The Radiation Environment at the CMS
Experiment at the LHC

Mika Huhtinen
CERN
CH-1211 Geneva, Switzerland

Dissertation for the degree of Doctor of Technology to be presented with due permission for public examination and debate in Auditorium F1 at Helsinki University of Technology (Espoo, Finland) on the 31st of May, 1996, at 12 o’clock noon.

Helsinki 1996
This work was started and to a large extent performed at the Research Institute for High Energy Physics of the University of Helsinki. The initiative to study the radiation environment at CMS came originally from Risto Orava. In the course of the work the encouragement and support of Jorma Tuominiemi, especially his insistence to concentrate all of my efforts on the radiation studies, became more and more important. Some of the work was completed during my Fellowship in the Radiation Protection Group at CERN. I am grateful to Manfred Höfert for the possibility to spend some time on CMS-specific problems and work for this thesis, even though my other duties have necessarily suffered from this.

Without relaxing any of his requirements on scientific quality, my supervising professor, Rainer Salomaa, has always shown a flexible attitude regarding the contents and time-table of my postgraduate studies and this thesis work.

Among numerous other members of the CMS collaboration Thomas Meyer, Ernst Radermacher and Tejinder Virdee carry a major responsibility on the contents of this thesis, since it was them, who posed most of the interesting questions to be studied.

During the last two years I have had the privilege to work with Graham Stevenson and to profit from his expertise in the fields of radiation protection, induced radioactivity and the shielding and safety aspects at hadron accelerators. From Alberto Fassò I have learned a great deal on advanced Monte Carlo techniques.

I warmly remember the hospitality of the PSI staff during our pion irradiations and especially the invaluable help of Roland Horisberger.

But amongst them all, it is easy to name the two persons who have had a major influence on all of my studies and research activities: Pertti Aarnio and Martti Pimiä.

Since 1987, when I started as a summer student in the Laboratory for Nuclear Engineering, Pertti has supervised my studies and gradually evolved from a teacher to a close collaborator and co-author. As for Martti, I think he has only rarely – if ever – given professional advice on the subject of this thesis. But I suspect he carries a non-negligible responsibility for getting me into a position within the CMS collaboration, where there was no way out of doing the work, which is now summarized in this thesis. Certainly I also have enjoyed our common studies of French Wine & Food.

But after all it has to be finally confessed, that in largest debt I am to my wife Sirpa, who has not only taken care of me, but also of our little daughter Kukka-Emilia.

Grants from the Emil Aaltonen Foundation and the Jenny and Antti Wihuri Foundation are gratefully acknowledged.
Abstract: A review of the radiation environment, expected at the CMS experiment at the LHC, is given. Special emphasis is put on the radiation exposure and the resulting radiation damage of silicon detectors. Computational methods to estimate the radiation damage in silicon due to energetic hadrons are discussed and an experiment aimed at determining the damage constants for positive pions is described. Results of this irradiation experiment are discussed and it is shown that the life-expectancy of silicon detectors at LHC is a critical parameter for the design of the experiments. The radiation background in the muon spectrometer of CMS is discussed and methods to reduce it to tolerable level are proposed. These include massive shielding and an optimized beam pipe geometry. The dose rates and neutron fluxes in the calorimeters of CMS are reviewed and first estimates of induced radioactivity are given.
## Contents

1 Introduction 1
   1.1 Physics at 1 TeV and beyond ........................................ 1
   1.2 Multi-TeV Colliders .................................................. 2

2 The Large Hadron Collider 3
   2.1 Machine Parameters ................................................... 3
   2.2 Minimum Bias Events at the LHC .................................... 5

3 The CMS Experiment 7
   3.1 The Detector .......................................................... 7
   3.2 The Experimental Region ............................................ 8

4 Simulation of Cascade Propagation 10
   4.1 Hadronic Interactions in Matter .................................... 10
   4.2 Particle Transport .................................................... 12
   4.3 Variance Reduction Methods ........................................ 13
   4.4 Fluxes, Currents and Doses .......................................... 13

5 Shielding at Hadron Accelerators 14
   5.1 Neutrons ............................................................... 14
   5.2 Photons ............................................................... 16
   5.3 Charged Particles .................................................... 16

6 Hadron Fluxes in the Tracking Cavity and around the ECAL 17

7 Displacement Damage in Silicon Detectors 19
   7.1 Non-ionizing Energy Loss ............................................ 20
   7.2 Calculation of Displacement Damage ............................... 21
   7.3 Measurement of Pion Induced Displacement Damage .............. 23

8 Background in the CMS Muon System 25
   8.1 Sensitivity Factors .................................................... 26
   8.2 Background Suppression .............................................. 28
   8.3 Beam Pipe Geometry .................................................. 28

9 Dose Rates in the Detectors 29

10 Shielding of Occupied Areas during LHC Operation 30

11 Induced Radioactivity 31

12 Summary 34
This thesis forms an introduction to the following publications.


The paper describes simulation results of the charged hadron and neutron fluxes to be expected in the tracking cavity of the CMS experiment. It is the first paper which clearly points out the fact that pion fluxes can dominate silicon damage close to the interaction point at LHC. Results reported in this paper were also used in the CMS Letter of Intent [1].

The disputant did the calculations and wrote most parts of the paper.


The paper gives a summary of the work on predicting pion induced displacement damage in silicon [2]. Such a prediction was motivated by the observations made in publication I. This paper, which is the first attempt to estimate the pion induced displacement damage in silicon at energies around the $\Delta$-resonance, emphasizes the need for an experimental measurement of the pion damage at low energies.

The disputant did the calculations and wrote the paper.


The paper describes the pion irradiation experiments motivated by the results of publications I and II and performed at the Paul Scherrer Institute. Among the results of the other groups which performed independent irradiations within the same SIRAD collaboration, this paper is the first to report experimental results of pion induced displacement damage in silicon.

The disputant was involved in proposing the experiment, participated in the irradiation, did the IV analysis, co-supervised the CV analysis and wrote major parts of the paper.


The paper describes results of an optimization study of the LHC beam pipe materials and geometry. These studies form a part of the simulations of the neutral background in the muon system of CMS which are summarized in publication V. Based on the results reported in this paper the CMS experiment has adopted a conical beam pipe in the Technical Proposal [3].

The disputant did the radiation calculations and wrote corresponding parts of the paper.


The paper provides a summary of a dedicated simulation project to design a shielding system which reduces the neutral particle background in the CMS muon spectrometer to acceptable levels. It also provides an update of publication I by providing a more systematic study of the effects of moderators, including the new PbWO$_4$ calorimeter option. The shielding configurations proposed in this paper formed the basis for the shielding as defined in the CMS Technical Proposal [3].

The disputant did the calculations and wrote the paper.


The paper provides an overview of the radiation environment simulation performed for the CMS experiment and reviews the shielding proposed in the Technical Proposal. It also addresses issues of personnel radiation exposure and gives some indications of the activation of materials due to LHC operation. The disputant did the calculations and wrote the paper.
1 Introduction

The Large Hadron Collider (LHC) will be the next accelerator to be built at the European Laboratory for Particle Physics (CERN). According to the present time-table its commissioning will take place in 2004.

In addition to the discovery of the Higgs boson or any alternative symmetry-breaking mechanism the main tasks anticipated for the LHC are studies of CP-violation in the B-sector and an exploration of alternatives to the Standard Model of particle physics.

When operating at its peak luminosity, the LHC will produce almost $10^9$ proton-proton collisions per second. These create an extremely hostile radiation environment around the experiments. Since no large-scale particle physics experiment before has been confronted with radiation fields of comparable intensity, completely new detector concepts are needed for the LHC. A thorough understanding of the radiation environment at the LHC is required to guide the design of detectors and of the radiation shielding.

Most aspects of the radiation environment at the Compact Muon Solenoid (CMS) experiment are described in the publications of this thesis [PubI, PubII, PubIII, PubIV, PubV, PubVI]. This introductory paper is intended to be a general review of the simulation techniques, of radiation physics and of the radiation related problems encountered at the LHC.

In the following the motivation for building the LHC is briefly recalled. Chapters 2 and 3 describe the LHC machine and the CMS experiment. In chapter 2.2 the characteristics of LHC events and some available Monte Carlo event generators are discussed. Chapter 4 is a general review of the state of the art in hadronic shower simulation and chapter 5 discusses shielding strategies at hadron accelerators. From chapter 6 onwards the publications of this thesis are discussed in roughly their order of appearance. Emphasis is put on recent changes in the design of CMS and latest refinements to the calculations, which partly have not yet been published.

1.1 Physics at 1 TeV and beyond

From the very beginning of science, the structure of the universe and of the matter surrounding us has been subject to intensive research. The Greeks of antiquity postulated that there must be a basic building block of nature, which cannot be further subdivided. It was called the atom. At the end of the 19th century scientists indeed believed to have found those smallest pieces of matter, which thus deserved the name atom. But soon thereafter the atom was found to have an internal structure, including a nucleus surrounded by a cloud of electrons. Even the atomic nucleus was shown to be an assembly of smaller particles, protons and neutrons. It took until the 1960s before the substructure of protons, neutrons and other hadrons was verified. Electron scattering experiments showed that they consisted of small point like objects – quarks.

As a result of the fast progress in science and technology, which has taken place during the last 100 years, physicist nowadays use the worlds’ largest experimental facilities to discover ever finer details of matter. Large accelerators are needed to achieve sufficiently high energies. Like a microscope a particle beam is able to probe the structure of matter only at a scale comparable to the wavelength $\lambda = h/p$ of the particles.

According to our present knowledge the basic building blocks of matter are leptons and quarks grouped in three families of 2+2 each. The forces between these elementary particles are mediated by gauge bosons. Some of these are massless, like the photon, while others, like the $W^\pm$ and $Z^0$, are among the heaviest particles so far discovered. All these particles are collected in Table 1. The $Z^0$ can decay into any kind of neutrino which is lighter than half the $Z^0$ mass. The lifetime of the $Z^0$, as measured at the Large Electron Positron Collider (LEP), indicates that there are exactly three neutrino species accessible and so the existence of more than the three families listed in Table 1 is very unlikely.
Table 1: Discovered particles of the Standard Model. Evidence for $\nu_\tau$ is only indirect [4]. All common matter is composed of particles belonging to the first family. Each quark and lepton also has a corresponding antiparticle.

The particles of Table 1 and the interactions between them are governed by the Standard Model (SM), which forms the cornerstone of particle physics. Although having been extremely successful in predicting and explaining all experimental discoveries so far, the SM has some serious drawbacks. In addition to including numerous free parameters, it has to postulate a mechanism known as spontaneous symmetry breaking in order to allow for particles to have mass. This mechanism, however, involves the introduction of a new particle: the Higgs boson ($H$).

Although the largest existing particle accelerators, the LEP at CERN and the Tevatron at Fermilab, have provided deep and detailed insight into the structure of the Standard Model, the Higgs-boson has until now escaped all attempts of detection. Since in the Standard Model unitarity constraints require a Higgs type particle at an energy $\lesssim 1$ TeV [5], the LHC should uncover this presently anticipated “last secret”.

The recent experimental discovery of the top quark at the Tevatron implies that a SM Higgs is likely to have a mass in the range 100–400 GeV. This does not exclude the detection of a Higgs at LEP2 well before commissioning of the LHC, but means that most likely such a Higgs would not be of the simple SM type.

The most popular extension to the SM, the Minimal Supersymmetric Standard Model (MSSM), for instance would predict two neutral and two charged Higgs particles. If any of the supersymmetric Higgs particles would be detected at the LEP upgrade (LEP2), it would immediately imply the existence of a rich spectrum of new physics beyond the LEP2 reach: in the MSSM each particle known to us would have its supersymmetric partner. The observation of these particles and thus the verification of the MSSM model would then form the main task of the LHC.

### 1.2 Multi-TeV Colliders

In electron-positron colliders, like LEP, the colliding particles annihilate and all beam energy is available for production of new particles. Therefore the discovery potential of $e^+e^-$ colliders is huge even with relatively low beam energies.

In hadron colliders two composite particles, consisting of quarks and gluons, commonly called partons, traverse each other. Some of the partons can exchange momentum, i.e. scatter from each other. If the parton-parton collision is head-on the momentum exchange can amount to their full available energy. The other constituents of the colliding hadrons, however, continue with almost their original momentum. Hence only a small – but random – fraction of the total beam energy is available for particle production. With a tiny probability one parton can carry almost the total momentum of the hadron. If, however, the scattering partner
has a lower momentum the collision is strongly boosted and most of the energy is wasted into longitudinal momentum of the produced secondaries, much like in fixed target experiments. In order to collect at least a few of the rare events where both colliding partons have roughly equal and sufficiently high momenta, the total number of hadron-hadron collisions has to be maximized. This results in events being often accompanied by a huge pile-up of other simultaneous collisions. Therefore hadron colliders, compared to $e^+e^-$ machines, provide a particularly “dirty” environment. Additional experimental complications arise from the fact that the uncertainty in the momentum of colliding partons prevents the use of longitudinal momentum conservation.

Circular $e^+e^-$-machines reaching the 1 TeV energy scale, however, are not feasible, since the energy loss due to synchrotron radiation increases as the fourth power of the inverse of the rest mass of the particle [4]. LEP2, operating at the limit of which can be considered feasible, can reach a beam energy of $\sim 100$ GeV. In order to achieve an energy of 500 GeV, the circumference of the ring would have to be increased to 135 km. Thus linear accelerators provide the only alternative to obtain $e^+e^-$ collisions at 1 TeV. Their feasibility has been extensively studied during the last years, but the realization of such a facility is unlikely to be seen before the third decade of the next century.

Since $m_p/m_e = 1836$, protons are essentially not affected by synchrotron radiation losses and the energy limit for a given ring diameter of a proton collider is posed by the bending power of the magnets. Thus a circular $pp$-collider reaching far into the TeV regime can be obtained with comparatively moderate cost.

The environment at a hadron collider does not allow the same kind of precision measurements to be made, as has been and will be practiced at LEP and other $e^+e^-$-colliders. Due to their wider energy reach hadron colliders can be seen as exploratory machines, whereas $e^+e^-$-colliders are needed for subsequent high precision measurements. Consequently, results obtained at the LHC certainly will influence the design and energy coverage of any future high energy $e^+e^-$ facility.

## 2 The Large Hadron Collider

### 2.1 Machine Parameters

The LHC [6] has been approved as the next joint worldwide effort to provide the first look into matter at the mass scale of 1 TeV. The LHC, to be installed in the existing LEP tunnel after the decommissioning of LEP2, will provide the highest energy densities in the center of mass, ever created artificially. The schematic layout of the LHC is shown in Fig. 1.

The LHC will collide proton beams of 7 TeV energy. In order to create a sufficient number of high-energy parton-parton collisions, the LHC is designed to reach an unprecedented peak luminosity of $10^{34}$ cm$^{-2}$s$^{-1}$ at its two high luminosity interaction points in octants 1 and 5.

The existing LEP tunnel fixes the circumference of the LHC to 26.659 km. A total of 1104 superconducting 8.4 T dipole magnets of 14.2 m length are needed to force the proton beams into paths with bending radii of 2784.32 m in each of the eight arcs. Eight superconducting cavities operating at a peak voltage of 2 MV will supply a maximum energy of 334 MJ to each of the two circulating beams.

