire chamber. Cathode pads allow accurate measurement in the bending plane, while the time measurement gives the third coordinate (parallel to the beam axis). Typical accuracies are 500 μm in the radial direction, 100 μm in the $r\phi$ (bending) direction, and 500 μm in z. The high number of measurements along the track path, typically of the order of 100, allows some particle identification by $\mathrm{d}E/\mathrm{d}x$. An example of performance in this domain is shown in Figure 49. The capability reflects the variation with energy of $\mathrm{d}E/\mathrm{d}x$ by collisions, and is best adapted to low-momentum particles (below 5 GeV). The maximum practical operating rate of such a detector is

given by the drift time to the end plates, about 40 μs, i.e. about 20 kHz. This was perfectly matched to LEP (22.5 μs between bunch crossings), and is also adequate for LHC operation with ions (the rate of inelastic lead–lead collisions is about 10 kHz at the nominal luminosity of $10^{27}\mathrm{cm}^{-2}\mathrm{s}^{-1}$). An average density of 0.1 particle per cm^{-2} is considered a (safe) maximum.

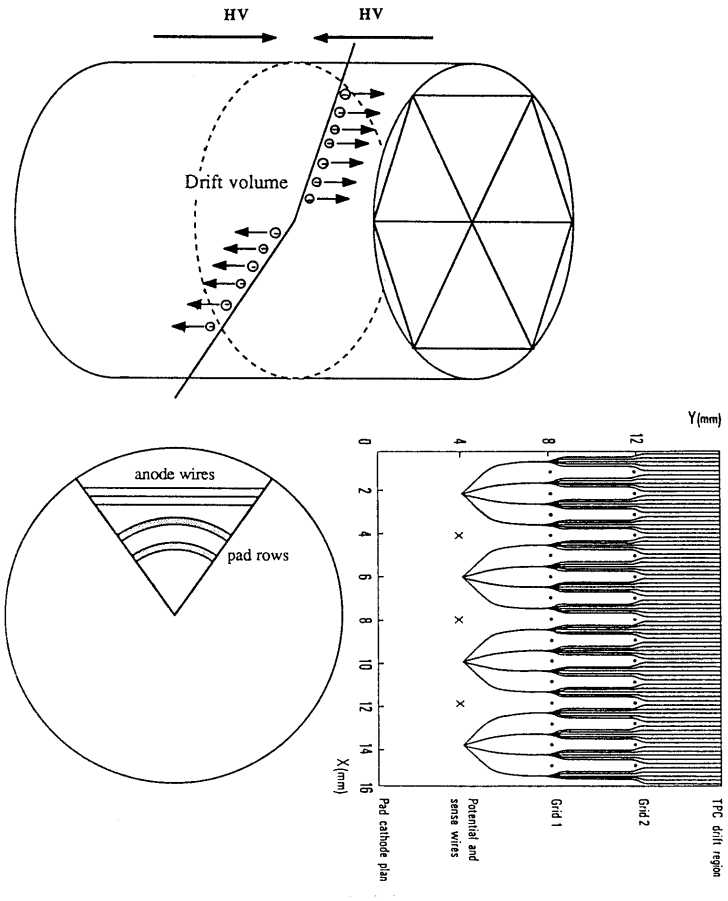

Figure 48: *Sketch of a Time Projection Chamber.*

4.1.2 Detectors and geometries at the LHC

In order to match the bunch-crossing rate at the LHC, one should consider detectors in which the signal duration is less than or about 25 ns. This leaves semiconductor detectors (silicium, GaAs, or possibly diamond), gas detectors with short drift path (MSGCs and small straw tubes), and scintillating fibres. The latter were discarded by ATLAS and CMS for various practical reasons. However, they form the basis of the D0 upgrade [58].

Given the small thickness of those detectors, the geometry of an LHC tracker has an 'onion shell' structure of cylinders (barrel part) and disks (forward part). This is illustrated by Figure 50 which shows a perspective view of the CMS tracker.

Most detectors (Si strips, straws, MSGCs) give as result a line on a surface, while pixels offer a 3-D point. Clearly this is much richer information, motivating much R&D activity to make these detectors suitable for large-scale application.

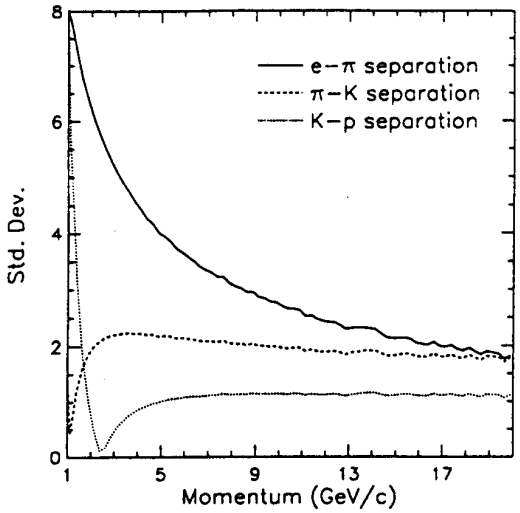

Figure 49: *$e/\pi/K/p$ identification capability of the Aleph TPC [60].*

In order to have a quantitative appreciation of the pattern recognition problem, one may evaluate the detector occupancy (probability of a cell being hit for a given triggered event) as a function of radius. At the nominal high luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ typical numbers are as follows:

- straws at radius 60 cm: occupancy $\simeq 20\%$ for cell size $s \simeq 1$ m $\times$ 4 mm;
- Si strips at radius 25 cm: occupancy $\simeq 1\%$ for cell size $s \simeq 0.1$ m $\times$ 100 μm;
- pixels at radius 12 cm: occupancy $\simeq 0.01\%$ for cell size $s \simeq 50$ μm $\times$ 300 μm.

They show that the increased granularity beats off the $1/r^2$ behaviour, provided the adequate technique is used.

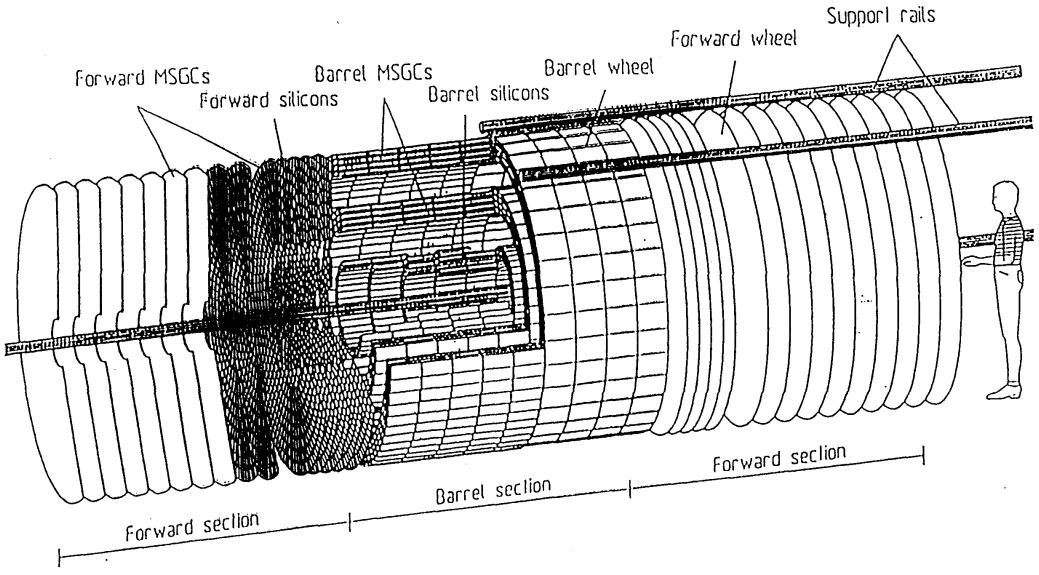

Figure 50: *View of the CMS tracker.*

4.1.3 Function and limitations

The minimum function of the tracker, at the nominal high luminosity, was defined in ATLAS so as to:

- recognize and reconstruct high p_T tracks in roads defined by the muon spectrometer or by the calorimeter;
- check for their impact parameter;
- check for the existence of a low-energy companion (for converted γ's or Dalitz pairs) and when possible
  - search for secondary vertices within a jet (jet-tagging);
  - perform full tracking within a jet;
  - perform full tracking of the triggered event (at luminosities $\leq 10^{33}\text{cm}^{-2}\text{s}^{-1}$).

Finally the tracker should participate in the electron and muon trigger at level II, by allowing fast track-finding within roads. As shown in the following paragraphs, technologies now exist or are in the process of being developed which should allow such functions with good efficiency. However, practical limitations related to material accumulated in the tracker from the different layers, power to be extracted by cooling, and cost, will inevitably limit the achievable performances. As an illustration of the first point the amount of material corresponding to the CMS tracker is shown, as a function of pseudorapidity in Figure 51 (the plot is qualitatively similar for ATLAS). Clearly, a material amount of up to 0.5 X_0 in the path of an electron, in a large magnetic field, is going to seriously limit the calorimeter intrinsic performance.

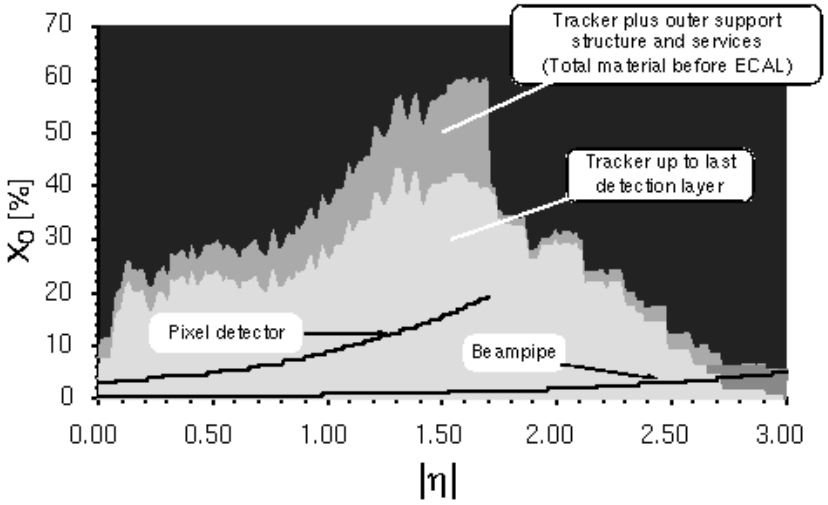

Figure 51: *Material in the CMS tracker.*

4.2 Semiconductor devices

Although semiconductor detectors had been used for quite some time in nuclear spectroscopy (see Section 3), their use for precision tracking is more recent, having been linked first to the measurement of short-lived 'charmed particles' [61]. Since then, this technique has seen an extremely fast development, and it is now considered a rather standard tool of high-energy physics experiments. The use as vertex detectors in Collider

experiments has, however, revealed fundamental problems associated with radiation damage (see below), which are better understood bearing in mind the basics of this detector.

4.2.1 Signal formation in a depleted junction

As is well known, energy levels of atoms become energy bands in crystals. At very low temperature the upper (conduction) band of an intrinsic semiconductor (Si, Ge) would be empty, and the lower one (valence) would be full. The gap distance between the two bands, in Si, is 1.12 eV. As temperature rises, some electrons move to the conduction band, leaving an equal number of holes in the valence band. In the presence of impurities there are electrons in the conduction band (n-type) or holes in the valence band (p-type) even at low temperature. In all cases the product of e and h densities is a function only of temperature [62].