This energy is far more than sufficient to cause significant damage to any accelerator component. In order to prevent accidental losses of the full beam, a fast dumping system with a kicker magnet rise time of less than 3 $\mu$s has been designed. In order to minimize induced activity and to obtain some longitudinal dilution of the cascade the beam dump will have a graphite core surrounded by aluminium and iron shielding. Since no material would be able to sustain the dumping of the full LHC beam into one spot, a scanning system is
needed which sweeps the beam over the face of the dump so that no part of the graphite is heated above the vaporization threshold.

It has been estimated that the superconducting magnets would quench if the rate of beam particles lost into a small azimuth region of the beam pipe exceeds $7 \times 10^6$ s$^{-1}$ m$^{-1}$ [6]. In order to guarantee a safe operation of all magnets along the ring, the LHC needs a highly efficient beam cleaning system. This is achieved by two cleaning insertions, with betatron cleaning in octant 3 and momentum cleaning in octant 7. The first one restricts the beam envelope to $10 \sigma$, whereas the second ensures that the momentum spread does not exceed $|\Delta p/p| = 3 \times 10^{-3}$.

Most of the created $pp$-events do not contain any new physics but add up to a huge total track density and energy deposition in the experimental area. This implies significantly more severe radiation problems than would be encountered at an $e^+e^-$ collider with the same discovery potential. While the immense background at the LHC sets new standards for accelerator, detector and data-acquisition technology, it also requires more detailed and careful predictions of the radiation environment than any existing high energy accelerator [7]. This of course includes assessments of radiation safety, but to an even larger degree the estimation of performance and aging of the detectors to be used in this hostile environment. Radiation damage and induced activity are related to integrated luminosity, whereas occupancies and run-time radiation exposure depend on the instantaneous luminosity. For all radiation issues, however, the bunch structure of the LHC beam is insignificant. The most important parameters for an assessment of radiation issues are collected in Table 2.

The nominal beam current of 530 mA is achieved with 2835 bunches of 7.5 cm length containing $1.05 \times 10^{11}$ protons each. The bunch spacing is 24.95 ns. At the interaction vertices the beams cross at an angle of $200 \mu$rad. The interaction spot has a horizontal and vertical rms-spread of 16 $\mu$m. The rms-spread along the beam axis is 54 mm.

For an inelastic $pp$ cross section of 70 mb the peak luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ implies $7 \times 10^8$ interactions per second. The nominal beam and luminosity lifetimes are 22 h and 10 h, respectively. This leads to a day-averaged luminosity of roughly half the peak value, which can be achieved with either one fill of 20 h or two fills of 8 h [8]. Assuming 180 days of operation per year [8] the annual number of $pp$ interactions is $5 \times 10^{15}$. Another common definition for a “LHC year” is that the machine operates at a luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ for $10^7$ seconds. This gives a total of $7 \times 10^{15}$ events, which is in reasonable agreement with the previous definition.

The LHC will not reach its nominal luminosity immediately after commissioning. An assumption recommended for assessments of the radiation environment is that in the first year 10% of the design luminosity is reached. This increases to 1/3 and 2/3 during the following two years and from the fourth year onwards the LHC operates at its nominal luminosity [9]. The time required to obtain a physics discovery, i.e. to have enough statistics accumulated, is best measured in terms of integrated luminosity. The standard LHC physics programme is based on an integrated luminosity of $\int \mathcal{L} dt = 5 \times 10^5$ pb$^{-1}$. Taking into account the low luminosity startup phase, the numbers in Table 2 suggest that this can be reached in 9 years.

The LHC has been designed to be able to operate also as an ion-ion collider. The energy per charge unit can be maintained at 7 TeV, giving for lead ions $E/A=2.76$ TeV. The luminosity can reach $1.95 \times 10^{27}$ cm$^{-2}$ s$^{-1}$, i.e. almost seven orders of magnitude lower than for the $pp$-option. Although the Pb-Pb cross section and average multiplicity are by factors of $\sim 35$ and $\sim 200$ larger than the $pp$ ones, the average radiation background during a Pb-Pb run remains three orders of magnitude below that of the high luminosity $pp$ operation. However, the instantaneous track density with which an ion experiment is confronted during a central Pb-Pb event is much higher than at the $pp$ experiments. This imposes the use of high granularity detectors, especially in the tracker region, but radiation damage or safety problems are less significant than for the $pp$-experiments.
2.2 Minimum Bias Events at the LHC

Except for the small contribution from machine related background the radiation field around the high luminosity LHC experiments is determined by the minimum bias events, i.e. the bulk of all events, which usually do not contain any interesting physical signatures.

Since no existing hadron collider is able to provide collisions with 14 TeV energy in the center of mass, there is no exact knowledge about the $pp$ cross sections and the structure of minimum bias events at the LHC. Simulation codes tuned to available experimental data at lower energies and cosmic ray measurements, as well as analytical extrapolations, agree on the fact that the inelastic non-diffractive cross section should be around 70 mb to which diffractive events add $10$ mb [7, 18].

Even more uncertainty than in the cross sections is to be found in the multiplicity of events. A commonly assumed charged particle multiplicity in the central pseudorapidity range is 6 per unit of $\eta$, but depending on the structure functions different event generators can give anything between 4 and 10. Both of these extreme values certainly are most unlikely to be true. However, the cross sections and multiplicities tend to be correlated and should always be considered together [19].

Fig. 2 shows a comparison of charged multiplicities and transverse momentum spectra between the three event generators most frequently used for LHC studies. The DTUJET92 generator has been replaced by a more recent version of the same code, DTUJET93 [20], which has a more accurate treatment of minijets [18] and is therefore the recommended one [21]. The PYTHIA 5.7 [22] sources are obtained with parameters specially tuned for LHC physics studies [19]. Table 3 further illustrates the uncertainties in the particle contents obtained with the three generators. In all cases single diffractive events have been excluded. The “standard” $p_T$ cut of 150 MeV/c excludes some particles, which are important for the inner tracker. Values without this cut are shown in parentheses [7].
Figure 2: Transverse momentum and multiplicity distribution of events expected at the LHC. Curves are scaled to correspond to 70 mb cross section.

<table>
<thead>
<tr>
<th></th>
<th>DTUJET92</th>
<th>DTUJET93</th>
<th>PYTHIA 5.7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>70 mb</td>
<td>70 mb</td>
<td>65 mb</td>
</tr>
<tr>
<td>$p_T$ (GeV/c)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$dN/dp_T$ ($\eta&lt;2.5$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proton</td>
<td>2.39</td>
<td>1.30</td>
<td>1.71</td>
</tr>
<tr>
<td>Neutron</td>
<td>2.35</td>
<td>1.39</td>
<td>1.58</td>
</tr>
<tr>
<td>$\pi^+ \pi^-$</td>
<td>25.71</td>
<td>22.62</td>
<td>24.67</td>
</tr>
<tr>
<td>$K^+ K^-$</td>
<td>2.33</td>
<td>2.89</td>
<td>2.65</td>
</tr>
<tr>
<td>$K^0$</td>
<td>2.34</td>
<td>2.78</td>
<td>1.33</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>29.26</td>
<td>26.82</td>
<td>17.60</td>
</tr>
<tr>
<td>$e^+ e^-$</td>
<td>0.37</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>$\mu^+ \mu^-$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Charged hyperon</td>
<td>1.13</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>Neutral hyperon</td>
<td>1.06</td>
<td>0.45</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>66.94</td>
<td>58.74</td>
<td>49.68</td>
</tr>
<tr>
<td>Total charged</td>
<td>31.92</td>
<td>27.30</td>
<td>29.16</td>
</tr>
</tbody>
</table>

Table 3: Particle multiplicities in the central and forward regions for non-diffractive events. A $p_T$ cut of 150 MeV/c has been applied. Hyperon decays are performed in the PYTHIA run, but not in the DTUJET runs. PYTHIA multiplicities have been rescaled to $\sigma = 70$ mb.

Cut [19] show that the PYTHIA events contain significantly more low $p_T$ particles than the DTUJET events. If particles down to $p_T=0$ are included PYTHIA gives a significantly higher multiplicity, in particular the photon multiplicity roughly doubles [19].

It has been shown that differences which arise from the event generators alone can introduce an uncertainty of $\sim 30\%$ to the flux of $pp$-secondaries [19]. Except for the regions close to the vertex, which are directly exposed to the $pp$-secondaries, uncertainties in the geometry description and in the physics models used for cascade propagation are likely to dominate.
Figure 3: Longitudinal cut through the CMS detector[3]. EB, EF = electromagnetic calorimeter, HB, HF = hadron calorimeter, MB, MF = muon stations, VF = very forward calorimeter. The polar angle, \( \theta \), is usually replaced by pseudorapidity \( \eta = -\ln \tan(\theta/2) \).

3 The CMS Experiment

3.1 The Detector

The Compact Muon Solenoid (CMS) was proposed as a general purpose high luminosity experiment for the LHC in 1990 [12]. Since then the detector has evolved over several intermediate designs to its present day form as defined in the Technical Proposal [3] and shown in Fig. 3. Despite some improvements over the years it is the only one of the four detectors having presented an expression of interest in 1992 [13] which now has received conditional approval without any major revision of the original design. In all its essential parts, i.e. a simple and redundant muon system, a high precision inner tracker and all calorimetry inside a strong solenoidal magnetic field, the detector is very similar to the first proposal.

At the core of CMS is the superconducting solenoid providing a 4 T magnetic field of 13 m length and 5.9 m diameter. The iron of the flux return yoke houses four layers of muon chambers.

The high magnetic field provides 12 Tm bending of power up to a pseudorapidity of \( \eta = 1.5 \). Even at \( \eta = 2.4 \), which is the acceptance limit of the muon spectrometer, the bending power exceeds 4 Tm. This facilitates a precise muon momentum measurement without strong demands on chamber accuracy. Since the muon identification is performed in the iron, the overall design is very compact.

The present design of the central tracker includes in the barrel region two layers of silicon pixel detectors at radii of 7.7 cm and 11.5 cm. These are followed by a three-layer silicon tracker and further 7 layers of microstrip gas chambers (MSGC). The support structure is provided by a carbon fibre space frame. In the forward region detectors are perpendicular to the beam and the total number of planes is 14-16, enough to ascertain the detection of 10–12 hits per straight track. The beam pipe in the central region is beryllium in
order to minimize multiple scattering before the first sensitive detector layer. Especially to facilitate the detection of the potentially important decay channel $H \rightarrow \gamma \gamma$, strong demands have been set on the resolution of the electromagnetic calorimeter (ECAL). The best solution to achieve the required performance is a homogeneous crystal calorimeter. In the Letter of Intent [1] the preferred crystal was CeF$_3$, which has superior performance but the needed $\sim$25 m$^3$ of CeF$_3$ were found to lead to an intolerable total cost of the ECAL.

Among the cheaper solutions has been a more conventional lead/scintillator option ("shaslik") [14]. Although the performance of such a sampling calorimeter is not as good as that of a homogeneous crystal, it is by far the cheapest of the alternatives considered. In the Technical Proposal PbWO$_4$ is adopted as the crystal for the ECAL [15]. It is significantly cheaper than CeF$_3$ but – except for a smaller light yield – has similar properties [15].

Except for narrow cable paths and a few alignment holes the ECAL covers hermetically the region up to $|\eta| = 3$. Its task is to absorb and measure all the energy carried by primary photons, electrons and positrons. While 23 cm of PbWO$_4$ corresponds to almost 26 radiation lengths $\chi_0$, it amounts to just a little more than one hadronic interaction length $\lambda_I$.

The operation of the muon chambers relies on the fact, that they are only rarely reached by particles other than muons. In order to achieve the required suppression of hadronic cascades $\sim 10 \lambda_I$ are required between the vertex and the first muon station. Hadron calorimetry is not likely to be an essential tool for the detection of the Higgs which over almost all of the expected mass range would be discovered through its decay into photons or muons. A hadron calorimeter with good resolution and large coverage of solid angle can, however, be vital for the detection of possible supersymmetric particles. For these one important signature is missing transverse energy. Since the CMS hadron calorimeter (HCAL) is situated inside the coil, it has to be non-magnetic. The solution adopted by CMS is a copper/scintillatorsampling calorimeter.

The first muon chamber provides an accurate muon momentum determination by measuring the exit angle of the muon from the central magnetic field. Since the vertex point is known in the transverse plane this angle gives the bending radius of the muon. Despite the attenuation provided by the calorimeters the first muon stations, especially MF1, are exposed to hadronic punchthrough which may influence the measurement. In particular, if the muon is within a jet, it may be impossible to obtain a reliable direction measurement.

A second independent muon momentum determination is provided by the full set of four muon stations. Their information can be combined to a curvature measurement in the saturated (1.8 T) iron yoke. Although multiple scattering in iron limits the accuracy of this measurement, the advantage is that the massive yoke – together with the requirement of a signal in at least three out of the four stations – significantly reduces the hadronic background. The information of the outer muon stations is also essential for a reliable muon identification.

In $\gtrsim 95\%$ of the cases the track segments observed in the muon stations can be matched with information from the inner tracker. In the region $|\eta| < 2.0$ this three-fold redundancy provides a momentum resolution of 0.5–10% over the transverse momentum range $p_T=10$–1000 GeV/c. The information from the tracker is also used to resolve secondary vertices. This is essential for B-physics but is also needed to verify that muons indeed have originated from the primary vertex, as assumed by the momentum determination in the first muon stations.

### 3.2 The Experimental Region

The CMS detector will be housed in an underground hall in octant 5 of the LHC, almost 100 m below the earth surface. The experimental hall has a diameter of $\sim 26$ m and a full length of 53 m. The free space between the
low-beta quadrupoles, which provide the high luminosity at the interaction point, is 46 m. The hall geometry, as it is approximated for the radiation environment simulations, is shown in Fig. 4.

Two shafts are needed for lowering detector components to the underground area. The larger main shaft, which is elliptic with major and minor axes of 23 and 17 metres, penetrates directly into the experimental hall. It has only negligible influence on the radiation field around the detector, but needs to be considered when studying the environmental impact of the LHC [16]. In order to prevent “sky-shine”, i.e. neutrons scattering back from the atmosphere onto the earth surface, this shaft has to be sealed with a concrete plug during LHC operation [16]. The smaller side shaft has a circular cross section and leads into the service cavern next to the experiment.

Table 4 shows a rough energy partitioning in the experimental region as calculated from direction cosines and energies of the secondaries in the simulated pp events. About 95% of the energy is carried by ~25% of the pp-secondaries into the very forward direction. Here these energetic particles would impinge directly on the superconducting quadrupoles. The estimated power, which would be deposited in each of the first quadrupoles (Q1), is ~250 W which is far beyond the quench limit. Copper collimators of ~2 m length will be positioned between the IP and the magnets to reduce the power absorbed in the coils. With their inner radius of only 1.5 cm these collimators also prevent any accidental beam loss to impinge directly on the experiments.

In order to open the CMS detector most of the shielding around the collimators has to be removed so that the endcap can be moved up to the level of Q1 [3]. In order to minimize the time required for this operation, special emphasis has been put on having a movable shield between the VFCAL and the Collimator region. The very forward calorimeters will be moved to specially prepared excavations situated under the collimators.