The p–n junction is the most important structure as far as detectors are concerned. It consists of layers of opposite doping joined together. In the absence of external polarization, some electrons drift from the n to the p side (and vice versa for holes) until the 'built-in' potential $V_{\rm bi}$ stops the flow. At this steady point the Fermi levels are equalized, and there exists a thin 'depleted region' (no free charge) on either side of the transition. Polarizing the diode with the n side at the negative potential allows the evacuation of the charges. This is the forward biasing, which presents no interest as a detector. On the contrary, when biasing the n side at the positive voltage, one increases the depth of the depleted region. Eventually, for a large enough voltage, the junction is fully depleted. This is the situation of interest for a detector. Electron-hole pairs created by ionization (on average 3.6 eV in Si to create a pair) move under the electric field in the junction, and create the signal of interest. A few quantitative statements can be made:

– The depth of the depletion region is easily calculated using 'Gauss's theorem' (see Figure 52).
If ρ is the charge density after depletion (reflecting the doping level), one has:

$$d \mid \overrightarrow{E} \mid /\mathrm{d}x = \rho/\epsilon_0\epsilon_r$$

which gives a full depletion voltage (for a bulk thickness d):

$$V_{\rm d} = \rho d^2/2\epsilon_0\epsilon_r$$

from the Ohm law applied to the forward biased mode, one relates the charge density ρ, the resistivity r and the mobility μ by $\rho = 1/r\mu$, leading to:

$$V_{\rm d} = (1/r\mu)(d^2/2\epsilon_0\epsilon_r)\ .$$

For currently used dopings on a typical n-side of 300 μm, the full depletion voltage is about 100 V.

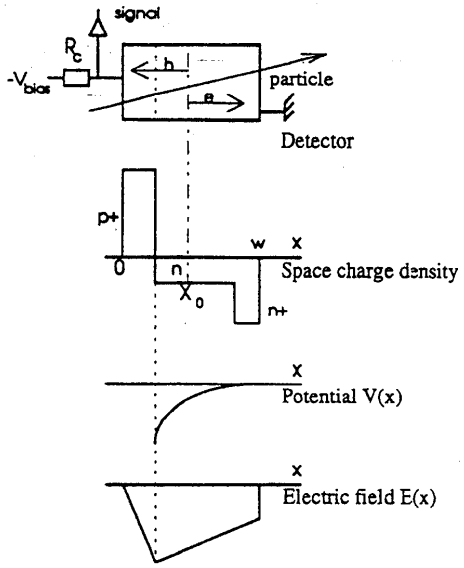

Figure 52: *An n–p junction*

– An electron-hole pair created at a distance x from the junction side, in a detector of thickness d gives rise to:
  – an electron current $i_- = ev_e/d$ for a duration x/v_e
  – a hole current $i_+ = ev_h/d$ for a duration$(d-x)/v_h$.
  This is well illustrated by Figure 53, taken from Ref. [63]. Signals from α sources are produced right of a test detector at the surface and give rise to an electron (hole) signal if produced on the junction (back) side. Because of the lower drift speed (typical mobility of 500 instead of 1500 $cm^2/V.s$) the hole signal is significantly slower, but still fast enough in terms of the requirements mentioned earlier. For a mip (signal from a minimum-ionizing particle), both carriers contribute. For the standard 300 μm thickness, the average signal collected is about 20 000 e^-.

4.2.2 Some practical aspects and limitations

Signal, noise, charge division The signal level being given, it has to be compared to the achievable noise. Like for the liquid argon calorimeter, the noise depends primarily on the capacitance to be read out, and on the preamplifier characteristics. For a simple geometry of adjacent strips on the junction side, with a typical strip size of 10 cm length by 100 μm pitch, the effective capacitance (predominantly to neighbouring strips) is about 20 pF. Typical preamplifiers adapted to a high level of integration (bipolar or CMOS processes) have a transconductance of typically 10 mA/V (for a few mW power dissipation) leading to a noise charge of about 1500 e^- for a fast bandwidth (response to a delta current peaking after about 15 ns).

Direct readout The simplest layout consists in directly coupling to a preamplifier metal strips laid on the thin (p^+ side) of the detector. In this way the resolution in position can hardly be better than the strip pitch divided by $\sqrt{12}$. A practical limit is about 50 μm pitch, or 15 μm resolution. In this scheme any leakage current (increasing with the

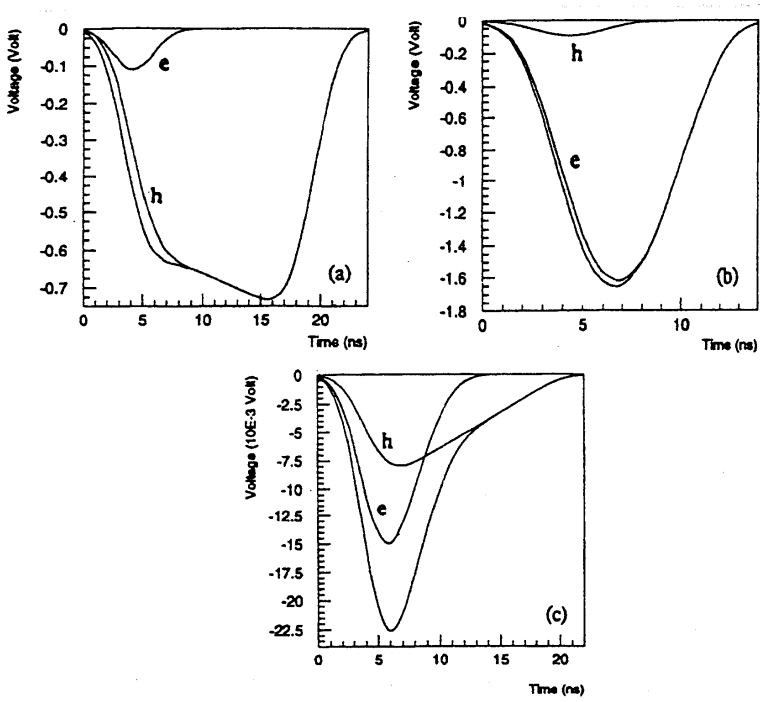

Figure 53: *Simulated alpha signals on the back side (a), the junction side (b) and mip signal (c).*

irradiation level) has to be evacuated through the preamplifier [64]. The vertex detector of the CDF experiment at the Fermilab Tevatron was built according to this concept, and operated successfully [65]. In particular the impact parameter resolution in the transverse plane was measured to be $[13 + 40/p_{\mathrm{t}}(\mathrm{GeV})]$ μm.

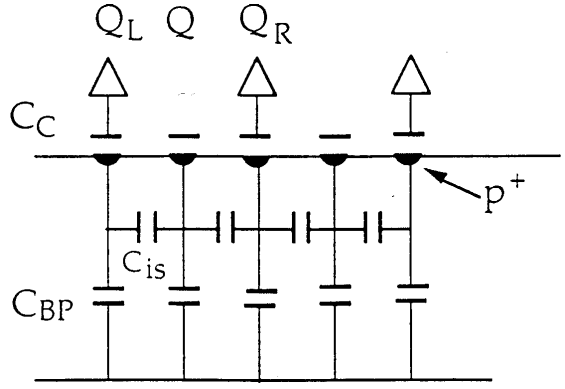

Figure 54: *Equivalent network in a configuration of charge division.*

Charge division An alternative scheme, somewhat more complex, consists in coupling the metal strips to the detector through a thin, insulating SiO_2 layer (AC coupling). The HV polarization is provided through a biasing (high) resistance network, which eliminates also the leakage current. This layout offers the interesting possibility of 'charge division'.

It consists in equipping only one strip out of say three or four with a preamplifier. In this case the detector can be approximated by the network of Figure 54. It is a simple matter to see that the signal from drifting charges divides itself between the two closest strips equipped with preamplifiers, in a way which gives signal proportional to the distance ratio between the hit strip and the readout ones. Measuring the charge on these two strips thus allows one to interpolate, and to obtain an accuracy better than that given by the pitch of strips equipped with preamplifiers.

One of the limiting factors in this way of proceeding is the noise, which has to be significantly smaller than for the previous scheme. A noise-to-mip ratio of 12 is considered a minimum. As an illustration one can quote the results obtained with the first DELPHI vertex detector [66], which are displayed in Figure 55. A parametrization of the resolution in impact parameter is $[24 \oplus 69/p_t(\mathrm{GeV})]$ μm.

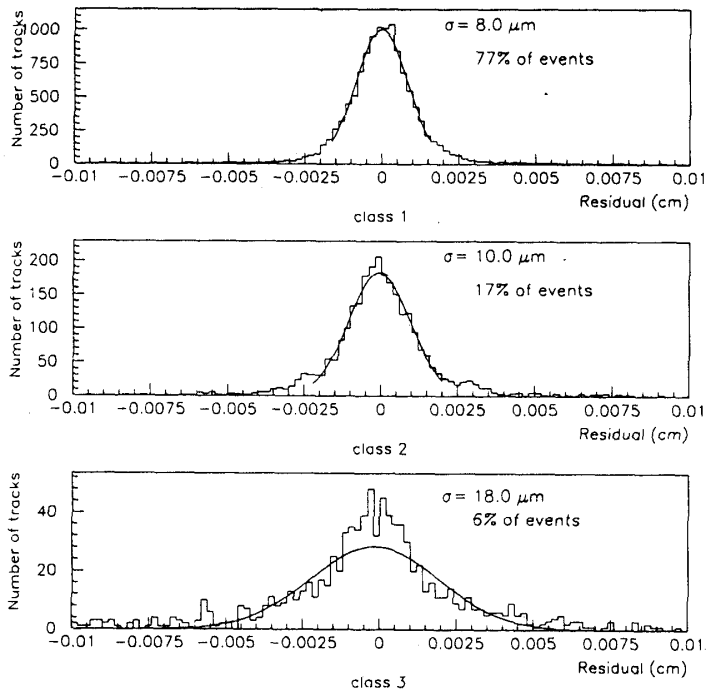

Figure 55: *Distribution of the residual per plane with the DELPHI vertex detector using charge division. The three classes of events reflect the hit multiplicity along a given track [66].*

Double-sided detectors By putting n^+ implants and metal strips on the back side of a n-bulk detector, one can also read out the back side, frequently referred to as the ohmic side, as opposed to the junction side. In this way a single detector offers a two-directional readout. A simplified drawing of this scheme is shown in Figure 56, taken from Ref. [67]. As an illustration of this scheme, Figure 57 shows the pulse-height correlation observed between the two sides (since the two signals are built out of the same drifting charge, they should be highly correlated). This scheme offers the obvious benefit of saving one

detector thickness out of two (i.e. 300 μm of Si). Depending on the material associated with other elements (electronics, support, cooling.) this advantage may overwhelm, or not, the increase in difficulty, and cost, associated with double-sided wafer processing.

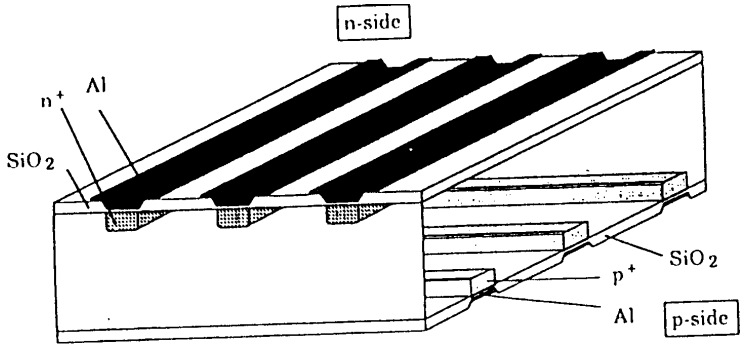

Figure 56: *Simplified drawing of a double-sided semiconductor detector.*

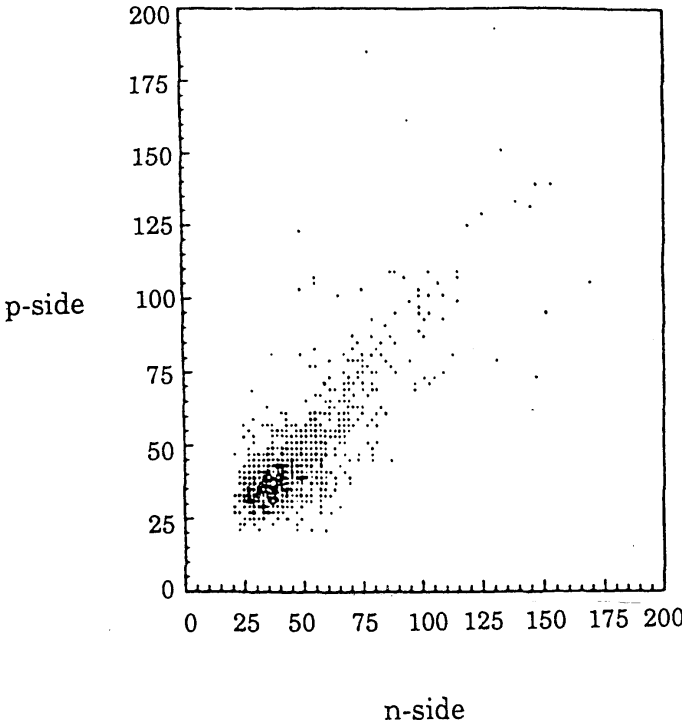

Figure 57: *Charge correlation between the two sides of a double-sided detector.*

As an illustration Figure 58 shows the performance obtained in this way with the DELPHI vertex detector [68]. The inferior accuracy obtained in the z direction is mainly associated with the inclination of tracks in a cylinder geometry.

4.2.3 Radiation damage. Type inversion

A major issue in the use of silicon detectors at the LHC is their behaviour after irradiation. The most important effect is associated with non-ionizing energy loss (NIEL) [69], which is mainly due to hadrons (see Figure 59). Nuclear interactions create defects in crystals, which alter their behaviour. The fundamental effects that occur are not totally

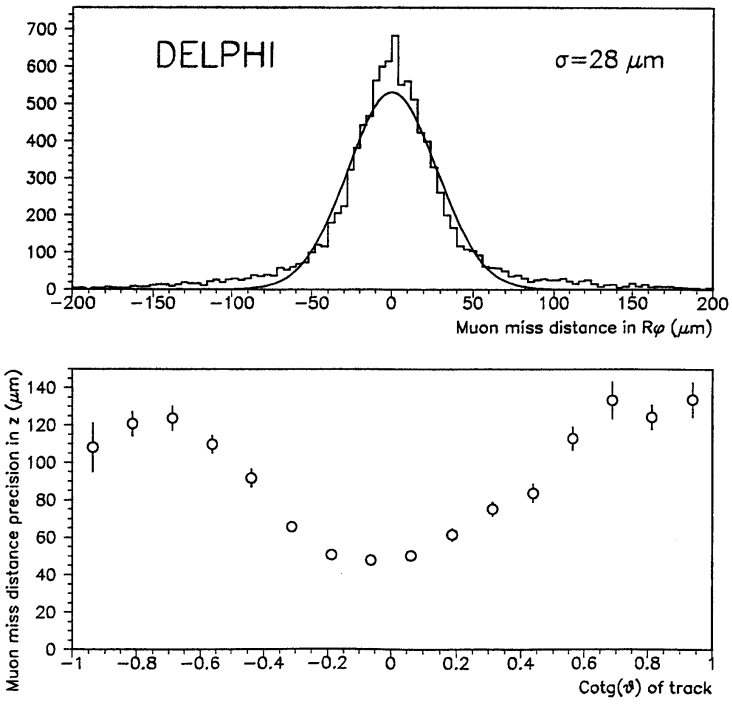

Figure 58: *Extrapolated distance to the vertex for muons with the two-sided vertex detector in DELPHI.*

clear yet, and a sound solid-state-physics-based model is still lacking. However, the observed behaviour is reproducible, and rather well described by phenomenological models (see for example Refs. [70] and [71]).

Usually one starts from a junction in which the bulk is n-type. One observes that, under irradiation, the n-doping equivalent concentration decreases, reaches zero for a dose between 1 and 3×10^{13} n/cm^2 (1 MeV neutron equivalent), and then the substrate behaves as p-type, with an equivalent p-doping which increases with the dose (see for example Figure 60 taken from Ref. [70]). This is called type inversion. A direct consequence is that the junction side moves from where it was originally to the other face of the detector (originally the ohmic side).

If one were to start from a p-bulk, its doping concentration would simply increase (see Figure 60). So the observation is consistent with donor removal and acceptor introduction. Mechanisms invoked for donor removal are, for example, combination with a lattice vacancy. Alternative explanations exist, like the creation of deep level defects (more or less at the middle of the forbidden band), which can ionize either as p or n types [72].

The above has extremely important consequences for LHC detectors. Following the apparent behaviour of impurity concentration, the voltage for full depletion would first decrease, and then, after inversion, increase continuously. Since one can not go further than a certain practical limit (250 V or so), eventually, the detector will no longer be fully depleted. This is illustrated schematically in Figure 61. In this case the readout of strips on the initial junction side essentially no longer works because of the undepleted layer which shorts all strips together. On the contrary, the readout on the initially ohmic side

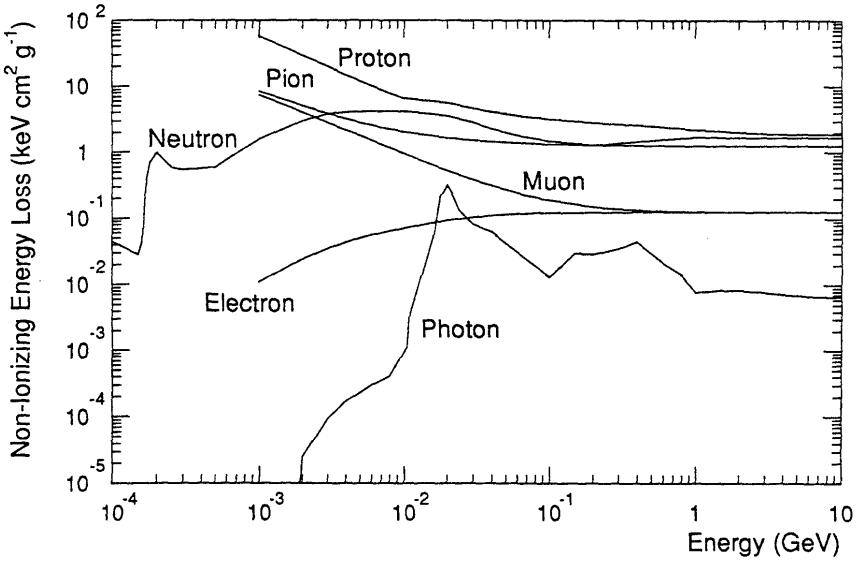

Figure 59: *Non-ionizing energy loss of different particles as a function of their kinetic energy [69].*

still works properly thanks to the n^+ implants, but with a reduced signal because of the incomplete depletion.

Such a scheme, with strips on the initially ohmic side, and referred to as 'n^+n', is being seriously considered for the (single-sided) readout of the LHC detectors, although the choice has not yet been made.

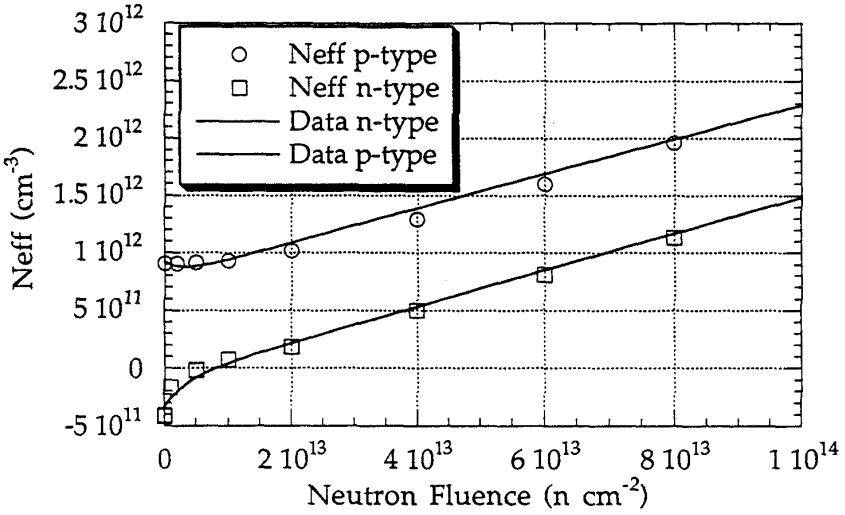

Figure 60: *Effective doping concentration of p-type and n-type silicon detectors after an increasing irradiation (1 MeV neutron equivalent).*

Another important aspect of radiation damage is the increase of the reverse current (or leakage current) associated with the defects created. This current I_{L} contributes to the power dissipation and, most importantly, generates a shot noise ($i_{\mathrm{n}} = \sqrt{2qI_{\mathrm{L}}B}$ in a

bandwidth B) which amounts typically to 500 electrons r.m.s for a current of 1 μA.

Finally a very important aspect is the temperature of the detector, both during operation and maintenance periods. Given a temperature of operation, say 5°C, one observes for a short time, when warming up, some decrease of the irradiation effects (annealing), and then a definite worsening of them (anti-annealing). Furthermore the leakage current is a fast varying function of the temperature. In total, the present trend is to design trackers for cold operation (about 0 to 5°C), with the requirement that they will not be warmed up to 20°C more than one month or so per year. An illustration of the consequences for the planned ATLAS tracker is shown in Figure 62.

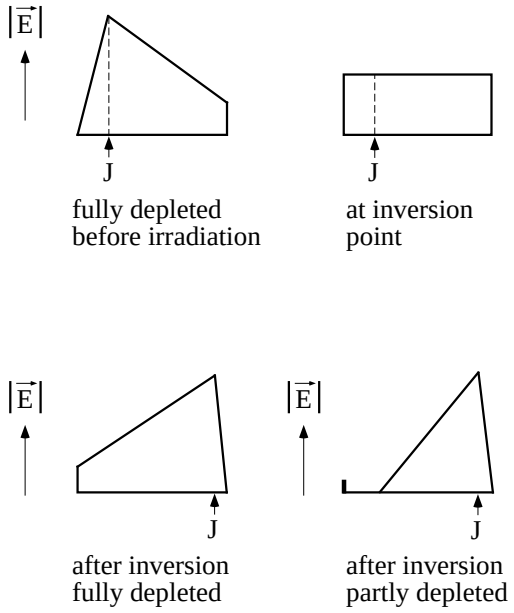

Figure 61: *Schematic type inversion.*

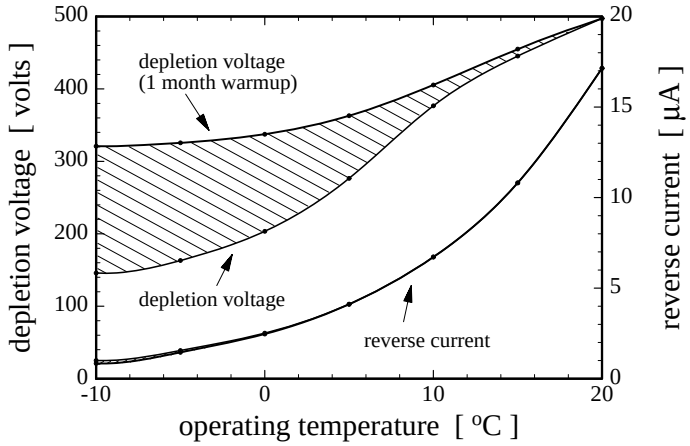

Figure 62: *Full depletion voltage and leakage current in the ATLAS silicon layer at* $R = 30$ *cm, as a function of operating temperature, for a 10-year exposure.*

4.2.4 Pixel layers

As recalled at the beginning of this section, there are clear motivations for high-granularity and high-precision detectors close to the vertex. Pixel detectors represent a further challenge beyond strip ones, mainly because a new way had to be mastered for connecting detector pads and preamplifiers in two dimensions.