Beam particle losses at limiting apertures together with beam-gas events around the ring will generate a halo in the LHC tunnel, which could affect the experiments. Except for high energy muons all components of this machine background can be suppressed by sealing the tunnel entry with concrete blocks. Most of the low-β string will reside inside the tunnel, so the sealing is needed also to reduce the relatively intense background emerging from these machine elements [17].
Table 4: Average incident energy (in GeV per event) and multiplicity (in % of total) in different regions of the CMS detector and hall according to the event generators discussed in chapter 2.2. The numbers are obtained from direction cosines of $pp$-secondaries.Decays and the magnetic field inside the detector may slightly affect these values. The values in the right column of DUJET93 include diffractive events.

4 Simulation of Cascade Propagation

4.1 Hadronic Interactions in Matter

The inelastic interaction of a fast hadron with a nucleus can be divided into three quite distinct phases. In the multi-GeV energy domain the first phase consists of interactions at parton level. The quarks of the projectile interact with the quarks of some of the nucleons in the target nucleus. Due to confinement these interactions do not lead to free quarks, but the strings start to fragment and finally hadronize into mesons and baryons. The final state of this first phase consists of a shower of fast particles and the target residual. This first phase can to a good approximation be described as a superposition of several elementary hadron-nucleon collisions. Since the hadronization of the strings takes a finite time, most leading secondaries are formed outside of the target nucleus and high energy interactions of a second generation have a reduced probability [23].

At a slightly longer time scale the nucleons of the target residual undergo a reordering process. Some of them may have been knocked by the projectile hadron and have sufficient energies to leave the nucleus or knock neighbouring nucleons out. In general this intranuclear cascade terminates with a nucleus still far from the valley of stability. Although it can maintain a substantial excitation energy the residual approaches quickly thermal equilibrium. This pre-equilibrium phase is the most difficult one to treat theoretically, since it is a kind of mixture between high energy and nuclear physics involving both, particle production and nuclear effects. Most of the high-energy approximations are not valid but at the same time energies are high enough to prevent the use of nuclear structure models. Recent pre-equilibrium intranuclear cascade codes have shown promising results in this important energy range between few MeV and few GeV [24].

The last phase of the hadronic interaction consists of particle evaporation from the residual nucleus at thermal equilibrium. The evaporation probabilities are determined by the amount of available excitation energy, the mass difference between initial and final states and the height of the Coulomb barrier. Due to the last constraint neutrons are the most abundant particles in the evaporation spectrum. After evaporation the target nucleus will be either stable or – which is more likely – in the neighbourhood of stable nuclei. Some excitations decay by prompt photon emission. Multifragmentation and fission are processes which compete with evaporation [25]. Except for the heaviest nuclei, which are close to the fission barrier, both processes usually require excitation energies of several MeV per nucleon.

In principle radioactive decay of the final residual nucleus can be added as a fourth phase to the chain described above. This, however, takes place at a time scale which, depending on the nucleus, can be anything between a fraction of a second and millions of years and is not considered in any particle physics simulation code.
Figure 5: Left plot: equidose contours of $\pi^+$ initiated cascades tungsten. Pion energies are 100 MeV, 1 GeV, 10 GeV and 100 GeV. The contours are spaced by one decade and range from $10^{-13}$ Gy/primary to $10^{-16}$ Gy/primary. Integration in the transverse plane is performed over ±2 cm. Right plot: corresponding radially averaged neutron fluxes. Table below: spectrum of charged hadrons hitting the barrel and endcap regions of the CMS ECAL. Numbers are in percent.

<table>
<thead>
<tr>
<th>Energy</th>
<th>&lt; 0.1 GeV</th>
<th>0.1 – 0.32 GeV</th>
<th>0.32 – 1 GeV</th>
<th>1 – 3.2 GeV</th>
<th>3.2 – 10 GeV</th>
<th>10 – 32 GeV</th>
<th>32 – 100 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barrel</td>
<td>15.30</td>
<td>22.32</td>
<td>28.67</td>
<td>28.20</td>
<td>5.24</td>
<td>0.23</td>
<td>0.04</td>
</tr>
<tr>
<td>Endcap</td>
<td>6.33</td>
<td>17.60</td>
<td>32.45</td>
<td>31.45</td>
<td>10.55</td>
<td>1.55</td>
<td>0.07</td>
</tr>
</tbody>
</table>

In the FLUKA Monte Carlo code [26] the high energy interactions are simulated within the Dual Parton formalism [27]. In the present FLUKA95 version interactions at intermediate energies are treated by two distinct models: a code based on a model of resonance production and decay covers the region between 1.3 GeV and 5 GeV, while a sophisticated pre-equilibrium cascade code [24] treats the energy range below 1.3 GeV and also simulates the propagation of the intranuclear cascades. The evaporation module of FLUKA95 uses latest available nuclear mass and level density tabulations [28].

Fig. 5 shows energy deposition contours and neutron fluxes of hadron cascades in tungsten ($\lambda_f=9.6$ cm) for four different energies as simulated with FLUKA. The enhanced penetration of leading particles along the cascade axis becomes evident at the two highest energies. At $\lesssim 1$ GeV almost no buildup can be observed and the maximum energy deposition and neutron fluxes are found immediately below the surface where the particle entered. The table in Fig. 5 displays the charged hadron spectra incident on the CMS ECAL. A comparison with the associated figures shows that most of the neutron albedo is caused by pp-secondaries with an energy around 1 GeV. The tail of high energy particles in the LHC pp-secondary spectrum is responsible for calorimeter punchthrough and other deep penetration phenomena. But also these high energy cascades always terminate in a huge amount of low energy protons and neutrons. This emphasizes the fact that most radiation issues, even at the LHC, are to a large extent determined by physics phenomena at relatively low energies.
Particle transport is fundamentally a ray-tracing problem, which can at least in theory be solved analytically. In practice, however, analytical approaches are restricted to simple geometries and require simplifications of physical processes. In a true four dimensional multi-medium environment of arbitrarily complex structure the issue becomes complicated. Although paths can be calculated explicitly even in the presence of magnetic fields, a multitude of possible interactions usually precludes analytical methods. Essentially all particle transport problems in high energy physics are solved with the Monte Carlo method where the solution to the transport equation is obtained by sampling interactions from the appropriate probability distributions while stepping along the trajectory of each particle.

Interactions to be taken into account during transport include

1. point like inelastic interactions: nuclear scattering, capture reactions, annihilation, air production and the photoelectric effect,
2. point like elastic interactions: Coulomb and nuclear scattering, the Compton effect, bremsstrahlung and δ-electron production,
3. particle decay,
4. quasi-continuous interactions: multiple Coulomb scattering and ionization energy loss in materials and
5. continuous processes which are not associated with a collision: Čerenkov and synchrotron radiation.

All point-like events are characterized by a cross section $\sigma$. For atomic mass $A$ and density $\rho$ this corresponds to a mean free path $\lambda_I = A/NA\rho\sigma$ while the decay length $\lambda_D = \gamma\beta c\tau$ [4] depends on the life time $\tau$ and velocity $\beta c$ of the particle.

The ionization energy loss is due to small momentum transfers to atomic and free electrons and is described by the Bethe-Bloch formula [4]. It has a shallow minimum at $\beta\gamma \sim 3$ and increases logarithmically towards higher energies. Below $\beta\gamma \sim 2$ the energy loss increases rapidly with decreasing energy. This increase gives rise to the Bragg peak, i.e. a large energy deposition at the end of the range of a particle. Except in the presence of magnetic fields or when the particle is actually stopping, continuous energy loss affects the tracking only through the energy dependence of cross sections.

The Coulomb potential has infinite range and although the nucleus is screened in atoms, a charged particle still has a huge cross section for scattering to a very small angle. These scatterings are so frequent that they are usually treated as continuous processes. Formulae for the lateral displacements and the changes of direction cosines due to multiple scattering exist [4], but they are not straightforward to apply to displacements in thin objects or close to boundaries. Decreasing the step indefinitely is not only prevented by efficiency considerations but also by the fact that the Molière theory breaks down below a certain material-dependent step size [4]. FLUKA uses a sophisticated algorithm which guarantees exact multiple scattering even in thin objects with a magnetic field [29]. Stepping over a boundary is done with a few elementary scatterings in order to avoid flux-infinities, which can be the result of an uncontrolled change of direction cosines exactly on the surface.

In FLUKA the transport of the electromagnetic cascade is done with EMF [30], an improved version of the EGS4 shower code [31]. Depending on the $Z$ of the material electrons, positrons and photons can be reliably transported to energies as low as 1–10 keV.

One of the most significant improvements in FLUKA versions later than FLUKA87 is the inclusion of a low energy neutron transport module. Neutron transport in FLUKA is based on multigroup transport theory using a
72 group structure. The central idea behind multigroup transport is, that cross sections are defined as averages

\[
\bar{\sigma} = \int_{E_{\text{min}}}^{E_{\text{max}}} \sigma(E) \phi(E) dE,
\]

where \(\sigma(E)\) is the actual cross section at energy \(E\). The flux \(\phi\) is the weight associated with the cross section. In general \(\phi(E)\) is \textit{a priori} unknown. Usually experimental knowledge of average fluxes is used to derive the group cross sections. A downscattering matrix describes the transfer probabilities of neutrons from one group to another. The polar scattering angle is usually sampled from discrete values, obtained from Legendre expansions of the differential cross sections. Some neutron transport codes use point cross sections, in which case large tables of \(\sigma(E)\) are needed. A point cross section approach does not require any assumption for the flux in the problem to be simulated and is therefore in general more accurate and universal. It is, however, important to recognize that the point cross sections themselves are retrieved from experimental neutron irradiation data under some assumption of \(\phi(E)\) for that particular experiment. Since the deconvolution of a neutron spectrum from any kind of measured data is not trivial, the errors in the point cross sections are not necessarily smaller than those in the group cross sections.

4.3 Variance Reduction Methods

It is well known that Monte Carlo integration converges slowly: \(\sigma \sim 1/\sqrt{N}\), where \(N\) is the size of the sample [32]. Sometimes special variance reduction methods can be applied to improve the convergence in some regions of phase space on the cost of others. In FLUKA several such techniques are available as alternatives to a fully analog simulation. The most important ones are: region importance biasing, Russian roulette/splitting in hadronic interactions, leading particle biasing in electromagnetic interactions, non-analog absorption of low energy neutrons and decay length biasing [24].

Region importance biasing causes particles moving to a region with higher importance to be split with appropriately reduced weights. Particles moving into the opposite direction are killed with a certain probability and the weight of the surviving ones is increased correspondingly. This technique is extremely powerful for deep penetration problems where – ideally – the dependence between the number of histories and the shield thickness can be turned from exponential to linear.

Leading particle biasing and Russian roulette are related: one or several particles emerging from an interaction are discarded and the weight of the surviving ones is increased – usually by the ratio of surviving and discarded energy. Since the survival probability is related to the energy, these techniques also increase the average penetration and can allow an increase of the number of histories due to the multiplicity suppression.

Weight windows are often essential to avoid excessive weight fluctuations introduced by some of the biasing methods. Such fluctuations could result in an increase of the variance instead of the expected decrease. Since a biased simulation cannot provide any information about the actual error, this has to be estimated from several independent samples.

4.4 Fluxes, Currents and Doses

A technical detail, which may appear insignificant but nonetheless repeatedly turns out to cause confusion is connected with terminology and definitions. The quantity which is of central importance for most radiation calculations is flux density [33], often called just flux [34]. It can be defined either as tracklength per unit volume per unit time, or as the number of particles hitting a sphere with unit cross-sectional area per unit time. [4]. Both definitions are equivalent and show that the unit can be expressed as \(\text{cm}^{-2}\text{s}^{-1}\). Fluence is
the time integral of flux [4]. The definition of flux as flux density times unit of area, given in Ref. [33], is in contradiction to the common convention quoted above and is rarely seen in the literature.

The quantity usually confused with flux is current. Both have the same unit, but the latter is defined as the net number of particles passing through a unit surface per unit time. In a directed radiation field flux and current are identical only for normal incidence to the surface. At all other angles the flux is higher by a factor of \(1/\cos \theta\). It can be shown that in a semi-isotropic radiation field the flux is exactly twice the current.

A further note on the unit of flux is appropriate: Hz/cm\(^2\) is used frequently, but when rigorously interpreted could lead to a misunderstanding. Hz is defined in relation with periodically occurring events, whereas background fluxes at the LHC follow the Poisson distribution. So the common notation of Hz/cm\(^2\) in connection with fluxes around LHC experiments should not be interpreted to imply a periodicity.

Dose is often used as a synonym for fluence, although the two have no universal one-to-one relationship. Dose is defined as absorbed energy per unit of mass and is expressed in Gy. Dose rate is dose per unit of time.

The radiation risk to human beings is expressed as dose equivalent, which has units of Sv. Physically dose equivalent is a poorly defined quantity. It is the absorbed dose (in Gy) weighted by a quality factor (Q) which is adjusted to reproduce the damage caused in tissue by a given type of radiation. Q values are expressed as functions of the linear energy transfer in tissue (L). The numerical values of quality factors are constantly under discussion and undergo revisions from time to time [35]. Typically they range from 1 for \(L < 10\) keV/\(\mu\)m to a maximum of 30 at \(L = 100\) keV/\(\mu\)m [35].

## 5 Shielding at Hadron Accelerators

### 5.1 Neutrons

Neutrons, which are predominantly released by nuclear evaporation and during the intranuclear cascade phase, are rather special particles. No other neutral hadron – with the exception of the antineutron – has a macroscopic life time comparable to \(\tau = 887\) s of the neutron [4]. While slow charged particles, which ionize the medium, have only a short range in bulk matter, neutrons are attenuated only by nuclear scattering processes and can therefore travel substantial distances.

Such scatterings can result in significant energy transfer called kerma. For instance in \(np\)-scattering a hydrogen nucleus often receives momenta which are sufficient to break atomic bonds. Since the slow proton recoil is heavily ionizing, neutrons can cause substantial damage in tissue or induce signals in detectors. In crystal or metal lattices neutron scattering events can distort the atomic structure which can lead to degradation of the performance of semiconductor devices or to changes in the properties of structural materials – a well known problem for the pressure vessels of nuclear power plants.

Low-energy neutron scattering can to a good approximation be described by non-relativistic formulae. In elastic scattering from a nucleus with mass \(A\) the neutron emitted to an angle \(\theta\) with respect to the original direction has an energy

\[
E(\theta, m, A) = \frac{m^2 E_0}{(A + m)^2} \left( \cos \theta + \sqrt{\frac{A^2}{m^2} - \sin^2 \theta} \right)^2,
\]

where \(E_0\) is the energy of the original neutron and \(m\) is the neutron mass. Equation 2 attains a minimum at
\( \theta = -\pi \), in which case it simplifies to

\[
E = E_0 \left( \frac{A - m}{A + m} \right)^2,
\]

which vanishes, if \( A = m \). Since this condition is realized for scattering from protons, hydrogen is often the most effective element to slow down neutrons. Scattering cross sections and inelastic channels of course also need to be included in the considerations, and it turns out that at higher energies heavier materials can provide better neutron shielding than a hydrogen-rich plastic.