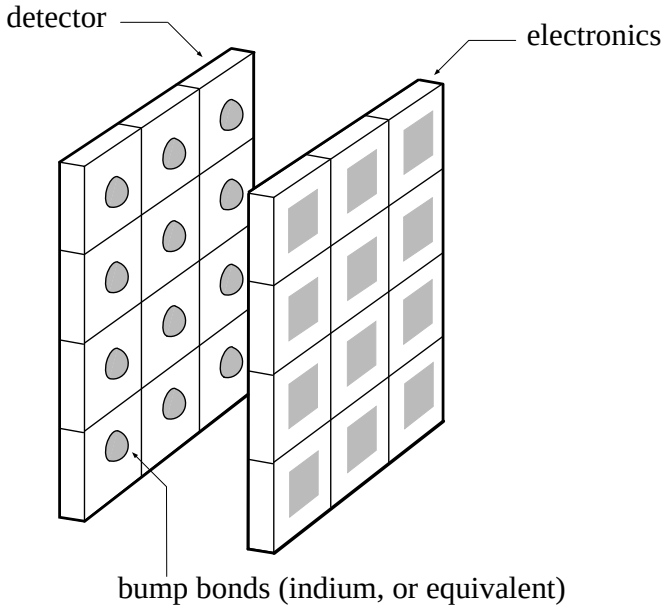

Figure 63: *Sketch of the pixel detector and electronics assembly by bump-bonding.*

As schematically shown in Figure 63, the current technique is to join together by 'bump-bonding' two semiconductor wafer elements, one carrying the detector (n-bulk-p junction), and the other the readout electronics (preamplifier and subsequent steps). For typical pixel sizes of 50 μm $\times$ 300 μm the bonding operation represents 6000 connections per cm^2, which is at the limit of what the microelectronics industry can do at present. Integrating the functionality of the readout on such a size (0.01 mm^2) has already been done. The pixel system of the Omega spectrometer (experiment WA97) represents the largest system used so far in an experiment [73], with a total of 72 kpixels.

Besides bump-bonding, the next difficulty with pixels is to organize the readout in such a way that the potentially enormous amount of data produced by such a detector does not flood the readout. This is possible because of the low occupancy (thanks to the small size), and the excellent signal-to-noise ratio: because of the small cell size, the noise can be kept below 500 electrons (for LHC adapted shaping) for a signal which is still 20 000 electrons if a thickness of 300 μm is used. So, using a threshold of say 5000 electrons one has virtually no noise hits, and more than 99% efficiency.

A practical scheme consists in organizing the readout in a 'column architecture', in which all information is transmitted along a column of pixels (typically 64 or 128 elements) and none along the rows. One scheme proposed for ATLAS consists (schematically) in loading, when the pixel is hit, the pixel address in the column, in an 8-bit column shift register which is clocked down the end of the column at each bunch crossing. When the address appears at the bottom of the column, a bunch-crossing (bc) counter started by

the 'or' of the column, tells when this pixel was actually hit, thus giving the full address: bc number, column number, and address in the column. This information is kept until a match can be made with the level-I trigger bc number. If a match occurs, the hit is kept and transferred off-chip using optical fibres.

Both ATLAS and CMS plan two layers of pixels over the full tracker acceptance, at typically 12 and 16 cm radius in the barrel part, amounting to about 10^8 channels. At such a small distance from the interaction vertex, the irradiation problems are frightening. However, the depletion voltage problem can be eased, given the good S/N ratio, using a thinner detector (250 μm), and the leakage current noise is not critical because of the small detector cell size. These are clear advantages of the pixel layout. What remains, however, is the electronics radiation hardness, which is the subject of vigorous R&D in collaboration with industry (using for example the processes SOI3HD or DMILL, see Ref. [74]). At the time of the present report, there are good indications, but no firm demonstration, that a solution exists.

Finally it should be said that, as far as radiation hardness is concerned, there is a good propect of Schottky (metal-semiconductor) GaAs detectors (see for example Ref. [75]) resisting up to irradiation levels of 10^{15} n/cm^2 (1 MeV equivalent neutrons).

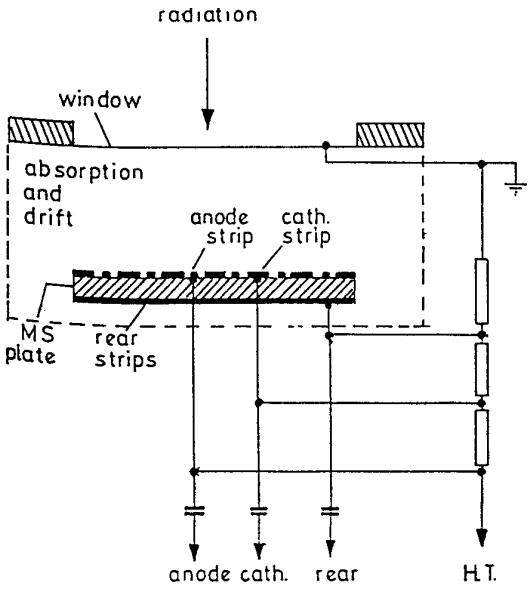

Figure 64: *Schematic drawing of an MSGC and of its polarization.*

4.3 Gas devices

4.3.1 Microstrip gas chambers

MSGCs are an attempt to reach—almost—the granularity of silicon strip detectors, using the cheaper approach of gas detectors. The key point is to replace wires (see for example Ref. [76]), which provide amplification in a MWPC, by strips deposited on a glass substrate, (see Figure 64). The gas gap is typically 3 mm, providing about 30 electron–ion pairs for a mip. A very thin and narrow anode strip allows amplification (see a typical field map on Figure 65 taken from Ref. [77]) by a factor of a few thousands, giving signals comparable in magnitude to those from Si although somewhat longer due to the drift time

in the gas (50 ns). Very specific gas mixtures, Ar-DME or CO_2-DME, have to be used in which the dimethyl-ether avoids secondary emission and prevents ageing.

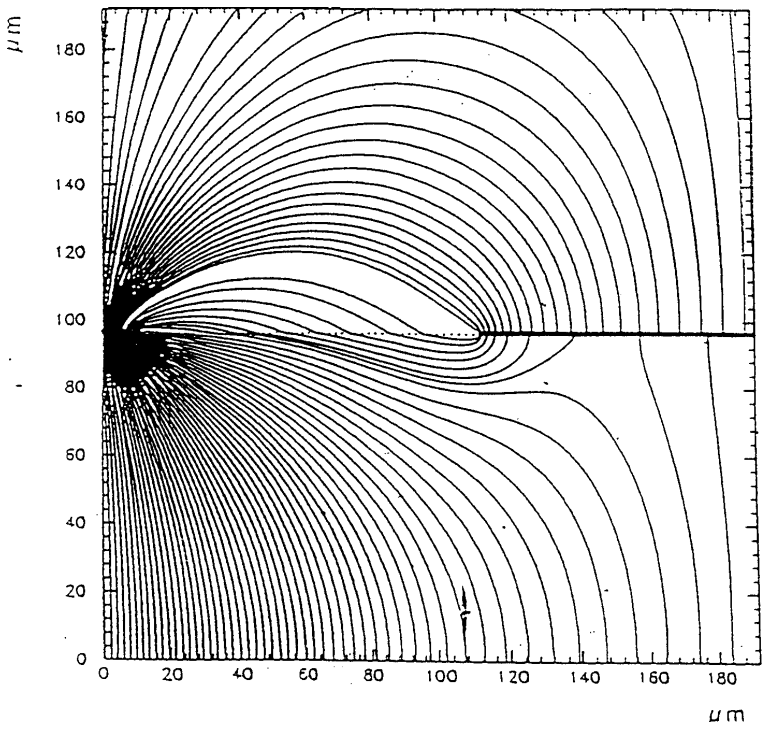

Figure 65: *Calculated field map in an MSGC. The density of field lines in the bottom part (substrate) depends on the backplane voltage. The anode is at the point where the field lines converge.*

Associated with this modest amplification, it was found that an instantaneous counting rate of 10^6 particles/mm^2 could be tolerated, which is adequate at the LHC. As the active material is primarily the gas, the detector was thought to be intrinsically radiation resistant. Wire ageing, a phenomenon known to happen in classical wire chambers (see the muon section), was however found to limit the integrated flux of operation of MSGCs to at most 0.1 C/cm (in extremely clean conditions, with gold electrodes), corresponding to about the LHC flux integrated over 10 years, for a detector at 30 cm from the beam.

The most difficult aspect of these detectors was found to be associated with the glass (or other material) substrate. Unlike wire chambers were the ions drift in the gas away from the anode, up to the cathode (metallic or rather good conductor), in a MSGC, part of the ions are led to drift in the substrate, close to its surface (the density of field lines in the substrate depends upon the backplane voltage). In these circumstances, the substrate charges up, which modifies the potential distribution, and therefore the chamber gain.

It has been found that stable operating conditions (i.e. independent of the instantaneous rate) can be obtained with a bulk conducting glass of resistivity about 10^{10} Ωcm, or with surface conducting substrates of resistivity around 10^{12} $\Omega/\Box$. The best known example of the first type is S8900, a glass containing heavy metals ($X_0 \simeq 4$ cm), and conducting by electron transfer. A well-known example of the second type is DESAG

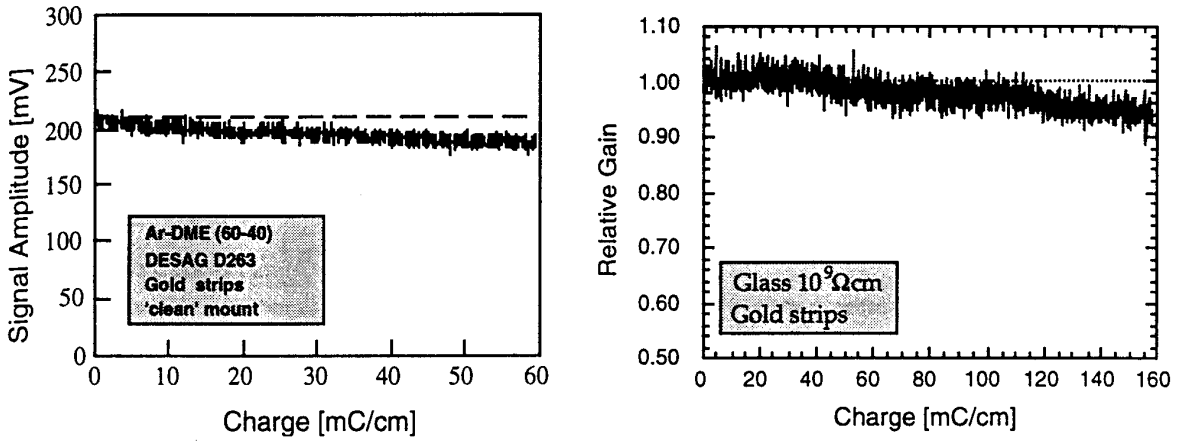

Figure 66: *Ageing properties of MSGC prototypes with surface conducting (left) or bulk conducting (right) substrates.*

D263. With both of them, and using extremely clean gas distributions systems, and gold electrodes, good results were obtained as shown in Figure 66. The lighter D263 glass is preferred for reasons of transparency, although its behaviour is somewhat inferior to the other one. A solution could be D263 with a thin coating of S8900 (or diamond).

Many prototypes of MSGCs have been tested in the last few years (see for example Ref. [78]). For a typical anode pitch of 200 μm, the resolution is currently 40 μm for tracks perpendicular to the chamber. In case the track path in the chamber projects onto many strips (up to 10 at 45°), the resolution and the efficiency degrade rapidly. In the CMS barrel part, the layout includes a tilt angle of about 12° to compensate for the Lorentz angle (see Section 2.5).