At energies of the order of 20 MeV and above nuclear structure effects become less important and cross sections start to be smooth resembling those of other high energy particles. At energies \( \gtrsim 1 \text{ MeV} \) materials of medium weight are usually ideal for neutron attenuation. They have short mean free paths for neutrons but do not introduce the problem of neutron multiplication in inelastic events. The best known case of such multiplication is certainly neutron induced nuclear fission, but especially \((n,2n)\) reaction channels exist also in non-fissile nuclei like lead. One excellent neutron shielding material is stainless steel. Pure iron is less effective, since in some energy regions it has very low neutron scattering cross section. Scatterings push the neutrons into these “windows”, from where they can be scattered out only by impurity atoms.

Essentially at all energies \( \lesssim 1 \text{ MeV} \) the best choice for neutron attenuation is hydrogen. Among the best neutron moderators are water \((\text{H}_2\text{O})\), paraffin and polyethylene \((\text{both } \sim \text{CH}_2)\). All of these combine one relatively light element with two hydrogen atoms and so provide the highest hydrogen densities of all commonly used materials. Often concrete is not a bad alternative, especially if a cheap general purpose shielding material is needed. It has a relatively large hydrogen content and due to its high density and heavier constituents is effective also in the high energy domain. Concrete can also be loaded with ferrites or heavier elements to obtain even better attenuation of high energy particles. Heavy dopants like barium have the disadvantage of producing long-lived radioisotopes in the high energy radiation field around hadron accelerators.

When a neutron has experienced numerous elastic scatterings, it will have slowed down to a limit, where its average momentum transfer in all subsequent collisions is zero. Most neutron spectra are dominated by these thermal neutrons, which at room temperature have an energy of \( kT=25 \text{meV} \). By definition thermal neutrons do not initiate any high energy recoils. The neutron capture cross sections of most isotopes, however, rise towards lower energies and can reach values of several kb, i.e. more than three orders of magnitude larger than typical high energy hadron-nucleus cross sections. Therefore most neutrons are slowed down to thermal energies before being absorbed by a nucleus.

Neutron capture reactions may constitute a problem since the new isotope is usually formed in an excited state. Although the neutron energy may have been less than 1 eV the resulting nuclear excitation can amount to several MeV. This excess energy is released in form of capture gammas, which then contribute locally to the electromagnetic dose rate. The capture gamma energy in hydrogen, for instance, is 2.2 MeV corresponding to the deuteron binding energy. Other materials, like cadmium, can emit several photons for each captured neutron. Therefore, in configurations with efficient photon shielding, neutrons can increase the effective penetration of gamma dose rates. The number of energetic photons can be reduced by doping the neutron moderator with lithium-7 or boron-10. Both have high capture cross sections but very low gamma energies. Lithium in fact is one of the few isotopes which do not produce a capture gamma at all: the formed Be-8 is unstable with respect to two helium nuclei. With an average of 2.5 kb cadmium has – with the exception of the rare gadolinium – the highest thermal neutron capture cross section of all natural elements. Although it produces numerous capture gammas and is highly toxic it is one of the most common thermal neutron shielding materials. To reduce the photon flux, cadmium layers can be surrounded with lead or other efficient photon shielding material.
5.2 Photons

Shielding a pure photon field is relatively simple. Since photon interaction cross sections increase rapidly with the atomic charge of the material, the usual choice is to use the material with highest Z available. In practice this often means lead.

If the photon field is accompanied by other types of radiation – which always is the case around hadron accelerators in operation – some caution has to be taken in the choice of the material. Although lead has excellent shielding properties for photons, every lead atom is loaded with neutrons. During the intranuclear and evaporation phases, which follow after a high energy hadronic interaction, several of these can be released. Since pure lead is also very transparent to low energy neutrons a lead layer in an environment dominated by high energy hadronic radiation can lead to an increase of the total radiation background.

5.3 Charged Particles

From the shielding point of view charged particles can be divided into three groups: electrons, muons and hadrons.

Electrons and positrons are governed by the same physical laws as photons. At high energies they together constitute the electromagnetic cascade. Therefore an efficient photon shield always provides also good shielding against incident electrons.
Muons are unstable but have a half life sufficient for a range of several kilometres. As the heavy partners of the electrons, muons are subject to the same electromagnetic interactions but due to the larger mass radiative energy losses of muons are suppressed and become important only when the energy exceeds ~100 GeV. Since muons, in addition, are not affected by the strong nuclear interaction in the same way as hadrons, they experience very small attenuation in matter. This facilitates the operation of the muon detection systems of high energy physics experiments, but simultaneously introduces a possible radiation hazard. Being charged particles muons are slowed down by ionization energy loss, which is 1.7 MeV cm\(^2\)/g for a minimum ionizing particle in concrete and 1.45 MeV cm\(^2\)/g in iron. Thus muons with energies below few GeV can be suppressed with shielding, but a TeV-muon can traverse several kilometres of rock. Since these high energy muons cannot be shielded, once they are produced, all one can do is to deflect them with a magnetic field or to go out of their way. This indeed is possible since these muons are emitted in the direction of the original projectile with an extremely small sideways spread [36]. At the LHC possible hazards due to high energy muons which are produced by proton losses along the ring are minimized by positioning occupied areas on the inside of the ring.

The only method to actually suppress the high energy muon background is to prevent them from being produced. Most muons in the accelerator environment are formed in meson decays. Therefore the number of muons can be reduced by suppressing these decays. Since the typical decay length of a charged pion is tens of metres, it is often possible to force pions to interact before they have had a change to decay.

Pions and other charged hadrons experience both the strong and the electromagnetic force. Therefore they lose energy by ionization or can undergo inelastic nuclear scattering. In the latter case, depending on the energy, a number of secondary particles with reduced energy are formed. All these particles form the hadronic cascade, which is almost completely attenuated after 10–12 hadronic interaction lengths (\(\lambda_T\)) even at LHC energies. Around the iron mass a typical value for one \(\lambda_T\) is ~15 cm, dropping to 9.6 cm for tungsten [4].

The general recipe for an accelerator environment is to reduce the flux of high energy particles with dense but medium \(A\) material, like iron, placed close to the source. Then to slow down neutrons with hydrogenated materials and to capture them either in the hydrogen or with boron or cadmium. Finally the photons can be suppressed with a layer of lead.

Fig. 6 illustrates the radiation field around a shielded 1 metre long and 5 cm thick iron target which is hit by one 20 GeV proton. The relative neutron and photon attenuation efficiencies of iron, concrete, polyethylene and lead are clearly visible. It is however to be emphasized that Fig. 6 is very schematic since it displays a single particle and does not include any spectral information, which always is an important quantity associated with particle flux. The relative doses, given in Fig. 6, are estimated from 2000 primaries for each configuration and provide a more quantitative measure of the shielding efficiencies.

\section{Hadron Fluxes in the Tracking Cavity and around the ECAL}

The tracking cavity of CMS is limited by the ECAL to a radius of 1.3 m and a total length of 6 m. The opening in the endcap ECAL has a radius of \(r=30\) cm, corresponding to \(\eta=3\). Further downstream, at \(z \sim 10\) m the very forward calorimeter (VFCAL), intercepts particles in the \(\eta\)-range 3–5.

For radiation calculations it is usually a sufficiently good approximation to group detector planes at representative radii and to replace actual support structures with homogenized material distributed over the whole tracker volume [PubI, PubV, PubVI, 19].

Three fundamentally different sources of hadron background in the central cavity are illustrated in Fig. 7 and discussed below.
Figure 7: Schematic view of hadron tracks in the CMS inner tracker. a) fast or neutral particles, b) slow charged \(pp\)-secondaries c) slow charged particle produced by an inelastic interaction in the tracker material, d) albedo neutrons emerging from interactions in the calorimeters.

Figure 8: Fluxes of charged hadrons and neutrons at different radii in the CMS inner tracker. Solid arrows show the averages over \(z\) and dotted arrows show what \(1/r^2\) would predict when normalized at \(r=20\) cm [PubVI].

According to Fig. 2 \(dn/d\eta\) is almost constant in the region covered by the tracker. In the absence of a magnetic field this implies that the flux \(\phi \sim 1/r^2\), where \(r\) is the transverse distance from the beam line. In a uniform magnetic field along the \(z\)-axis the maximum radius which a charged particle can reach depends on its transverse momentum \(p_T\):

\[
r[mm] = \frac{p_T}{0.3B} \frac{[\text{MeV/c}]}{[\text{T}]},
\]

which means that the shape of the \(p_T\) distribution is reflected in the radial dependence of fluxes. As can be seen from the fluxes shown in Fig. 8, the \(1/r^2\)-dependence is not exactly valid in the 4 T magnetic field of CMS but the uniformity along \(z\) is still obeyed.

Although the tracker represents less than one radiation length and only a fraction of an interaction length some of the \(pp\)-secondaries can initiate inelastic interactions in the tracker material. In these cases several new particles with lower momentum are formed. Like decay products, but unlike low \(p_T\) \(pp\)-secondaries, these particles are uniformly distributed in the cavity. In particular they tend to loop close to their point of production, repeatedly penetrating the plane where the inelastic interaction occurred. These particles are responsible for the low energy flux observed at outer layers [3], but their contribution to the total background is relatively small [19].

The third source of background, which dominates at large radii, is calorimeter albedo. The \(pp\)-secondaries generate interactions in the calorimeters. While the fast particles emerging from these interactions are predominantly directed away from the cavity, the evaporated neutrons are emitted isotropically and have typical energies around 1 MeV. As will be seen, the damage function of low energy neutrons in silicon drops by an order of magnitude below 100 keV.

It has been shown in Refs. [PubI, PubV] that the material of the ECAL can affect the > 100 keV neutron flux
by an order of magnitude. Calorimeters which include high $A$ elements are shown to produce the most intense albedo. Any intrinsic hydrogen content – provided by the plastic scintillators of a “shaslik” ECAL, for instance – reduces significantly the neutron albedo. If intrinsic hydrogen is not present in the calorimeter itself, layers of hydrogenated material in front of the calorimeters can substantially reduce the amount of albedo neutrons and soften their energy spectrum. In particular, for the PbWO$_4$ calorimeter an efficient moderation of the albedo has been shown to be essential [PubV].

The fluxes shown in Fig. 8 are for the Technical Proposal configuration of CMS, including a PbWO$_4$ crystal ECAL equipped with a lead preshower and polyethylene moderators [3]. With this shielding the charged hadron flux dominates over the neutron flux at radii $\lesssim 70$ cm.

Close to the ECAL, neutrons are responsible for most of the damage in silicon devices. Some protection for the preshower detector can be obtained by placing moderator layers between the preshower and the ECAL [PubV]. On the back of the endcap ECAL, at $\eta > 2$, radiation-hard vacuum devices have to be used since neutron fluences can reach $10^{14}$ cm$^{-2}$ for $\int E dt = 10^6$ pb$^{-1}$ [PubVI]. Readout electronics behind the ECAL is exposed to similar neutron fluxes but a reduction by a factor of 3–5 can be obtained with a polyethylene layer of 5–10 cm thickness between the ECAL back plane and the electronics boards [PubVI].

7 Displacement Damage in Silicon Detectors

Displacement damage, often also referred to as bulk damage, is a distortion of the crystal lattice of a semiconductor. Since the CMS central tracker will utilize silicon pixel and microstrip detectors, their degradation due to displacement damage is one of the most crucial issues connected with the LHC radiation environment.

During the doping procedure lattice distortions are introduced in a controlled way in order to reduce the bandgap of the pure semiconductor and to enhance the number of thermally activated carriers. Effects caused by the displacements created during irradiation are in some respects similar to those obtained by doping but the defect formation mechanisms and defect types are uncontrolled. Usually radiation induced defects are detrimental to the detector performance.

Radiation exposure can lead to three types of changes in the characteristics of a semiconductor detector:

1. Crystal defects introduce new allowed states into the bandgap which increase the number of thermally activated charge carriers and lead to an increase of the bulk generation current, which usually forms the dominant part of the leakage current.

2. The same crystal defects also act as trapping and recombination centers which decrease the lifetime of minority carriers and so decrease the charge collection efficiency.

3. Some defects can neutralize the donors of the substrate, while others behave like acceptor-type dopants. Thus irradiation can first turn the silicon into effectively intrinsic and subsequently invert the originally n-type substrate into p-type.

Several empirical parametrizations have been presented for the degradation of silicon devices as a function of particle fluence $\phi$. The increase of the leakage current, $I$, and the change of the effective doping concentration $N_{\text{eff}}$ are often expressed as

$$I(\phi) = I_0 + \alpha \phi$$

and

$$N_{\text{eff}}(\phi) = N_0 e^{-\alpha \phi} - \beta \phi - N_{\text{ao}},$$

19
respectively. $I_0$ and $N_0$ are the pre-irradiation values and $N_{a0}$ is the initial effective acceptor concentration. An intuitive physical interpretation for the parameters $e$ and $\beta$ has been provided in Ref. [37]. Some of the introduced defects compensate the original donors by forming neutral complexes with them. This results in an exponential decrease of the number of active donors. Other defects, for instance divacancies which are formed by combination of two neutral vacancies, can behave like acceptors. Such acceptor-like defects are introduced at a rate proportional to the fluence. Although Eq. 6 fits well to experimental data, the correctness of the interpretation quoted above has not been established at the microscopic level [38].

Annealing effects are important, but often complicated. It is for instance well established that the leakage current annealing can be described by a sum of exponentials with different time constants [39]. However, it is not yet clear if one of these constants is infinite, or if the silicon ultimately would reach its pre-irradiation characteristics. Leakage current annealing is faster at elevated temperatures and is found to be essentially stopped at $-40^\circ\text{C}$ [PubV].

Significant annealing effects are also observed for the changes of $N_{\text{eff}}$, but these are even more complicated. Immediately after irradiation $N_{\text{eff}}$ starts to anneal towards its pre-irradiation value corresponding to n-type silicon. But after about two weeks at room temperature this annealing process is overwhelmed by another one, known as reverse- or anti-annealing. Apparently new acceptor like impurities are formed in the silicon during this process so that the substrate starts to evolve towards p-type. It has been found, however, that this reverse annealing can be frozen at temperatures around $0^\circ\text{C}$, where the beneficial annealing still takes place. Thus there is an operating “window” for silicon detectors around $0^\circ\text{C}$ [40]. At lower temperatures the beneficial annealing of both $N_{\text{eff}}$ and the leakage current is stopped and above $\sim5^\circ\text{C}$ the reverse annealing gets activated.

The voltage required to fully deplete a detector of thickness $d$ depends on $N_{\text{eff}}$, the electron charge $e$ and the dielectric constant of silicon $\varepsilon$:

$$V_{\text{dep}} = \frac{e d |N_{\text{eff}}|}{2\varepsilon}.$$  \hspace{1cm} (7)

It can be seen from Eqs. 6 and 7 that during irradiation the depletion voltage first tends towards zero but then – after the type inversion – increases linearly with a slope $\beta$. At the same time the leakage current also increase with a slope $\alpha$. If the detectors are all the time operated at full depletion, the power consumption – i.e. heat load – is proportional to $\phi^2$.

The leakage current in a silicon detector depends on the temperature $T$ according to

$$I(T) \sim T^{3/2} \exp \left( -\frac{E_g}{2kT} \right),$$  \hspace{1cm} (8)

where the bandgap in silicon at $20^\circ\text{C}$ is $E_g = 1.121 \text{ eV}$. Thus an increase of power consumption would heat up the silicon, which would increase the leakage current leading to ever higher heat load – a phenomenon known as thermal runaway [41]. Taking into account that elevated temperatures accelerate reverse annealing processes some consequences of such thermal runaway are not reversible.

### 7.1 Non-ionizing Energy Loss

For a given type of radiation the number of crystal defects depends on the number of inelastic interactions, which has a linear relationship to the particle flux $\phi$. From Eq. 5 it then follows that also the leakage current must be linearly related to the number of crystal defects. It is, however, not obvious at all if the constant of proportionality is universal or if it depends on the distribution of the defects. Experimentally we know for instance that the value of $\alpha$ in Eq. 5 is quite different for electron and proton irradiations and depends also
on the energy. The question is if this difference is only due to a certainly different total number of defects, or also to their distribution.

Upon scattering momentum is transferred to the target atom. The recoil energy \( E_r \) is divided into non-ionizing energy loss (NIEL) and into ionization of the medium. The NIEL resulting from a single scattering process is function of recoil type and \( E_r \) and can be expressed formally as \( \delta(A, Z, E_r) \). If \( E_r \) is low \( \delta \) can amount to almost the total recoil energy. For high \( E_r \) and a recoil around the silicon mass \( \delta \) approaches a value of 300–400 keV. This asymptotic value decreases with decreasing atomic number of the recoil ion [42, 43].

Coulomb scattering has a very large cross section but most of the events are characterized by low momentum transfers. When an electron traverses the silicon lattice it will transfer momentum to many silicon atoms – but always small amounts. This energy can very well be in excess of \( \sim 20 \) eV needed to dislodge the silicon atoms from their lattice sites [44], but in most cases the recoils will not get very far from their original position and only rarely will be able to displace any further atoms. Nuclear scattering of hadron projectiles, on the other hand, often results in momentum transfers of several MeV. A single recoil atom with a few MeV kinetic energy can initiate an atomic cascade in the lattice. Thus hadronic interactions produce clusters of substantial lattice distortion, which cannot be produced by electrons or muons. Prior to any reordering the number of displacements is estimated to be the NIEL divided by 50 eV [45]. Most of the crystal defects heal within very short time scale, but some are at least metastable.

Although no definitive answer has yet been found if correlations between defects are important, most experimental findings seem to support the NIEL hypothesis, which states that the observed damage should be directly proportional to the number of displacements, i.e. the NIEL.

Basically this statement means that the bulk damage after exposure to any kind of radiation can be determined by calculating the total energy loss in silicon and subtracting the ionization loss. For charged hadrons the latter always dominates and it has been shown in Ref. [PubVI] that the average NIEL for charged hadrons traversing 300 \( \mu \)m of silicon is \( \lesssim 0.1\% \) of the total energy loss. This emphasizes the fact that bulk damage in semiconductors is proportional to particle fluence but not necessarily to the radiation dose.

### 7.2 Calculation of Displacement Damage

The preceding discussion already indicated that the NIEL hypothesis forms a suitable basis for the calculation of displacement damage in silicon.

Since sufficiently accurate analytical models exist for the calculation of stopping powers of atomic cascades, the NIEL is a quantity which can be estimated without sophisticated solid state physics simulations [46, 47]. If experimental data are available to establish the conversion constant between the NIEL and observed damage, NIEL calculations can be used to estimate the damage for any kind of irradiation.

Since the atomic cascade is initiated by the primary recoil, the key quantities are energy and type of the recoiling atoms together with their production cross sections.

For an incoming particle \( X \) of energy \( E \) the average total NIEL (denoted here as \( \Delta \)) in silicon \((A_0 = 28, Z_0 = 14)\) can be formally expressed as

\[
\Delta(X, E) = \delta(A_0, Z_0, E_r)\sigma_{el}(X, E_r) + \sum_{i=1}^{A_0} \sum_{j=1}^{Z_0} \delta(i, j, E_r)\sigma_{in}(X, E_i, j, E_r). \tag{9}
\]

The elastic cross section \( \sigma_{el} \) to produce a recoil of energy \( E_r \) is relatively easy to calculate [2]. The energy transfer is fixed by the scattering angle, the mass of target and projectile and the collision energy. In a silicon
lattice the recoil is always a silicon atom. The differential cross sections can be obtained either from experimental data or theoretical models. At sufficiently high energies a rather suitable model is provided by the Glauber formalism [48]. Since the Glauber model is based on the impulse approximation, it is unable to take nuclear effects properly into account and consequently is not valid at very low energies.

The inelastic cross sections $\sigma_{in}(X, E, i, j, E_r)$, needed in Eq. 9 are significantly more problematic. Due to the huge number of partial cross sections, each corresponding to a different final state, an analytical calculation is impossible. Nucleons and other particles, produced in an inelastic event, usually do not undergo another interaction in the thin silicon detectors and contribute a negligible amount to the NIEL. So the first sum in Eq. 9 does not necessarily have to be started at $A=1$. From the target, however, one residual or sometimes several fragments are left with a few MeV of kinetic energy.

Typical recoil energies are so low, that the Fermi-energy of nucleons inside the target nucleus cannot be neglected. In fact most of the recoil energy is due to the differences in the Fermi motion of the nucleons in the original and the residual nucleus. By momentum and energy conservation the recoil energy is in principle known. Unfortunately data on exclusive inelastic events is scarce. Due to the difficulties to interface high energy physics and nuclear structure models, there are significant uncertainties in calculations which use simulated recoil spectra. The pre-equilibrium model [24] of FLUKA95 is in principle able to provide recoil spectra, but given the lack of experimental data their correctness is difficult to ascertain [49].

A semi-theoretical method of estimating pion induced displacement damage in silicon has been developed in Refs. [PubII, 2]. The method is based on the fact that elastic recoil energies, which due to the mass difference are lower for pions than for other hadrons, can be calculated analytically and only the second term in Eq. 9 remains unknown. To circumvent the problem of estimating inelastic recoil energies on a microscopic basis, experimental proton damage data has been used. This corresponds to hiding all partial cross sections in Eq. 9 by defining a function

$$\Omega(X, E) = \frac{1}{\sigma_T(X, E)} \sum_{i=1}^{A_0} \sum_{j=1}^{Z_0} \delta(i, j, E_r) \sigma_{in}(X, E, i, j, E_r)$$

(10)

where $\sigma_T(X, E)$ is the total inelastic cross section for particle $X$ with energy $E$ to interact with silicon.

By subtracting from the observed total proton damage the part which can be explained by elastic proton scattering, the contribution which has to come from all inelastic proton-silicon events, whatever they are, can be deduced. This is exactly $\Omega(p, E)\sigma_T(p, E)$.

A relation between recoil energies in $\pi$–Si and $p$–Si inelastic events has to be postulated. Such a postulate, however, is likely to be much less in error than a direct estimation of the recoil spectra. One reasonable assumption is, that protons and pions with the same momentum produce identical recoil spectra. The NIEL due to pions can then be estimated by

$$\Delta(\pi, E) = \delta(A_0, Z_0, E_r)\sigma_{el}(\pi, E, E_r) + \sigma_T(\pi, E)\Omega(p, E'),$$

(11)

where $E'$ is the proton energy which corresponds to the same momentum as a pion energy $E$. This formula is equivalent to the one presented in Ref. [PubII] if $\delta(A, Z, E_r)$ of the actual recoil is assumed to be equal to that of a silicon ion with equal energy and the recoil multiplicity per interaction is assumed to be 1. Since this method is sensitive to the assumption of the relation between proton and pion recoil spectra, it should be understood that it cannot provide more than a rough estimate.

Fig. 9 shows the result of such an estimation. The pronounced $\Delta$-resonance in the $\pi$-Si cross section is reflected in the damage curve. At high energies the pion damage drops to 2/3 of the proton damage – corresponding to the ratio of inelastic cross sections. Fig. 10 shows the particle spectra at the r=20 cm tracker layer
Figure 9: NIEL for neutrons, protons and pions in silicon. The pion curve is based on the calculation in Ref. [PubII]. Values are relative to 1 MeV neutron NIEL.

Figure 10: Kinetic energy spectrum of neutrons, protons and pions at the r=20 cm tracker layer of CMS. Averaging is performed over ±1 m in z.

7.3 Measurement of Pion Induced Displacement Damage

The predicted high damage of pions around 200 MeV, combined with the high abundance of such pions at the inner tracker layers of LHC detectors, motivate an experimental verification of the damage prediction.

Only a few facilities in the world can provide sufficiently intense low energy pion beams for such an experiment. One of these is at the Paul Scherrer Institute (PSI), where a series of pion irradiations were performed during a five-week period in summer 1994. These tests, performed by three independent groups in the high intensity \( \pi E1 \) beamline, provided the first determinations of pion damage at low energies. The first results are reported in Refs. [PubIII, 50]. A detailed description of the experimental methods can be found in Ref. [PubIII] and only a short summary is presented in the following.

Several diodes of \( 5.5 \times 5.5 \text{ mm}^2 \) area and 300 \( \mu \text{m} \) thickness were irradiated at \( \pi^+ \) momenta of 200, 300 and 400 MeV/c, which cover the interesting \( \Delta \)-resonance region. An extension significantly beyond this energy range is problematic, since the beam intensity drops rapidly when moving away from the resonance.

At each energy individual diodes were irradiated to different fluences. During irradiation the diodes were kept in an insulated box which was flushed with cold nitrogen to maintain the temperature below \( -40^\circ \text{ C} \). No annealing effects could be detected, which would not have been frozen at this temperature. Therefore a correction for annealing during irradiation was not needed. Leakage current versus voltage (IV) curves [51] of the diodes were taken a few minutes after removal from the box and were repeated at several later instants of time for each diode. Capacitance versus voltage (CV) curves [51] could be measured only after the diodes had spent some 20 minutes at room temperature, therefore some annealing had already taken place. Correction
for these effects was practically impossible, so care was taken to have a constant \( \Delta t \) between removal from the box and the first CV-measurement.

The pion fluence was monitored with pieces of aluminium foil, which had the same size as the diodes. One such foil was positioned on the back of each diode. The cross section for the reaction \( ^{27}\text{Al}(\pi^+, X)^{24}\text{Na} \) is reasonably well known [52]. Because the the half life of \(^{24}\text{Na} \) is only 15 h the beam had to be constantly monitored in order to correct for intensity variations during the longest irradiations which lasted up to 35 h.

In Ref. [PubIII] all measured data for the leakage current and the change of effective doping concentration as a function of pion fluence are given together with results showing the annealing behaviour of these two quantities during the first 50 days at room temperature. Figs. 11 and 12 show observed leakage current and \( [N_{\text{eff}}] \) values for 300 MeV/c pions just after irradiation and after 50 days of annealing. It must be emphasized that \( [N_{\text{eff}}] \) already has bypassed its minimum value and by coincidence the \( \beta \)-value at 50 days happens to be almost equal to \( \beta \) just after irradiation [PubIII].

Since the increase of leakage current is theoretically better understood than the changes of depletion voltage, a comparison between \( \alpha \) and the predicted NIEL of pions is particularly interesting. The NIEL obtained from the prediction has to be related to some proper damage value. The “predicted” values in Table 5 are obtained assuming \( \alpha = 8 \times 10^{-17} \text{ A/cm} \) for 1 MeV neutrons for which the NIEL is \( \sim 95 \text{ MeV mb} \) [53].

Although the shape of the predicted curve is relatively well reproduced in the experiments the predicted maximum of the pion \( \alpha \) is about 20% lower than the experimental value from Ref.[50] and 50% lower than the value reported in Ref. [PubIII]. This underestimation is not very surprising. Since pion absorption is a dominating process around the \( \Delta \)-resonance, the assumption that pions and protons would produce identical recoil spectra in inelastic interactions probably underestimates the pion NIEL. Upon absorption the pion deposits all of its rest mass in the nucleus. Although this energy is predominantly released through emission of two fast nucleons, the average recoil energies can still be expected to be higher than in non-absorptive events, where the projectile maintains some of its forward momentum. Sometimes pion absorption can also lead to
Table 5: Predicted and experimental values of $\alpha$ (in $10^{-17}$ A/cm) for pions. The predicted values are obtained assuming $\alpha = 8 \times 10^{-17}$ A/cm, for 1 MeV neutrons. Normalization is performed with respect to the 300 MeV/c values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>5.9 0.79</td>
<td>10.4 $\pm$ 0.7 0.85</td>
<td>7.8 $\pm$ 0.71 0.8</td>
</tr>
<tr>
<td>300</td>
<td>7.5 1.0</td>
<td>12.2 $\pm$ 0.5 1.0</td>
<td>9.7 0.7 1.0</td>
</tr>
<tr>
<td>350</td>
<td>7.4 0.99</td>
<td>--</td>
<td>8.8 $\pm$ 0.2 0.91</td>
</tr>
<tr>
<td>400</td>
<td>6.8 0.91</td>
<td>10.7 $\pm$ 0.7 0.88</td>
<td>10.5 1.0 1.1</td>
</tr>
<tr>
<td>450</td>
<td>6.1 0.81</td>
<td>--</td>
<td>8.7 $\pm$ 0.8 0.90</td>
</tr>
</tbody>
</table>

1 Interpolated from 150 MeV/c and 250 MeV/c values.

fragmentation of the target nucleus, in which case there can be several recoil fragments, which increase the NIEL over one single recoil. Although apparently of minor importance, these effects together may add up to the observed discrepancy.

It is more difficult to explain why the two experiments observe a different discrepancy. The reason might be in the experimental methods: the diodes of Ref. [50] were irradiated at room temperature and include an annealing correction, which was not needed in Ref. [PubIII]. More likely, however, is that the response of the devices is different. It is well known that different experiments find a large variation for the the 1 MeV neutron $\alpha$ [54] and the chosen value of $8 \times 10^{-17}$ A/cm is not necessarily correct for an individual device. Recently performed irradiations indeed indicate that the neutron $\alpha$ of the devices used in Ref. [PubIII] is about $10 \times 10^{-17}$ A/cm [55]. Since $\alpha \sim 8 \times 10^{-17}$ A/cm is claimed for the devices used in Ref. [50], the experimental pion/neutron damage ratios of both experiments seem to be in good agreement.

According to Figs 9 and 10 the NIEL varies by a factor of $\sim 2$ around the “standard” value for 1 MeV neutrons over most of the energy and particle range in the LHC spectrum. But the errors of the NIEL curves can still be significant. Additional uncertainties arise from the device type to be finally used, the annealing effects in the actual LHC environment and errors in the flux estimation. In view of these it appears to be an unjustified complication to fold the hadron spectra with the damage curves. A more appropriate procedure is to “define” average damage constants for the total flux in the LHC tracker region and to apply these to the fluxes given in Fig. 8.

Such an average for the annealed leakage current constant at 0$^\circ$ is of the order of $0.6 \times 10^{-17}$ A/cm. The effective doping concentration after inversion increases with a slope of $\sim 1.4 \times 10^{-12}$ V/cm. The breakdown limit of most detectors is $\sim 200$ V, which is reached at a fluence of $\sim 1.5 \times 10^{14}$ cm$^{-2}$ s$^{-1}$. At this fluence the leakage current of a 300 $\mu$m thick detector is $27$ $\mu$A/cm$^2$ leading to a heat load of $5.4$ mW/cm$^2$. This heat is uniformly produced in the whole bulk of the detector and if there would be no heat losses the temperature of the silicon would increase by $\sim 6^\circ$C per minute. In order to prevent thermal runaway, very sophisticated cooling systems are needed [3].