The MSGC technique looks promising, but is still evolving. Such detectors have been retained by CMS to equip the tracking volume at radii larger than 30 cm.

4.3.2 Straw chambers

To equip the outer part of its tracker, ATLAS has chosen straw tubes. These detectors are less granular than MSGCs (a typical straw is 50 cm long and 4 mm in diameter), but being much more transparent, a much larger number of layers can be afforded for the same amount of material, opening the possibility of so-called 'continuous tracking'. Straw tubes have been used for a long time to equip vertex detectors, in place, or right behind semiconductor layers. In their principle they are no different from the circular tubes discussed in Section 2.5. However, their size, location, and use has led to a rather different class of problems.

An interesting aspect of the ATLAS straws is their use as Transition Radiation Detectors. When an ultrarelativistic particle crosses many boundaries between media with a different index of refraction (stack of polypropylene foils) it emits γ quanta (of a few keV) which convert in the high Z gas (xenon) used in the straws. The practical use of this property is to help identify electrons (which emit γ's) out of a hadron background (see for example Ref. [79]) which do not.

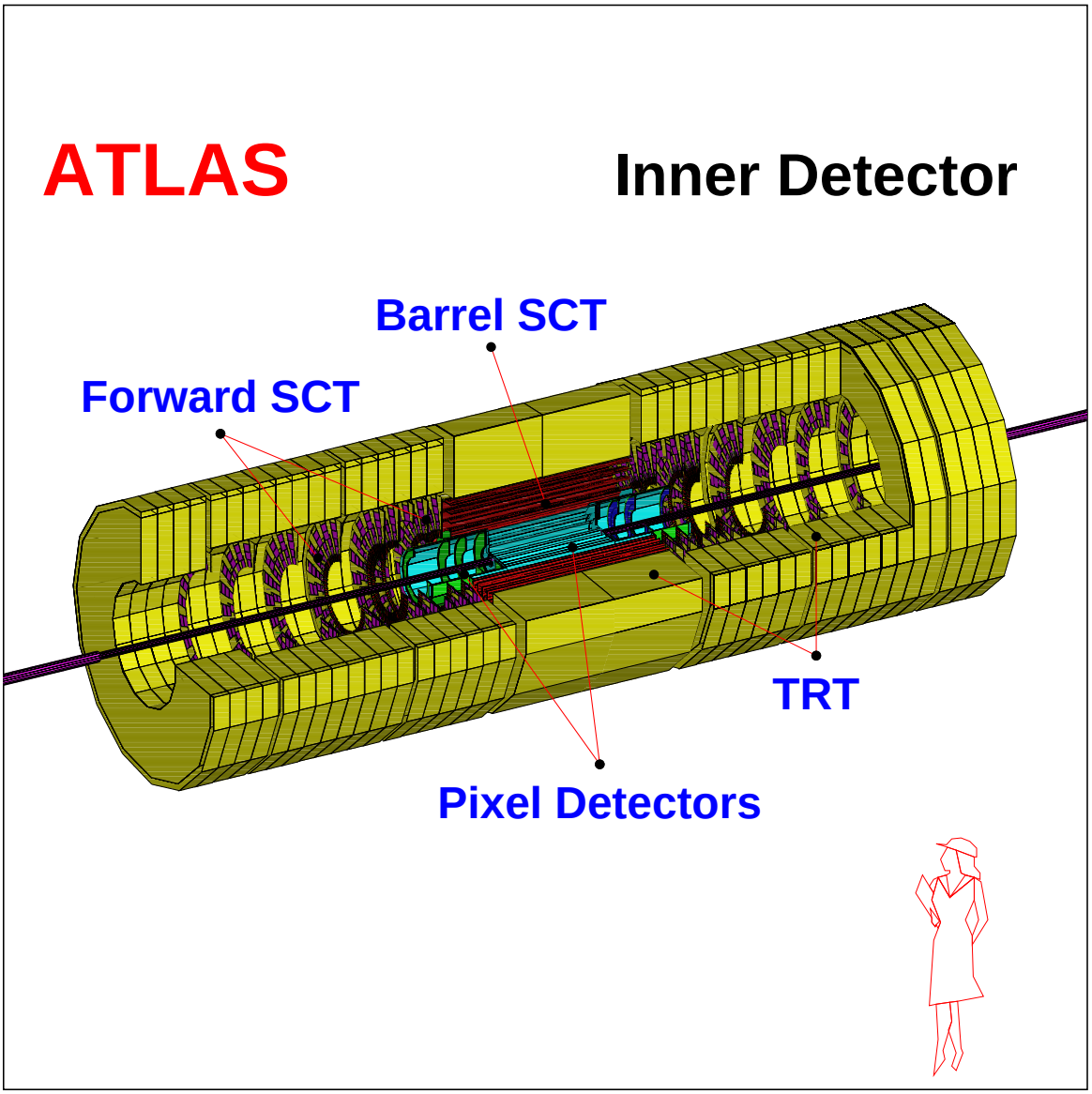

Figure 67: *Perspective view of the ATLAS tracker showing its major elements.*

The geometry of the ATLAS tracker is shown in Figure 67. The TRT occupies the large radius part of the inner detector cavity. Straws are arranged parallel to the beam in the barrel part, and orthogonal to it in the end-cap part, to optimize the use of the magnetic field for momentum measurement. The total thickness of the detector (gas, wires, straws and radiator) is about 0.15 X_0.

Besides the difficulties associated with the mechanical setting up of a (transparent and accurate) 0.4 million straw system, the main problems of the straws in ATLAS are related to the high rate leading to an occupancy in the 10–30% range. In particular, the signal shape, which has a long $1/t$ behaviour (reflecting the ion drift speed) should be electronically compensated by an appropriate network.

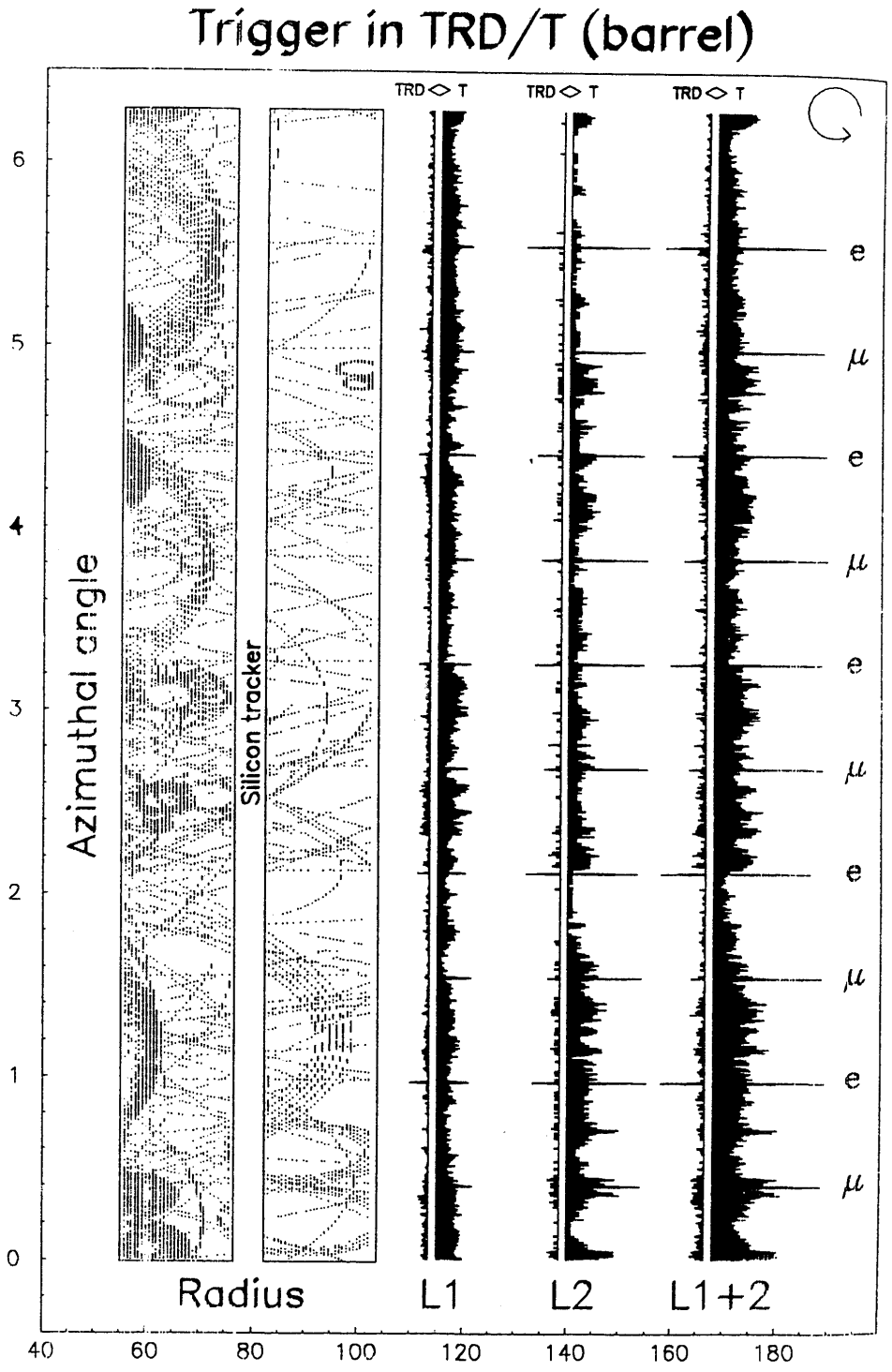

Figure 68: *Simulation of tracks in the ATLAS straw tracker. The histograms on the right indicate how a fast trigger could recognize high p_T tracks by applying a threshold to the number of hits in a given solid-angle bin.*

With a typical voltage of 1.8 kV applied to 50 μm wires, the gain is about 2×10^4, and the maximum drift time 40 ns (with a $Xe/CF_4/C0_2$ mixture). Ten years of LHC

running correspond to about 1 C/cm on the inner wires. Tests have shown that this is rather easily achievable (note the difference with respect to MSGCs). An electronic noise of typically 2000 electrons with the proper bandwidth allows the measurement of the drift time out of the first electron reaching the anode wire. A consequence of high occupancy is the association of wrong hits with the track. Detailed simulations indicate a resulting mean accuracy per straw of 160 μm, and an equivalent accuracy for the straw stack in the barrel part of 60 μm.

The main strength of the straw tracker is pattern recognition associated with the continuous tracking. This is illustrated in Figure 68 showing a simulated ATLAS event in which high p_{T} tracks are easily seen, both visually and with a simple histogramming method which can be used at the trigger level II.

4.3.3 Some technical aspects

For optimum operation and mechanical stability, the detector temperature has to be carefully regulated, despite the power dissipation of the tracker electronics, amounting to typically 50 kW for either experiment. One way is to use as cooling medium circulating in small pipes, a liquid containing small ice crystals in suspension (binary ice), much of the power being taken by the ice/water phase transition.