8 Background in the CMS Muon System

Since muons are likely to provide the only reasonably clean signals of new physics at the LHC the reliability and efficiency of the muon system is of central importance for the performance of the CMS experiment. The critical rate limit, posed by the saturation of the Resistive Plate Chambers (RPC) and by the occupancy of the Cathod Strip Chambers (CSC), is of the order of 1 kHz/cm$^2$ [56, 57].
Figure 13: Signal generation in a multilayer muon chamber by various particles and illustration of the shielding strategy. Stars indicate observable signals. a) A fast charged particle penetrates the whole chamber, b) a photon stops in the iron/lead, c) a slow charged particle stops in the iron, d) a neutron slows down and stops in the polyethylene and the capture gamma is stopped in the lead and e) a neutron penetrates the chamber, generates one recoil on the way and stops in the iron. One capture gamma returns to the chamber and converts to an $e^+e^-$-pair.

The innermost muon chambers, especially MF1, have to cope with substantial hadronic punchthrough which adds to the muon signal rate. If not properly shielded, the signal rate due to neutrons and photons can reach comparable levels.

According to Ch. 5 part of the neutron kerma goes into ionization which can be observed as a signal in a gas chamber. According to Eq. 2 protons are the most important recoils since heavier ions usually do not receive sufficient energies to overcome the signal threshold of a chamber. Although capture gammas have to initiate a further reaction before a visible signal is obtained, they are usually more important than recoil ions.

Since the produced electrons or recoil ions have in general very low energies their effect is only local, often almost point-like. In a multilayer detector like the muon stations of CMS, they only rarely give a signal in more than a few of the sensitive volumes before being stopped in the wall separating two detector layers from each other. Fig. 13 illustrates the signal generation mechanisms in the CMS muon stations and indicates some shielding alternatives.

### 8.1 Sensitivity Factors

Since the probability of obtaining a signal from a photon or a neutron is very low the number of energy deposition events is small. It is not feasible to score these events directly in a simulation. Significantly better statistics can be obtained by scoring the neutral fluxes and weighting them with a sensitivity factor, defined as the probability per unit flux to obtain a signal. Since these factors depend on the incident spectrum, the chamber type, and its surroundings, experimental sensitivities are useful only if the experimental conditions really agree with the assumptions of the simulation. In general this is not the case. A dedicated simulation to obtain the sensitivity factors is confronted with uncertainties in the signal generation mechanism like amplification in the gas and the signal threshold of the chamber. The most reliable way is to perform an irradiation experiment with a typical chamber and a roughly LHC-like neutron spectrum. The signal thresholds in
a simulation model can be tuned with the obtained data by simulating exactly the experimental conditions. These tuned parameters can then be used to obtain sensitivities of the actual CMS chambers exposed to the simulated LHC spectrum.

Unfortunately such irradiation tests of the CMS chamber prototypes have not yet been performed. Thus simulations of the chamber sensitivity have been based on rough assumptions of signal thresholds [PubV]. This introduces the dominant uncertainty into the final signal rate estimates for neutral particles.

Typical neutron sensitivities of RPC and CSC detectors, surrounded by different materials, were found to be $1 - 3 \times 10^{-4}$ excluding capture gamma production. The photon sensitivities of $1 - 2 \times 10^{-2}$ are therefore dominating. Most of the photons, however, are due to neutron capture and if this contribution is included in the neutron sensitivity a value of $2 - 6 \times 10^{-3}$ is obtained [PubV]. Since an experimental distinction between capture gammas and recoil ions is practically impossible, this is also the sensitivity which a neutron irradiation experiment should observe. There is an obvious danger of double counting: when both neutron and photon fluxes are scored the capture gammas are already included in the photon flux and the neutron sensitivity factor should not include them. If on the other hand experimental neutron sensitivities are used, the corresponding gamma flux – i.e. capture gammas subtracted – is in most cases only a fraction of the total. From the sharp energy cutoff in the photon spectrum of Fig. 15 it can be concluded that essentially all photons are due to neutron captures. In such a case the best estimate for the photon sensitivity can be obtained by using the neutrons as incident particles but normalizing with the observed photon flux [PubV, PubVI].

Figure 14: Fluxes of neutrons, photons and charged particles in the MF1(AB) and MF4 muon stations of CMS.

Figure 15: Energy spectra of neutrons, photons and charged particles in the MF1 and MF4 stations of CMS.
8.2 Background Suppression

The collimator, which according to Table 4 is hit by an average energy of 2.3 TeV per event, is the principal source of background radiation in the experimental hall. All of the incident energy is not contained in the copper. The produced hadronic cascade has an extent of several metres and dissipates energy into surrounding shielding materials.

The absence of the shielding would result in an immense particle flux in the whole experimental hall and neutron and gamma fluxes in the outermost muon stations would reach 4 and 1 MHz/cm², respectively [PubV]. This would correspond to a signal rate of 10–20 kHz/cm², but the direct charged particle contribution, which has a sensitivity factor of 1, would be of equal importance.

It has been shown in Refs. [PubV, 3] that proper shielding can reduce the photon flux below 10 kHz/cm², corresponding to a signal rate of < 200 Hz/cm². The charged particle fluxes can give a slightly larger contribution, ranging locally up to rates of ~1 kHz/cm² [3].

The choice of shielding materials is dictated by the physics of radiation shielding as discussed in Ch. 5 and constraints posed by engineering, cost and available space. Compared to the other high luminosity LHC experiment, ATLAS [58], the muon system of CMS profits from the shielding provided by the massive iron yoke. The only directly exposed chamber is MF4. Therefore CMS is less sensitive to particle fluxes in the experimental hall. This allows a relatively thin shielding, which is easier to support and remove during detector maintenance. A significant amount of high energy particles are able to penetrate this thin shielding but these are mainly directed away from the detector towards the cavern walls. Since the spectrum scattered back from the walls is soft, a relatively thin shielding wall immediately in front of MF4 provides sufficient flux reduction [PubV, PubVI, 3]

8.3 Beam Pipe Geometry

It has been shown [PubIV, PubV] that a beam pipe consisting of one or several cones can significantly reduce the background in the muon system. A standard beam pipe would have a constant radius of the order of 6 cm and aluminium walls of 1.8 mm thickness. Neglecting the magnetic field and multiple scattering effects, a particle hitting the pipe at a distance of 8 m from the interaction point arrives at an inclination of only 7.5 mrad and thus could traverse 24 cm in the aluminium. Since \( \lambda_f = 40 \) cm in aluminium, half of these particles would interact in the pipe material.

Cones pointing to the vertex, are – ideally – not hit by any particles. In practice the cones cannot be made exactly pointing, since the vertex has a finite spread. So there are some particles which hit the cone walls at extremely small angles. These almost certainly interact in the aluminium. But the solid angle contributing this high interaction probability is very small and thus a significant overall gain is obtained by a conical pipe. It has been shown that the magnetic field does not significantly change this result [PubIV]. Based on the studies presented in Refs. [PubIV, 59] the CMS collaboration has proposed a conical beam pipe in the region between the central beryllium section and the VFCAL.

In Ref. [59] it was also shown that a massive vacuum pump inside of CMS would increase the background in the muon chambers. The conical beam pipe has the additional advantage of providing better conductivity so that the pump can be moved away from the IP and aligned with the front face of the VFCAL. In this position it has negligible effect on the background rates [PubV].
9 Dose Rates in the Detectors

The radiation dose absorbed during LHC operation can influence the performance of some detectors, destroy non radiation-hard electronics or degrade the properties of some organic materials.

With the aid of the Bethe-Bloch formula[4] charged particle fluence can be converted into ionizing energy loss. It was indicated earlier that non-ionizing effects give a negligible contribution to the total energy deposition of charged particles. For neutrons this is not true: recoil ions can dissipate a significant portion of their energy in non-ionizing processes. In non-hydrogeneous media the non-ionizing and ionizing energy losses of neutrons can be of the same order of magnitude [60]. But since the total neutron energy loss is always only a fraction of the ionization loss of charged particles, the energy deposition scored with FLUKA is a good approximation of the energy loss due to ionization.

Since ionizing radiation creates trapped surface charges in the surface layers of semiconductor components, standard front-end electronics is very sensitive to absorbed dose. There is a very distinct difference to bulk damage, which affects the detectors: bulk damage is caused predominantly by hadrons, whereas absorbed dose gets contributions also from the electromagnetic part of cascades, including neutron capture photons and even photons from decay of induced activity. Radiation hard electronics has been shown to survive doses of up to 1 MGy, which according to Fig. 16 is more than enough for LHC experiments. But such devices are significantly more expensive than conventional ones.

In scintillators the light yield and light transmission usually decrease as a function of absorbed dose. The degradation is not expected to be very critical for crystal scintillators, like the CMS ECAL [61]. Plastic scintillators, however, are not likely to maintain the required performance at doses, significantly in excess of 10 kGy [62]. Fig. 17 shows that the dose in the plastic scintillators of the HCAL reaches almost 10 kGy per 10^8 pb^-1. Therefore replacements are likely to be necessary in order to reach the LHC lifetime of \( \int L dt = 5 \times 10^8 \ \text{pb}^{-1} \). Even though the light yield and transmission itself may remain within acceptable limits, another issue is that any changes must be well under control in order to maintain the calibration and thus the energy resolution of a calorimeter. Especially problematic is that radiation doses in the calorimeters are not uniform. Even within one scintillator unit differences can be significant.

Fig. 17 illustrates and interesting feature of sampling calorimeters. According to the Bethe-Bloch formula the dose in plastic should be about 50% higher than in copper [4]. The difference found in the simulations amounts to a factor of 3–4, which can be explained by energy deposition of neutrons. This effect forms the basis of compensating calorimeters, i.e. devices with equal response to electromagnetic and hadronic cascades [63, 64]. Usually the response to the EM shower is larger, so that the energy resolution of the calorimeter is influenced by fluctuations of the \( \pi^0 \) content in hadronic cascades. If the two responses are equal, i.e. \( e/h=1 \), these fluctuations are not reflected in the resolution. This can be achieved either by decreasing the electron signal or by amplifying the hadron signal. Since a significant amount of the energy in a hadronic cascade is carried by neutrons the latter can be achieved by converting these neutrons to visible energy. Exactly this takes place in a plastic scintillator, where neutrons produce large amounts of hydrogen recoils.

Radiation damage of organic materials dictates the selection of support structures and layout of cables. The tracker support frame, which has to maintain its rigidity despite the high doses, is based on carbon fibres embedded in epoxy resins. In high rate irradiation tests such composites have not shown any sign of degradation at doses well in excess of 10 years of LHC operation. It is, however, not excluded that low rate irradiations could give different results. This is especially true if oxygen, which is known to play an important role in the damage mechanism, is present during the irradiation [65]. The diffusion of oxygen into the material is characterized by long time constants, so that the combined effects of oxygen absorption and simultaneous radiation exposure may not become visible in short high rate irradiations.
Figure 16: Radiation dose at five representative radii of the CMS inner tracker. The solid arrows indicate the averages over $z$. The dotted arrows show what $1/r^2$ would predict when normalized at $r=20\,\text{cm}$ [PubVI].

Figure 17: Radiation doses at some selected radii in the CMS endcap calorimeters. The spikes are due to energy deposition by neutrons in the plastic scintillator layers of the HCAL. Some histograms terminate due to the conical shape of the calorimeters [PubVI].

It is known that most conventional cable insulations decompose at relatively moderate doses of 0.1–1 MGy and are therefore not suitable for use in central parts of LHC detectors. Insulators with verified good radiation hardness up to $\sim 10$ MGy are the only ones to be allowed in the most exposed areas [65].

The extreme radiation environment around the experiments is emphasized by considering the energy absorbed by the collimator. The copper mass of $\sim 2000\,\text{kg}$ absorbs $\sim 200\,\text{W}$ of power. Without any heat losses this would be sufficient to heat up the whole mass with a speed of $\sim 1$ degree per hour. Since the collimator is fairly well embedded in shielding, heat leakage is likely to be small. Special cooling might become essential in order to avoid large temperature variations, which could influence the alignment of the quadrupoles within the same shielding structure.

10 Shielding of Occupied Areas during LHC Operation

Without any shielding around the collimators the dose equivalent in the experimental hall, when the machine is operating at peak luminosity, would reach $\sim 1\,\text{Sv/h}$ [66]. The shielding required by the muon system of CMS reduces this to $\sim 2\,\text{mSv/h}$ [16]. Such an area is accessible under severe restrictions on access time, but it is far above the CERN limit of $25\,\mu\text{Sv/h}$ for radiation controlled areas, i.e. zones which can be constantly occupied or accessed as part of the daily routine. The construction material of the experimental hall is standard concrete in which the typical value for dose attenuation length at hadron accelerators is 50 cm. In order to obtain the needed reduction by further two orders of magnitude the wall separating the experimental hall from any occupied area has to be at least 2 metres thick.

Different methods to estimate dose equivalent at hadron accelerators are proposed. These are compared in Fig. 18 and discussed below.

The first method consists of scoring total energy deposition and multiplying it with an average quality factor (Q). The uncertainty in this method arises from the Q-value to be used. An often quoted recommendation
is $Q=5$, which has also been found in ATLAS simulations[67] but several other studies indicate that $Q=10$ should be closer to reality at the LHC [PubVI, 16, 66, 68].

The second method is very empirical. It is based on average properties of hadronic cascades and just states that the dose in Sv is the number of stars per cubic centimetre in concrete multiplied by a factor of $4.5 \times 10^{-8}$. The concept of star density frequently appears in the estimation of radiological quantities at hadron accelerators [69]. A star is defined to be an inelastic interaction of a hadron which has more than 50 MeV of kinetic energy. The numerical values of conversion factors like the one quoted above have to be determined from comparisons between simulations and experiments. Therefore these factors include all contributions to the dose, but depend both on the experimental methods and the physics contained in the simulation code. For the latter reason they are strictly valid only if the code used for the predictive simulation is known to include the same processes and cross section as the one which was used to determine the conversion factors.

The third method, which in principle is the most accurate one, consists of folding the particle fluxes with conversion factors, which depend on the type and energy of the incident particle. These factors are usually calculated with dedicated transport codes so that their correctness depends on the quality of the calculations and on the geometrical conditions assumed. Another uncertainty here arises from the fact that recoil protons and capture gammas are usually implicitly included in the neutron fluxes and have to be separated from other protons and photons. This uncertainty, however, is usually small since the total proton and photon contribution to the dose in concrete is of the order of 10%, only [66].

## 11 Induced Radioactivity

It is claimed that $\sim 50\%$ of the residual nuclei produced in hadronic interactions in materials heavier than iron have radiologically significant half lives between 10 minutes and 10 years [70]. On average these emit 1.5 photons with a mean energy of 800 keV but some daughter nuclei may still be radioactive and emit further photons. Altogether this yields a rough rule of thumb: each hadronic interaction leads to one induced activity photon of 800 keV energy.

This rule for the average production rate is accompanied by a similarly average formula for the time dependence of dose rate. It has been found [71] that the dose from an object exposed to hadron radiation can be expressed as a function of irradiation and cooling times ($t_i$ and $t_c$):

$$D(t_i, t_c) \sim \ln \left( \frac{t_i + t_c}{t_c} \right). \quad (12)$$

Certainly this formula cannot be valid in cases like aluminium activation, where only a few isotopes contribute [72]. Like the production rule quoted above also this formula works only when the number of different isotopes produced is so large that no single one is dominating.

Neglecting activation by low energy neutrons, a radioisotope can be formed only as a result of a hadronic interaction. In full analogy with the previously described method to estimate run-time dose rates from the number of inelastic interactions it has been proposed to estimate induced activity doses from simulated star densities [73]. The conversion factor between star density and induced activity dose rate is referred to as $\omega$-factor and it has units of (Sv h$^{-1}$)/(star cm$^{-3}$ s$^{-1}$). By comparing experiments and simulations the $\omega$-factors can in principle be made very accurate. However, the same comments apply as for the star density to run-time dose conversion: the $\omega$-factors depend on the simulation code used but also on the radiation field. For instance, if the $\omega$-factors are determined in an environment dominated by high energy hadrons and the configuration to be studied is exposed to a significant flux of low energy neutrons then the contribution from thermal neutron activation is not taken properly into account. And it is well known that for some materi-
Figure 18: Attenuation of the dose equivalent in the side wall of the CMS cavern at the point of maximum incident dose. Values obtained from the three dose estimators are shown together with the fit of the analytical formula [16].

\[ H = 3000 \times \exp\left(-\frac{d}{\lambda}\right)/(d+1200)^2 \]

\( \lambda = 55 \text{ cm} \)
\( \lambda = 50 \text{ cm} \)
\( \lambda = 45 \text{ cm} \)

Figure 19: Induced activity dose rates at different positions with respect to a CMS tracker wheel and in contact with some detector modules of the CMS tracker. In all plots only \(^7\)Be, \(^{22}\)Na and \(^{198}\)Au are considered [PubVI].

als, like cobalt, it can dominate over high energy activation. Thus the \( \omega \)-factors are valid only if both the simulation codes and the physical configurations are compatible with those used for their determination.

The rigorous method of estimating induced activity dose rates proceeds over particle fluxes, which are multiplied with partial isotope production cross sections to obtain explicit production rates for each isotope. Using the extensive tabulations of radionuclide decay data [74] it is straightforward to calculate the time dependence of the dose rates from explicit radioisotope inventories. For some elements, like aluminium, where sufficient coverage of experimental cross sections is provided [75] such a procedure is feasible. In general, however, the problem is that partial cross sections for the needed radiation type, energy range and material are not available.

Whether explicit or average radionuclides are used as input, the estimation of dose rates in real geometries requires that the attenuation in the active medium itself or any surrounding shields is taken into account. The general formula for the dose rate due to photons of energy \( E \) (MeV) at finite distance from an active object of arbitrary shape can be written as a sum [72]

\[
\frac{dD}{dt} \left[ \frac{Sv}{h} \right] = \sum_j \frac{10^{-8} EA_j}{7r_j^2} \prod_i \exp\left(-\frac{x_i}{\lambda_i}\right),
\]

(13)
where $A_j$ is the number of photons emitted per second by a small volume element $j$ and $r_j$ (cm) is the distance to that element. The product runs over material layers of thickness $x_i$ traversed by the photon and $\lambda_i$ is the corresponding attenuation length.

Often buildup is not negligible and has to be included as a correction to the exponential attenuation term [34]. Such a correction is relatively straightforward in homogeneous media, but becomes complicated in the presence of multilayered shields [76]. Buildup effects also depend on the geometrical situation. If, for instance, radiation is incident at a very small angle on a slab of shielding material buildup effects may cause the sideways scattered contribution to dominate over the direct penetration, which is suppressed by the exponential term. In addition to such effects the ray-technique represented by Eq. 13 does not take into account backscattered radiation and therefore tends to underestimate doses behind corners. If, however, essentially all material is activated – which is the usual case around hadron accelerators – the dose rate to first approximation depends only on the activity close to the material surfaces. Then buildup appears only as a small correction and Eq. 13 becomes relatively accurate.

Equation 13 has been used to estimate the dose rate in the CMS hall due to activation of bulk objects [PubVI]. These calculations are based on star densities in all materials and corresponding $\omega$-factors, which certainly can be in error by a factor of 2–3 [72]. For all regions outside the CMS shielding the dose rate estimates remain well below 1 $\mu$Sv/h. Thus no access restrictions to the experimental hall will be needed after the beam is stopped. The issue of maintenance is quite different. According to Ref. [PubVI] dose rates close to the forward ECAL can reach ~500 $\mu$Sv/h, at the corner of the VFCAL ~10 mSv/h and in contact with the collimator front face ~30 mSv/h. The maximum dose rate – as expected – is found in the center of the collimator: ~0.5 Sv/h. These values are for 60 days of irradiation and 1 day of cooling. For an infinite irradiation and zero cooling time dose rates are estimated to be about a factor of three larger [69].

Certain low density objects and voids like air are better treated by scoring tracklength (or flux $\phi$) and using explicit isotope production cross sections. The activity concentration as a function of irradiation and cooling time ($t_i$ and $t_c$) can be expressed as

$$A(t_c) = \rho \frac{N_A}{M} \sum_j \sigma_j \phi \left[1 - e^{-\lambda_j t_i}\right] e^{-\lambda_j t_c},$$

(14)

where $\sigma_j$ is the production cross section of isotope $j$, which has a decay constant $\lambda_j$. $N_A$ is the Avogadro number, $M$ the molecular mass and $\rho$ the density of the sample. Equation 14 does not take into account the possibility of radioactive daughter nuclei.

If all materials have low densities it is often a sufficient approximation to neglect the attenuation term in Eq. 13. The tracker, which sits close to the vertex in the highest flux, is the most interesting object for such a study. It is also very suitable for such treatment, since it consists mainly of carbon, aluminium and silicon for which isotope production cross sections are available. The MSGC detectors might be equipped with gold strips, which require special considerations, due to the large thermal neutron activation cross section of gold. Calculations of the induced activity in the tracker modules as a function of time are described in Ref. [PubVI]. Fig. 19 shows the obtained doses at different positions with respect to a tracker wheel and in contact with single modules. The assumed LHC operation schedule is reflected in the dose rate, where the short-lived gold follows the fine-structure, $^7$Be is visible as a small slowly decaying contribution and the $^{22}$Na gradually approaches its saturation value.

It has to be emphasized that since all short-lived nuclei are neglected, the curves in Fig. 19 represent only a small fraction of the induced activity dose rate during LHC operation. Such a simplification is justified since any access to the tracker requires opening of the CMS endcap ensuring that nobody will be able to get close to the modules in less than ~1 week after machine shutdown. By that time essentially only the $^{22}$Na activity is left and even at saturation does not reach values which would be critical for maintenance.
As a conclusion, CMS is freely accessible during machine shutdowns, but certain parts – namely the endcap ECAL, the VFCAL and especially all beamline elements between the IP and Q1 – are medium or highly radioactive objects. Special shielding has to be provided around these objects when the run-time shielding of CMS is removed during maintenance periods. The other alternative is to remove the items requiring maintenance from the radioactive environment. For instance the tracker should not reside inside of the activated ECAL during time consuming inventions.

Air activation is an issue which is still under study [77]. Also here tracklength scoring is used and a total of 39 radioisotopes are considered. Most of the radionuclides produced in air are either radiologically relatively harmless, are produced with low cross sections or have very short half lifes. It is, however, possible that some air volumes, especially the collimator surroundings, require special precautions in order to prevent radioactive air leakages into occupied areas. The time required to remove the shields which block access ways during machine operation provides sufficient decay time for the short-lived nitrogen and oxygen isotopes. Thereafter the radioactivity in the air of the experimental hall is negligible and no access restrictions are needed.

12 Summary

The radiation shielding requirements and strategies at the CMS experiment have been reviewed. It has been shown that the unprecedentedly hostile radiation environment provided by the LHC has to be taken into account in the whole design of the detector. Hadronic simulation codes like FLUKA have in the last five years undergone significant development and the new physics models introduced have made them complete and accurate enough to face the LHC requirements [26].

The radiation background at CMS is entirely dominated by the $\nu\nu$-interactions. Beam losses in the machine, even if these take place close to the experiment, are negligible. Except for high energy muons this background is also easily reduced by sealing the interface to the machine tunnel with concrete.

Silicon detectors at $r=20\,\text{cm}$ survive $\int L \, dt = 2 - 3 \times 10^6\,\text{pb}^{-1}$ before the depletion voltage reaches the breakdown limit. This is only half of what is needed to complete the full LHC physics program. These considerations, however, are based on the assumption that high resistivity silicon of $300\,\mu\text{m}$ thickness is used. It has been shown that type inversion and the subsequent depletion voltage increase can be delayed by using low resistivity n-type material [78]. Since, according to Eq.7, the depletion voltage depends on the square of the thickness of the active layer, breakdown can be avoided by using thinner detectors or operating them underdepleted. All these alternatives sacrifice some of the signal to noise ratio of the detectors, but this might be the price to pay for a sufficient lifetime. Latest developments focus on possibilities called defect engineering: if certain impurities are introduced into the silicon substrate, they may be able to neutralize the defects caused by irradiation without affecting the properties of the silicon themselves. Although the silicon irradiation studies launched by the LHC project have given promises that a sufficient lifetime can be reached, they also have clearly proven that this is neither obvious nor simple.

The ECAL, which is the major source of neutrons inside of the detector, causes significant radiation damage to its own readout photodiodes and front-end electronics. Neutron fluxes integrated over the LHC lifetime can reach $5 \times 10^{14}$ [PubVI] which is twice as high as the hadron flux at the $r=20\,\text{cm}$ tracker layer. With moderators these fluxes can be lowered by a factor of 3–5 [PubVI], but even then they may constitute a problem which has not yet been studied sufficiently.

According to Fig. 17 the scintillators of the most forward part of the HCAL absorb a dose of $\lesssim 40\,\text{kGy}$ over the LHC lifetime. In this – fortunately small – region scintillators most probably have to be replaced every few years [3].
The quartz-fibre VFCAL of CMS is extremely radiation hard, but the problem is induced activity. Depending on the irradiation and cooling times, contact dose rates of 10–30 mSv/h are expected at the front edge of the VFCAL. These values are at the limit of an object which still is accessible and certainly access times have to be limited to a few minutes.

While the barrel muon system can be considered to be safe, the forward muon chambers are likely to operate at the limit of their capabilities. In fact none of the chambers could work without the massive shielding [PubV]. The thick steel insertion in the forward cone of CMS is essential to suppress background in MF1–3. In addition to the shielding around the collimator the safe operation of MF4 requires a dedicated shield wall just in front of the chambers. Despite the significant reductions obtained with the shielding, background rates in MF1 are still very close to the acceptable maximum. In fact estimated systematic errors do not exclude the possibility of having rates in excess of the critical 1 kHz/cm² limit. If rates indeed should turn out to be too high, then there would be no space for additional shielding and MF1 would have to be redesigned to be able to cope with higher rates.

Although radiation safety is recognized to be important, it is for the LHC experiments a smaller issue than detector performance. There are areas close to the collimator, which might not be accessible at all due to induced activity. Limited access times are likely to be imposed for maintenance of the VFCAL and the forward ECAL. Certainly this is the case also for the machine elements and the vacuum equipment close to the experiment. However, the experimental hall and most parts of the experiment itself are accessible without any special restrictions.

It is some kind of a paradox that the LHC, while being the highest energy accelerator ever, is the first machine which required hadronic simulation codes to evolve towards low energy nuclear physics models and ultimately to incorporate these. But this is just a reflection of the artifact that radiation physics is due to the bulk of the hadronic cascades which even at the LHC is dominated by low energy particles. Until the LHC pushed the intensity of these cascades to the extreme, they never were really interesting for high energy physicists. For radiation protection purposes rough approximate formulae, like those introduced in Chapters. 10 and 11, have been used with success for almost three decades to obtain order-of-magnitude estimates of radiation exposure.

The LHC is the first machine where physicist and radiation protection personnel are equally concerned: in order to avoid excessive and expensive overshielding the radiation exposure of personnel has to be calculated more accurately than before. But the most critical points are found in detector performance and life time, where safety factors are difficult to obtain. Even after shielding the feasibilities of most detector systems to be used in the LHC radiation environment depend on factors of the order of two.

References


[38] G. Hall, private communication.


[74] National Nuclear Data Center Data Service NNDC/ONL-92/8, ENSDF database (periodically revised).
Radiation Environment Simulations for the CMS Detector

Mika Huhtinen

Abstract

This paper provides an overview of the most important radiation issues at the CMS experiment at LHC and gives a review of the simulations performed and the main results obtained with the FLUKA code during the last three years. The main emphasis is put on recent results, reported in the CMS Technical Proposal or those obtained thereafter. It is shown that for most of the CMS subdetectors proper radiation shielding is vital in order to guarantee a reliable operation and sufficient life span of the detectors. In addition new results on induced activity are presented.

Presented at the
Second Workshop on Simulating Accelerator Radiation Environments,
CERN, Geneva, Switzerland, 9–11 October 1995

CERN, Geneva, Switzerland
14th December 1995
Abstract

This paper provides an overview of the most important radiation issues at the CMS experiment at LHC and gives a review of the simulations performed and the main results obtained with the FLUKA code [1, 2] during the last three years. The main emphasis is put on recent results, reported in the CMS Technical Proposal [3] or those obtained thereafter. It is shown that for most of the CMS subdetectors proper radiation shielding is vital in order to guarantee a reliable operation and sufficient life span of the detectors. In addition new results on induced activity are presented.

1 INTRODUCTION

Due to its high energy, but in particular because of its extremely high luminosity LHC will provide a radiation environment with which large scale particle physics experiments have never before been confronted.

Compared to lepton collisions hadron-hadron interactions are particularly 'dirty', generating a huge amount of minimum bias background, which is not very exciting from the physics point of view, but keeps radiation protection personnel occupied.

Hadronic interactions lead to activation of materials and give rise to a neutron background which requires very special shielding precautions.

But at LHC safety aspects are only one part of the whole story: it has become obvious that the hadronic minimum bias background is so massive that it starts to influence the detectors. It may damage semiconductor devices to such a degree that they would not survive the whole LHC program. Or it may lead to saturation of detectors or occupancies beyond tolerable levels. It has been shown that in some cases background reduction by orders of magnitude is required in order to safeguard reliable detector operation [4].

Mostly driven by the requirements imposed by LHC and the former SSC FLUKA, among some other radiation transport codes, has evolved during the last five years into a multipurpose simulation tool [2]. A particularly important improvements were the addition of a pre-equilibrium cascade model [5] and low energy neutron transport and an accurate multiple scattering formalism [6] to the traditionally first-class high energy event generation modules of FLUKA. Complemented with EMF (ElectroMagnetic Fluka [7]), an improved version of EGS4 [8], recent versions of FLUKA should contain all the best models for radiation simulations at high energy hadron accelerators. An important – and unique – feature of FLUKA is the possibility to use it as a fully analog code, in which case all correlations are preserved, or to use any of the powerful variance reduction techniques which are essential for deep penetration studies.

Benchmark experiments [9, 10] have shown that FLUKA is accurate within some tens of percent for dose and fluence estimation. Thus it is certainly among the most reliable simulation tools for studying radiation issues at LHC.