Another very difficult point is alignment. The accuracy aimed at in relative detector positioning is about 10 μm. As for the muon spectrometer, CCDs could be used, but they are not likely to be radiation-hard enough. Another elegant approach is to use 'frequency scan interferometry': by counting fringe numbers in a frequency scan between two interfering light paths from a laser and fibre network, one can measure distances of up to 1 m with an accuracy of better than 1 μm (see for example Ref. [80])

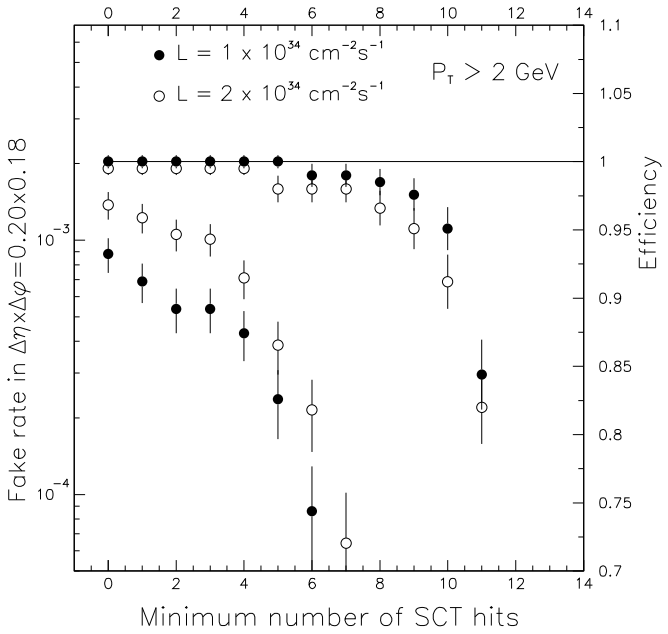

Figure 69: *Efficiency of track reconstruction (right-hand scale) and probability of fake track (left-hand scale) as a function of the number of hits in the silicon layers (ATLAS).*

4.4 Tracker performances

4.4.1 Pattern recognition

As mentioned above, this is one of the key points in order to have a tracker fulfil its task. Continuous tracking, as illustrated in Figure 68, gives a visual impression that high-p_T track finding is indeed possible. More thorough studies were made, using full simulation of 'mimimum-bias' events superimposed onto 'physics' events containing either leptons, or b-jets.

In ATLAS the performance for tracks above 2 GeV is illustrated by Figure 69 which shows the reconstruction efficiency of such tracks as a function of the number of spurious tracks found in the same window defined by hardware considerations (strip length). Requiring at least five hits in the semiconductor inner layers, the fake probability (a few 10^{-4}) is well below the pileup probability (0.7%) at design luminosity.

Another important aspect is the track-finding efficiency and the fake probability in a jet. In CMS, with a 97% track efficiency, the average fake probability is below 2% per track of the jet.

4.4.2 Momentum accuracy

The results obtained for muon tracks by a simulation of the CMS tracker are shown in Figure 70. The gain from the knowledge of the vertex position is clearly visible, especially at high momenta. The combined performance of the tracker and muon spectrometer was shown in Section 2. In ATLAS, the performance of the tracker is typically a factor 2 worse, reflecting the lower value of the magnetic field.

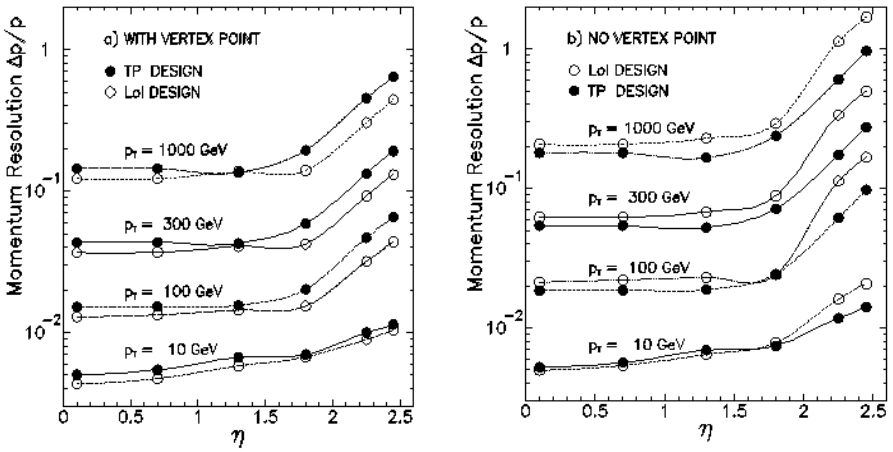

Figure 70: *Accuracy of muon momentum measurement from a simulation of the CMS tracker.*

As a figure of merit of the combined tracker and muon spectrometer of the two experiments, Figures 71 and 72 show the best expected mass resolution of a Higgs boson decaying into four muons. In case only the 'stand-alone' muon spectrometer could be used, the ATLAS performance would worsen by about 30%, and the CMS one by up to a factor 5 (see Figure 15), illustrating the difference of approach.

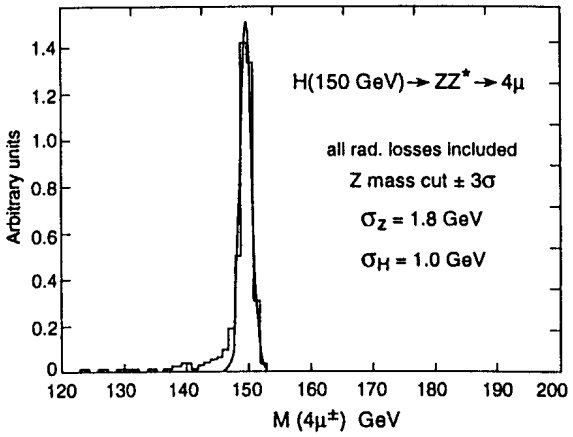

Figure 71: *Simulated 4μ reconstructed mass in CMS, for a 150 GeV Higgs (tracker and muon spectrometer combined).*

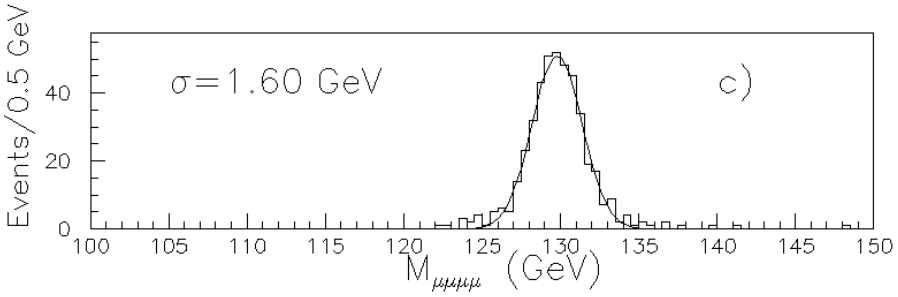

Figure 72: *Simulated 4μ reconstructed mass in ATLAS, for a 130 GeV Higgs (tracker and muon spectrometer combined).*

More difficult than the measurement of muons is the tracking of electrons which is largely spoilt by radiation in the tracker material. The performance achievable at high momenta is particularly important since the tracker is the only detector that can give the electron sign, notwithstanding the fact the accuracy in absolute value is dominated by the calorimeter, even down to 10 GeV or so.

In order to limit the degradation of momentum accuracy due to bremsstrahlung, one can incorporate in the track fitting the shower centre of gravity from the calorimeter. It is a simple matter to see that, in the small-angles approximation, this point is independent of radiation/conversions (see Ref. [81]). Even after this procedure the momentum resolution is 2.2% at 20 GeV in ATLAS while it was 1.2% at the same energy for muons.

In CMS, on account of the high magnetic field, the degradation is even more pronounced. This degradation affects not only the tracking result, but also the calorimeter measurement, as mentioned earlier. Despite the better crystal ideal resolution, the mass spectrum of a Higgs boson decaying into four electrons is, in the present state of the detector and its simulation, worse in CMS than in ATLAS (mass resolution of typically 1.8 GeV in CMS against 1.6 GeV in ATLAS for a Higgs mass of 130 GeV).

4.4.3 Vertexing

The vertexing capabilities are important for the B physics programme at the initial low luminosity of the LHC, and for tagging of b-jets to identify specific processes at all luminosities (top quark production, Higgs boson decaying to $b\bar{b}$, etc.) or to reject ($t\bar{t}$ and $Zb\bar{b}$) as backgrounds to a Higgs boson decaying into four leptons.

To assess the first point, the critical figure of merit is the accuracy in impact parameter measurement. This depends obviously on the detector layers, intrinsic performances, but also very critically on the radius at which the first layer is situated. In order to reach the best performance in the initial phase both experiments plan to install a 'B-layer' at 7 cm or so radius, which will be removed later on, leaving as first element a pixel layer at radius about 12 cm. Including this first layer, the planned performance for CMS is shown in Figure 73. It is to be compared to a mean impact parameter for tracks from a b-decay of typically 150 μm.

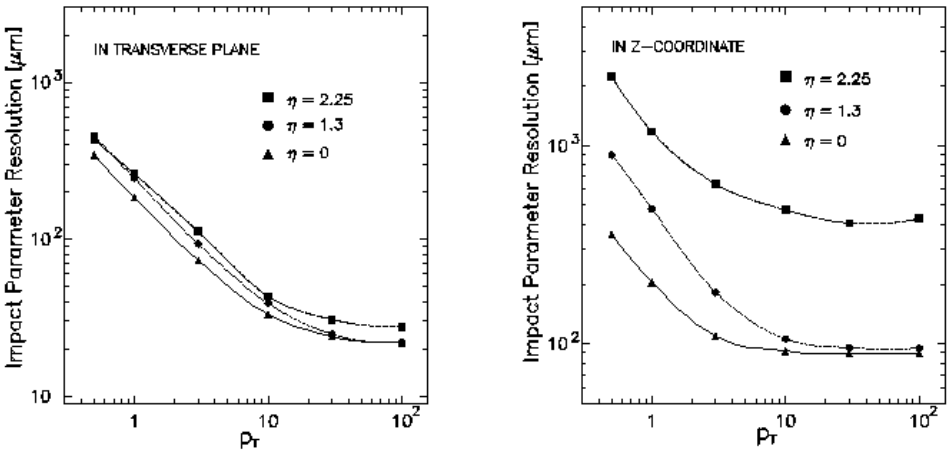

Figure 73: *Impact-parameter resolution calculated for the CMS vertex detector.*

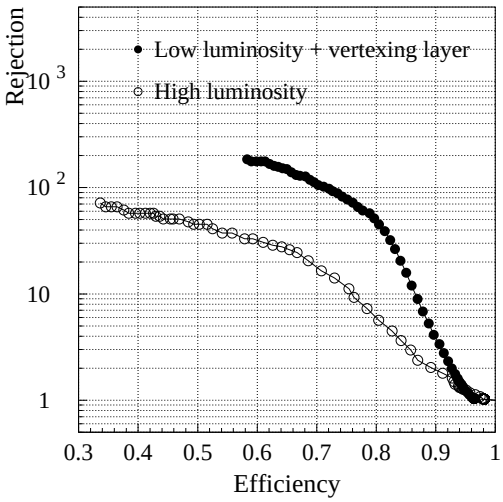

Figure 74: *Simulated non b-jet rejection as a function of b-jet tagging efficiency, as obtained by simulation of the ATLAS tracker.*

The vertexing capability at high luminosity depends not only on the intrinsic accuracy, but also on the pattern recognition capability in a dense environment (efficiency/fake rate, see above). As an illustrative performance, Figure 74 shows the simulated ATLAS capability at low and high luminosity. A rejection of non b-jets of a factor 20, with an efficiency of 60% to tag b-jets is considered extremely valuable.

5 SIGNAL PROCESSING, TRIGGER, AND DATA TRANSMISSION

These topics have become one order of magnitude more complex during recent years with the needs of new high-intensity experiments on CP violation in the kaon system and with the LHC experiments. The signal processing and trigger, which are strongly linked, are discussed. We also discuss briefly the data transmission.

5.1 Trigger

The trigger is a key part of any high-energy experiment with severe requirements such as:

- Very high efficiency in order not to miss rare signal events.
- Absence of bias in the selection.
- Capability to match the high rate coming from the front end with no or small dead time.
- Reduction of the data flow as early and as much as possible.

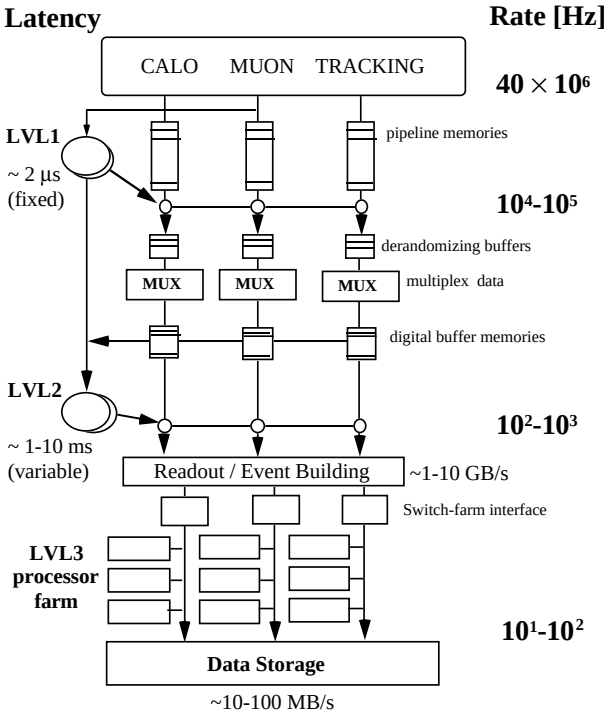

Figure 75: *ATLAS trigger architecture.*

For LHC experiments to take a decision in 25 ns and fan it out is obviously not possible and the trigger which consists of three successive levels and the data acquisition

have to be pipelined. A conceptual scheme of the ATLAS trigger is presented in Figure 75 with the foreseen latency and rates.

Table 5: *Level-I trigger requirement at high luminosity in ATLAS*

Trigger	Rate (kHz)
1 μ with $p_{\mathrm{T}} > 20$ GeV	4
1 e.m. cluster with $E_{\mathrm{T}} > 30$ GeV	20
2 μ with $p_{\mathrm{T}} > 6$ GeV	1
2 e.m. clusters with $E_{\mathrm{T}} > 20$ GeV	4
1 jet with $E_{\mathrm{T}} > 300$ GeV	3
Large E_{T} miss	1

At level I the information from different subdetectors cannot be combined. Only elementary operations are performed to define elementary trigger conditions (see Table 5) and the associated Region of Interest (RoI) around an e.m. cluster or a muon track in the spectrometer. In order to keep the information from events which will satisfy the first level the data are stored every 25 ns in pipelines (analog or digital). The data do not move in the pipeline and read/write pointers deal with empty cells and keep the address of level I accepted data. In Table 5 are summarized some of the level-I requirements of ATLAS/CMS with their rates. The expected rejection factor is about 1000.

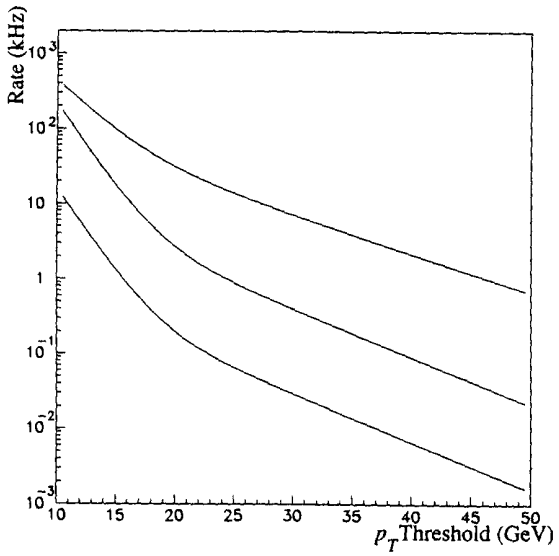

Figure 76: *Two-photon trigger rate as a function of E_T threshold. The upper curve is rate at level I without isolation, the middle curve after isolation and the bottom curve at level II with the preshower used.*

The level II trigger uses the RoI defined at level I to combine the subdetector information including inner-detector information. The full granularity of the calorimeter is used and the criteria refined: shower shape, track match with e.m. cluster, E/p, isolation, etc. The goal is to reduce the data flow by a factor 100 down to a level rate of about 1 kHz.

As an example the rate as a function of the transverse energy threshold is presented for the two-photon trigger in Figure 76. At level I the use of a crude isolation (because of large cell) allows one to lower the rates by a factor 10. Another factor 10 is gained at level II by using the fine granularity (strips in first compartment) and a better γ/jet rejection

The level III runs a full reconstruction algorithm and selects the most interesting events in the allowed bandwidth. As the average event size will be about 1.0 Mbyte, (see Table 6) and assuming a 100 Mbyte/s recording rate, the output of level III should not exceed 100 Hz.

Table 6: *CMS average event size at high luminosity*

Subdetector	No. channels	Occupancy (%)	Event size (kbyte)
Pixel	80 000 000	0.01	100
Inner tracker	16 000 000	3	700
Preshower	512 000	10	100
Calorimeters	250 000	10	50
Muons	1 000 000	0.1	10
Trigger	10 000	100	10
Total			970

5.2 Signal processing

The signal processing has to cope with the data acquisition requirements and is specific to each subdetector.

5.2.1 Inner tracker devices

An 8-bit accuracy is enough for the tracker information. Moreover in pixel detectors where the only information needed is the state of the pixel (signal or not), a simpler binary approach with a threshold is used. In detectors which need an analog readout for charge division, the solution is an analog pipeline (switch capacitor array) followed by an 8-bit Analog-to-Digital-Converter (ADC) at 100 kHz as the information is used only at level II.

5.2.2 Calorimeters

In an e.m. calorimeter cell, the noise is about 20 MeV and the maximum signal about 2 TeV, that is to say a 10^5 range. In order not to degrade the energy resolution, the noise has to be kept at the level of the least significant bit, so that a 17 bits range is desired. Handling such a dynamic range is very difficult. Various schemes have been considered:

- The Fermi Collaboration [82] proposes a dynamic range compression to 11 bits followed by a 40 MHz ADC and a digital pipeline. The board includes also many facilities for trigger signals as displayed in Figure 77. The main difficulty lies in the compression/decompression part. Some tests of this system have already been done with the liquid argon e.m. calorimeter and the hadronic tile calorimeter in ATLAS, and with the crystal calorimeter of CMS.

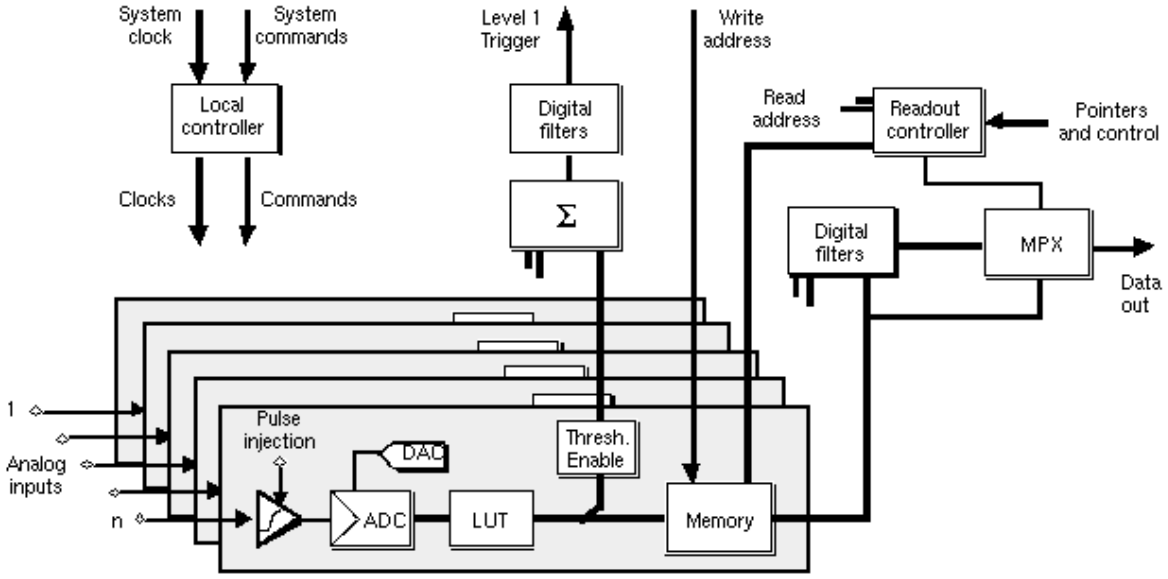

Figure 77: *The Fermi front-end system.*

- An equivalent solution can be performed with a 4-gain system (ratio 1, 3, 13, 50 typically) instead of a compression of the dynamic range. The data are stored with a track-and-hold system (or an analog delay) for about two clock cycles during which the best-suited gain is chosen depending on the energy in the cell. The signal is then digitized at 40 MHz with a fast ADC with a 10-bit range only. The major problem consists in finding the appropriate gain in a very short time without degrading the energy resolution. A solution of this type is planned in the NA48, KTeV and BaBar experiments.
- A third solution consists in storing the data in two parallel analog pipelines with gains 1 and 12 matching a 13-bit dynamic range (Figure 78). The data are kept inside the pipelines until a level I trigger decision is made. Accepted signals are then digitized by a 13-bit ADC. Tests with an ATLAS liquid argon e.m. prototype have already shown that with such a system the energy resolution reached is identical to the one already presented in the calorimeter section. Analog pipelines have been used successfully in the ZEUS experiment but with a lower dynamic range [83].

In all cases a careful understanding of the various sources of noise, coming mainly from the presence of both analog and digital signals on the same board, is the key problem. At the time of this report, the three techniques are still under development and some more evaluation tests are planned before a final choice is made in each LHC experiment.

In ATLAS the level I calorimeter trigger uses a reduced granularity, about 7000 cells with the elementary cells ganged in a 0.1×0.1 tower. The information of each trigger cell is digitized with an 8-bit dynamic fast ADC corresponding to 250 GeV full scale (above 250 GeV in a trigger cell the event is systematically stored for level II). A digital trigger processor runs a 2-D algorithm of cluster finding and applies a first isolation. Moreover, as the liquid argon signal extends over more than one bunch crossing, a digital pulse filtering (BCID) is used for bunch-crossing identification.

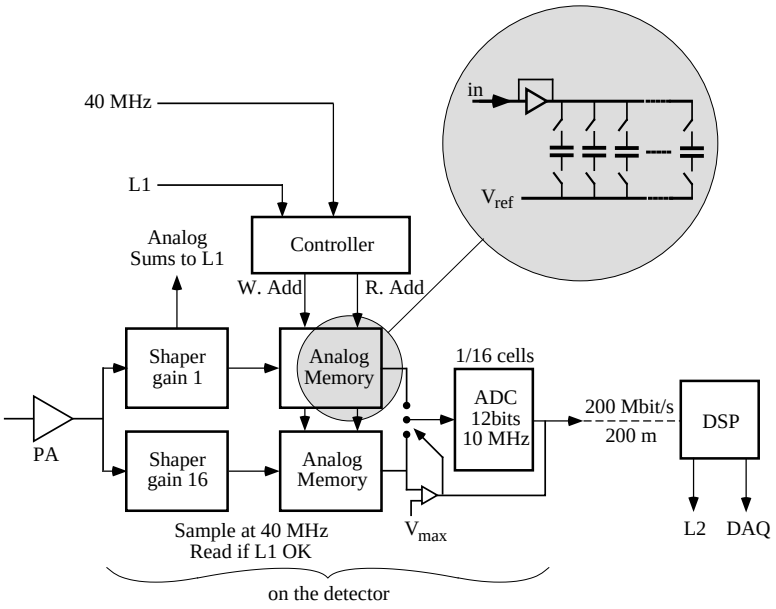

Figure 78: *Pipelines front-end system.*

5.2.3 Muon detectors

ADCs with 10-bit range (cathode strip chambers) or 8-bit TDC (drift tubes) fulfil easily the muon spectrometer requirements, which is a less challenging detector in terms of signal processing. The μ trigger is performed by coincidences between the layers of RPC of trigger stations (at different radius).

5.2.4 Data transmission – Mass storage

Trigger rate and data transmission speed are closely related. Each of the 7000 calorimeter trigger cell signals has to be extracted and sent to the logic board at 40 MHz, meaning 3500 fibres at 1 Gbit/s if digital data is transferred. After level I, the information has to move from the detector to the level II buffer (about 2000 boards) where a very high bandwidth networking is needed to combine the information from RoIs. Solutions are still under investigation. The events accepted by the level II trigger are sent to the event builder at about 1 to 10 GByte/s. A factor 10 reduction is expected with the level III processors which indicates that a 10^6 MIP processing power is required.

The mass storage of LHC data is about 15 TByte/day. The best recording devices up to now contain about 1 Gbyte in 3 cm^3 so that significant progress is necessary in storage capacity. However, no more than 1% of the data can be reasonably transferable to the participing institutes. A very strong reduction of the data and filtering should be made in situ at CERN and an optimized sharing of the capabilities in processing and simulation has to be set up between the collaborating institutes.

6 CONCLUSION

The goal of measuring accurately rare and complex processes in a high-energy, high-luminosity environment puts strong requirements on the LHC detectors.

The knowledge available in the high-energy physics community, complemented by strong detector development activities, and detailed simulations, have led to two new detector concepts, ATLAS and CMS. In some aspects, they are rather different and complementary:

- Solenoid against toroid for the muon detection.
- Crystals against noble liquid for e.m. calorimetry.

In other aspects, they use actually the same technique:

- Scintillating tiles for barrel hadronic calorimetry.
- Silicon strips and pixels for the tracking.

The pertinence of the choices, and of how they fit together, will be judged, in the end, by the quality of the results they produce.

We have said very little about other experiments, and this is certainly not doing justice to the clever developments which are being made for CP violation experiments, neutrino experiments, heavy-ion experiments, etc.

Nota Bene:

For the purpose of commenting on the detectors taken as illustration, numbers and sometimes interpretations have been given in this document. It is clear that they engage only the two authors, and in no case the original publications, to which the reader is invited to go for precise references.

References

[1] ATLAS technical proposal, CERN/LHCC/94–43 (1994).

[2] CMS technical proposal, CERN/LHCC/94–38 (1994).

[3] Textbooks on instrumentation:
Experimental techniques in high energy physics, Frontiers in Physics, T. Ferbel (ed.), Addison-Wesley (1987);
Instrumentation in elementary particle physics, C.W. Fabjan and J.E. Pilcher, World Scientific (1988);
Instrumentation in high energy physics, F. Sauli, World Scientific (1992);
Techniques for nuclear and particle physics experiments, W. Leo, Springer-Verlag (1993).

[4] The Large Hadron Collider CERN/AC/95–05 (1995).

[5] C. Albajar et al., Measurement of hadronic shower punchthrough in magnetic field, CERN-PPE/95–61 (1995).

[6] FLUKA code, A. Fasso et al., Proceedings of the IV International Conference on Calorimetry, World Scientific, Singapore (1994).

[7] ATLAS internal note TECH-NO-06 (1992).

[8] F. Bergsma et al., Nucl. Instrum. Methods Phys. Res. A357 (1995) 243.

[9] CEBAF-Report from 1986 study group.

[10] A. Yamamoto et al., Design study of a thin super-conducting solenoid for the SDC detector, KEK preprint 92–147 (Nov. 1992).

[11] V.A. Polychronakos, ATLAS internal note MUON-N0-38 (1994);
D. Carlsmith, CMS internal note TN-94-217 (1994).

[12] W. Blum, ATLAS internal note MUON-NO-24 (1993).

[13] P. Duinker et al., Nucl. Instrum. Methods Phys. Res. A273 (1988) 814.
[14] R. Santonico and R. Cardarelli, Nucl. Instrum. Methods Phys. Res. 187 (1981) 377.
[15] Review of Particle Properties, Phys. Rev. D50, No. 3 (1994).
[16] E. Longo and I. Sestili, Nucl. Instrum. Methods Phys. Res. A128 (1987) 128.
[17] R. Wigmans, Proceedings of the ICFA School (1987), World Scientific.
[18] M. Livan et al., CERN Report, CERN 95–02 (1995).
[19] D. Gingrich el al., Nucl. Instrum. Methods Phys. Res. A364 (1995) 290.
[20] RD36 Collaboration, CERN-PPE/95–152 (1995).
[21] K. Ahmet et al., Nucl. Instrum. Methods Phys. Res. A305 (1991) 275.
[22] B. Rossi, High Energy Particles, Prentice Hall (1952).
[23] U. Amaldi, Physica Scripta, 23 (1981) 409.
[24] C. W. Fabjan, Calorimetry in high energy physics, Frontiers in Physics, T. Ferbel (ed.), Addison-Wesley (1987).
[25] I. Wingerter-Seez, ATLAS internal note LARG-NO-18 (1995).
[26] T.S. Virdee, Proceedings of the Second International Conference on Calorimetry in High Energy Physics, Capri, World Scientific (1992).
[27] V. Fanti et al., Nucl. Instrum. Methods Phys. Res. A344 (1994) 507.
[28] R. Wigmans, CERN-EP/86–141 (1986).
[29] B. Andrieu et al., Nucl. Instrum. Methods Phys. Res. A336 (1993) 499.
[30] U. Behrens at al., Nucl. Instrum. Methods Phys. Res. A289 (1990) 115.
[31] F. Abe et al., Phys. Rev. D47 (1993) 4857.
[32] G. F. Knoll, Radiation detection and measurement, John Wiley & Sons (1989).
[33] H. J. Crawford et al., Nucl. Instrum. Methods Phys. Res. A256 (1987) 47;
J. Séguinot et al., CERN-LAA-PC/90–24 (1990).
[34] V. Aulchenko et al., Proceedings of the First International Conference on Calorimetry in High Energy Physics, World Scientific (1991).
[35] NA48 proposal CERN/SPSC/90–22 (1990).
[36] G. Barr et al., CERN-PPE/95–64 (1995).
[37] P. Cennini et al., Nucl. Instrum. Methods Phys. Res. A345 (1994) 230.
[38] Y. Kariotakis, presentation at the International Calorimetry Symposium, Beijing,25–27 October 1994.
[39] R. Zhu, G. Gratta and H. Newman, Nucl. Phys. B44 (1995) 88.
[40] E. Acker et al., Nucl. Instrum. Methods Phys. Res. A321 (1992) 60.
[41] Y. Kubota et al., Nucl. Instrum. Methods Phys. Res. A320 (1992) 66.
[42] BaBar technical report, SLAC-R-95-457, March 1995.
[43] BELLE letter of intent, KEK Report 94–2, April 1994.
[44] KTeV design report FN-580 (1992);
Kessler et al., Nucl. Instrum. Methods Phys. Res. A368 (1995) 653.
[45] R.Y. Zhu, Proceedings of the Crystal 2000 International Workshop, Editions Frontières (1992).
[46] RD18 'Crystal Clear Collaboration' status report CERN/DRDC/94–53, January 1995.
[47] A. Karar et al., Ecole Polytechnique, X-LPNHE/95–10 (1995).
[48] ALICE technical proposal, CERN/LHCC/95–71 (1995).
[49] J. Badier et al., CERN/DRDC proposal P50, CERN/DRDC 93–28 (1993).

[50] RD3 proposal, CERN/DRDC/90–31 (1990), CERN/DRDC/91-21 (1991), CERN/DRDC/92–40 (1992).
[51] J. Rutherfoord, GEM TN-91-27 (1991), GEM TN-93-410 (1993).
[52] R.L. Chase et al., ATLAS internal note LARG-NO-10 (1995).
[53] C. de La Taille, Electronic noise in LAr calorimetry, RD3 internal note 45 (1993).
[54] D.V. Camin et al., ATLAS internal note LARG-NO-11, (1995).
[55] R.L. Chase et al., Nucl. Instrum. Methods Phys. Res. A330 (1993) 228;
R.L. Chase et al., Nucl. Instrum. Methods Phys. Res. A343 (1994) 598.
[56] B. Aubert et al., Nucl. Instrum. Methods Phys. Res. A309 (1991) 438;
B. Aubert et al., Nucl. Instrum. Methods Phys. Res. A321 (1992) 467;
B. Aubert et al., Nucl. Instrum. Methods Phys. Res. A325 (1993) 116;
B. Aubert et al., Nucl. Instrum. Methods Phys. Res. A330 (1993) 405.
[57] D. Froidevaux et al., ATLAS internal note, PHYS-64 (1995).
[58] The D0 detector, Nucl. Instrum. Methods Phys. Res. A338 (1994) 185;
The D0 upgrade, Fermilab-Conf-177-E (1995).
[59] H. Aihara et al., IEEE Trans. Nucl. Sci. 30 (1983) 162.
[60] D. Buskulic et al., Nucl. Instrum. Methods Phys. Res. A360 (1995) 481.
[61] Vertex detectors, F. Villa (ed.), Ettore Majorana International Science Series, Plenum (1988); IEEE Trans. Nucl. Sci. 39 (1992) 629.
[62] S.M. Sze, Physics of semiconductor devices, J. Wiley & Sons (1981).
[63] F. Lemeilleur et al., Charge transport in silicon detectors, CERN/ECP 93–21 (1993).
[64] RD2 Collaboration, CERN/ECP/91–25 and IEEE Trans. Nucl. Sci. 39 (1992) 629.
[65] D. Amidei et al., Nucl. Instrum. Methods Phys. Res. A350 (1994) 75.
[66] N. Bingefors et al., Nucl. Instrum. Methods Phys. Res. A328 (1993) 447.
[67] L. Hubbeling et al., Nucl. Instrum. Methods Phys. Res. A310 (1991) 197.
[68] V. Chabaud et al., The Delphi microvertex with double-sided readout, CERN-PPE/95–86 (1995).
[69] A. Van Ginneken, Fermilab report FN-522 (1989).
[70] RD20 Status report, CERN/DRDC/94–39 (1994).
[71] G. Hall, Radiation damage to silicon detectors, Nucl. Instrum. Methods Phys. Res. A368 (1995) 199.
[72] J. Matheson et al., RD20/TN/36 (1994).
[73] E.M. Heijne et al., Nucl. Instrum. Methods Phys. Res. A349 (1994) 138.
[74] L. Rossi et al., ATLAS internal Note INDET-NO-85 (1994).
[75] K.M. Smith, Nucl. Instrum. Methods Phys. Res. A368 (1995) 220.
[76] A. Oed, Nucl. Instrum. Methods Phys. Res. A367 (1995) 34.
[77] RD28 progress report, CERN/DRDC/95–86 (1995).
[78] O. Adriani et al., Performance of a prototype of the CMS central tracker, Nucl. Instrum. Methods Phys. Res. A367 (1995) 189;
S.F. Biagi et al., A study of the response of thin oxyde MSGCs, Nucl. Instrum. Methods Phys. Res. A367 (1995) 193.
[79] B. Dolgoshein, Transition radiation trackers for ATLAS and HERA-B, Nucl. Instrum. Methods Phys. Res. A368 (1995) 239.
[80] A.F. Fox-Murphy et al., Frequency Scan Interferometry, ATLAS internal note INDET-NO-112(1995).

[81] D. Fournier, Noble liquid e.m. calorimetry, Nucl. Instrum. Methods Phys. Res. A367 (1995) 5.
[82] H. Alexanian et al. (Fermi Collaboration), ATLAS internal note DAQ-NO-16 (1994).
[83] A. Caldwell et al., Nucl. Instrum. Methods Phys. Res. A321 (1992) 356.