Independent simulation studies for CMS have been performed also with the MARS simulation code [11, 12, 13] but this paper will be restricted to FLUKA results.
2 LHC MACHINE PARAMETERS

The three LHC parameters of significance for the radiation background at experiments are the proton energy of 7 TeV, the day-averaged collision rate, estimated to be $3.5 \times 10^{30}$ s$^{-1}$ and the maximum collision rate, which is roughly twice the average value. The two latter are directly related to the peak and average luminosities ($10^{34}$ cm$^{-2}$ s$^{-1}$ and $5 \times 10^{33}$ cm$^{-2}$ s$^{-1}$, respectively) and the inelastic cross section at 7 TeV, which is expected to be 70 mb – diffraction excluded. The average luminosity depends, not only on the peak value but also, on the number of fills per day and the initial number of protons per fill.

To obtain year-averages for radiation studies the recommended[14] assumption is that during the first year LHC would reach 1/10 of the design luminosity and in the two following years 1/3 and 2/3, respectively. From the fourth year onwards LHC would operate at full luminosity. According to Fig. 1 the annual operation time, is $1.5 \times 10^7$ s. The low luminosity lead-run adds only a negligible amount to the radiation exposure. Integrated luminosity, which is the best measure for the time required to obtain a physics discovery, is usually expressed in terms of pb$^{-1}$. The standard LHC physics program is based on $5 \times 10^5$ pb$^{-1}$, which corresponds to $5 \times 10^7$ s at peak luminosity. With the quoted numbers and schedules this is equivalent to roughly 9 years of LHC operation including the low luminosity startup phase.

Obviously peak rates are important for determining maximum occupancies and background signal rates. Radiation damage and activation, however, have to be determined from long-term averages.

The proton bunch spacing is 25 ns. Since lifetimes of hadronic cascades are much longer the irradiation appears continuous as far as background is concerned.

3 THE CMS DETECTOR

The Compact Muon Solenoid (CMS) [3] is one of the two general purpose high luminosity experiments proposed for the LHC. It is characterized by a large solenoidal magnet, which provides a uniform 4 T field in the tracking and central calorimetry regions. Most of the muon spectrometer is embedded in the massive iron return yoke of the magnet.

The central tracking comprises three parts: two layers of pixel detectors at radii of 7.7 and 11.7 cm are followed by three layers of silicon strip detectors starting at $r=20$ cm and seven layers of microstrip gas chambers starting at $r=50$ cm. The strip detectors will be organized in a novel spiral design which allows easy routing of services. Along the beam axis ($z$) the tracker will be divided into 25 units called wheels, each having $\Delta z=25$ cm and $r=130$ cm. The support structure will be provided by a carbon fibre space frame.

The PbWO$_4$ crystal electromagnetic calorimeter (ECAL) provides excellent energy resolution which facilitates the detection of the potentially important $H \rightarrow \gamma \gamma$ channel. Since PbWO$_4$ crystals have a radiation length of only 8.9 mm they provide an extremely compact calorimeter. The ECAL will cover hermetically the pseudorapidity region $|\eta| < 3$.

The hadronic calorimeter (HCAL) comprises two parts. In the central rapidity region up to $|\eta|=3$ a copper/scintillator sampling calorimeter will be used. The forward region up to $|\eta|=5$ will be covered by the very forward calorimeter (VFCAL), which will be based either on parallel plate chambers (PPC) or quartz fibres (QF) embedded in steel or copper absorber.

The muon spectrometer which extends up to $|\eta|=2.4$ consists of four stations each having several independent position sensitive layers. In the barrel part the first station (MS1) comes immediately after the coil and provides an accurate stand alone muon momentum determination by measuring the exit angle from the field. It also provides the needed information for matching the muon track with the central tracker. The three following muon stations are interleaved with the iron of the return yoke and provide another independent determination of the muon momentum by measuring the bending radius in the saturated (1.8 T) iron.

Trigger information will be provided by resistive plate chambers (RPC) at central rapidities up to $|\eta|=2.1$. Cathode strip chambers (CSC) provide the spatial information in the endcap region while drift tubes will be used in the barrel muon stations, where particle rates are lower and the magnetic field is more uniform.

4 CMS EXPERIMENTAL AREA

The low-beta insertions allow a free space of 23 m between the end of the first quadrupole (Q1) and the interaction point (IP). At peak luminosity the total power carried by the secondaries from pp-interactions to either side of the experiment amounts to $\sim 800$ W. Most of this energy is deposited in the very forward region and could lead to quenches of the superconducting quadrupoles. In order to prevent this a copper collimator with a tentative length of 2 m, an outer radius of 20 cm and an aperture of 1.5 cm will be placed between the IP and Q1. The incident energy on the collimator is on average 2.3 TeV per event, from which it absorbs a power of $\sim 200$ W About 4.3 TeV continue through the collimator and are absorbed somewhere along the beam line. Only some 0.4 TeV out of the 7 TeV per event are absorbed in the CMS detector itself. Since particle generation, to a first approximation, depends linearly on the energy deposition, the collimator forms by far the dominant source of radiation background in the experimental cavern.

Due to the massive iron yoke and the solid VFCAL of CMS the radiation issues are divided into two almost independent regions: the interior of the detector including all of the tracking and the calorimetry system and the outside,
which comprises all equipment in the hall and the outermost layers of the muon spectrometer.

The radiation background inside the detector is determined by the multiplicities and pT-spectra of the primary events together with the materials present in the tracking and calorimetry regions. The number of particles able to penetrate the material of the calorimetry system and the iron yoke and having sufficient transverse momentum to escape the magnetic field is negligible. Thus the radiation environment in the experimental cavern is determined by the energy dissipated in the collimator and other objects close to the beam line.

5 SIMULATION METHODS

Primary events have been generated with the DTUJET93 event generator [16]. Single diffractive events are of negligible importance for the radiation background in the experimental area and have been neglected. The obtained DTUJET event file was randomized in order to obtain a smoother source for sampling. The resulting particles have been tracked with the FLUKA94/95 shower code [1] and normalization of results is according to the average DTUJET event multiplicity of 133.

The simulations have been done in several phases, and biasing methods, energy cutoffs and geometries have been selected accordingly. For instance, to estimate values in the tracker region the external parts (e.g. collimator) are not significant. In general, neutrons have been tracked to thermal energies and charged hadrons to rest. Photon cuts have been at 100 keV.

Magnetic fields were present in the detector (radially symmetric approximation) and quadrupoles.

6 PARTICLE FLUXES AND RADIATION DOSES IN THE TRACKING CAVITY

One central radiation issue concerning all LHC experiments is the damage induced in semiconductor detectors. The innermost pixel and strip detector layers of the CMS tracker will most probably be based on well established silicon technology.

The crucial question is if these silicon detectors can withstand the hadron fluxes in the vicinity of the LHC vertex for a sufficiently long time [17]. It has been well known that neutrons damage the silicon lattice by creating dislocations which lead to an increase of leakage current and depletion voltage and to a decrease of charge collection efficiency. Similar consequences have been observed for proton irradiation [18, 19, 20, 21]. Quite recently it was pointed out that pions, which are the most abundant particles close to the IP, induce at least the same amount of damage as neutrons and protons [22, 23, 24, 25].

Simulated charged hadron and neutron fluxes are shown in Fig. 2 for different radii in the tracking region. It can be shown analytically that in the absence of a magnetic field and scattering material the flux obtained from a flat pseudo-rapidity distribution is uniform along z and decrease as 1/r^2.

Despite the 4 T field the charged fluxes have maintained the uniformity along z but the 1/r^2 dependence is violated as indicated by the arrows in Fig. 2.

According to Fig. 3 the damage induced in silicon depends not only on the particle type, but also on its energy. In particular, neutrons with kinetic energy less than 100 keV do not damage silicon, whereas the damage function for protons rises towards low energies and the pion damage is expected to have a local maximum in the Λ-resonance region.

While charged particle fluxes are almost entirely determined by the primary events, neutron fluxes are mainly the result of calorimeter albedo. Therefore they are relatively
Table 1: Reduction of neutron flux in the tracking cavity as a function of forward (horizontal) and barrel (vertical) polyethylene moderator thickness in front of the PbWO<sub>4</sub> ECAL. Thicknesses are given in cm [27].

![Image](96x656)

Table 1

<table>
<thead>
<tr>
<th>Thickness cm</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction</td>
<td>1.0</td>
<td>0.64</td>
<td>0.52</td>
<td>0.39</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>in 100 keV</td>
<td>0.83</td>
<td>0.50</td>
<td>0.43</td>
<td>0.32</td>
<td>0.30</td>
<td>0.28</td>
</tr>
<tr>
<td>Flux reduction factor</td>
<td>0.73</td>
<td>0.48</td>
<td>0.39</td>
<td>0.31</td>
<td>0.31</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Figure 4: Radiation dose at five representative radii. The solid arrows indicate the averages over <em>r</em>. The dotted arrows show what 1/<em>r</em> would predict when normalized at <em>r</em>=20 cm.

Figure 4: Particle kinetic energy spectra in different regions of the CMS tracker. Dashed line: <em>r</em>=7.7–11.7 cm, solid line: <em>r</em>=20–41 cm, dotted line: <em>r</em>=80 cm. Two pseudorapidity ranges, 0–1 and 1–2, are considered. The solid and dotted neutron curves have been multiplied by 0.1 and 0.01, respectively in order to separate them from the dashed one.

uniform in the whole tracking cavity. It was pointed out in Ref. [23] that the absolute level of the neutron flux and the energy spectrum depend strongly on the calorimeter material. Hydrogenated neutron moderators in front of the calorimeter were shown to reduce the E<sub>＞</sub>100 keV fluxes significantly. An optimization of moderator thicknesses for the PbWO<sub>4</sub> calorimeter was performed in Ref. [27] and the obtained reduction factors in the middle of the tracking cavity are reproduced in Table 1. It can be seen that already a few centimetres of polyethylene reduce the fluxes by a factor of 3–4 and, in fact, a thickness increase beyond few centime-

eres does not provide significant further reduction. Detectors close to the forward ECAL, however, are more sensitive to the thickness of the forward moderator and their protection requires a thicker layer than indicated in Table 1. The presented results are for 4 cm PE in the barrel and 8.5 cm in the endcap regions. Both moderators are split in half in order to encapsulate the preshower [3].

Some position dependence of the particle spectra can be seen in Fig. 4. Due to primary neutrons and hyperon decays the neutron spectrum is hardest in forward directions and close to the vertex. When interpreting the charged hadron curves it must be remembered that all particles are included. Although pions dominate, protons and kaons introduce some smoothening and downward biasing effect.

Fig. 4 is based on DTUJET93, similar DTUJET92 spectra can be found in the CMS Technical Proposal (CMS-TP) Ref. [3] separately for π<sup>+</sup>, K<sup>+</sup> and p.π.

Except for minimum ionizing particles the relation between fluence and radiation dose is a relatively complicated one, involving a folding of the energy spectrum with the Bethe-Bloch formula. The FLUKA code can be used directly to score energy deposition in the tracker layers. The resulting dose rates are shown in Fig. 5. The deviation from 1/<em>r</em><sup>2</sup> is slightly larger than for fluxes. This is due to an interplay between the <em>pT</em>-spectrum, the 4 T field and energy dependence of dE/dx [28].

Ionizing energy deposition can create surface damage in semiconductors and thus mainly affects readout electronics. The lifetime of the detectors themselves is limited by bulk damage effects. These are – at least to a good approximation – linearly dependent on the non-ionizing energy loss (NIEL). Within a factor of two typical NIEL values are 100 MeV mb. Thus a charged hadron with a cross section of ~ 500 mb in silicon, deposits roughly 200 keV NIEL per collision giving an average NIEL of ~0.15 keV per 300 μm. A comparison with the mean ionization loss of 116 keV shows that radiation dose is not a proper measure...
7 NEUTRON FLUXES INSIDE AND AROUND THE CALORIMETERS

The calorimeters form the principal neutron sources inside CMS. This is especially true for the PbWO$_4$ ECAL, with both lead and tungsten being loaded with neutrons and oxygen as the lightest constituent not providing much intrinsic moderation.

The main issue is the radiation damage induced in semiconductor devices. The preshower with silicon pad detectors will be close to the front face of the ECAL. Readout photodiodes will sit on the back of the crystals. In addition, a significant amount of readout electronics will be positioned in the gap between the ECAL and the HCAL. Behind the ECAL radiation doses are significantly smaller than at the tracker, but may still require the use of radiation hard electronics. Neutron fluxes at the ECAL surface can reach very high values, which might severely restrict the lifetime of semiconductor devices if proper shielding precautions are not taken.

Fig. 6 shows the neutron fluxes in the most critical region, i.e. in front and behind the ECAL endcap. While the thickness of the front moderator is determined from the optimization studies of Ref. [27], the moderator layers between the ECAL and the HCAL are tentative and serve only as a pre-study of the effect of polyethylene in this region. The results suggest that 7 cm PE behind the ECAL reduce the neutron flux by a factor of four. The resulting flux at 50 cm is then comparable to the flux at 90 cm if no moderators were introduced. Thus it might be sufficient to use moderators only in the lowest part of the endcap ECAL, although some isolation layer towards higher radii might then be needed. It is also quite evident from Fig. 6 that the Cu/scintillator HCAL does not contribute significantly to the neutron flux in the ECAL–HCAL gap at the most critical small radii.

The effectiveness of polyethylene is nicely illustrated by the left plot of Fig. 7 where the neutron spectrum inside the calorimeter is shown. It clearly reflects the typical "spallation" and evaporation peaks around 70 MeV and 1 MeV, respectively$^1$. Since there is very little intrinsic moderation in the PbWO$_4$ calorimeter the spectrum has only a negligible tail to low energies. Behind the ECAL and after 7 cm of polyethylene the spectrum is completely different: the > 10 MeV neutrons have relatively small hydrogen scattering cross section and are not significantly affected by the polyethylene but the huge evaporation peak is almost completely suppressed. Most of the neutrons are either captured or appear at lower energies where they are not critical for silicon.

The right hand plot of Fig. 7 shows the corresponding spectrum inside the VFCAL, which in the simulation is assumed to be a solid block of pure copper. The absolute value of the flux decreases with radius and as a function of $z$, but the shape of the spectrum remains unaffected. Again even small amounts of hydrogen would provide significant moderation of < 10 MeV neutrons. Most important are the fluxes in positions where readout electronics might be positioned. As can be seen from the left plot of Fig. 8 the

---

$^1$ The actual production peaks are narrower but the spectrum is broadened to lower energies by scattering processes.
neutron fluxes at the back face of the VFCAL range from $10^6$ cm$^{-2}$ s$^{-1}$ to more than $10^7$ cm$^{-2}$ s$^{-1}$. Even on the outer side surface the fluxes are of the order of $10^6$ cm$^{-2}$ s$^{-1}$. According to Fig. 7 about half of this flux is above the critical 100 keV limit. It is worth to note that these fluxes are comparable to those encountered at outer layers of the inner tracker. Some polyethylene insertions may be useful to protect readout electronics, but it must be emphasized that the fluxes scored in the calorimeter bulk do not give an exactly correct picture of fluxes on the actual surface. The concrete shielding will provide moderation so that the flux on the VFCAL surface (copper/concrete interface) will have an at least partially softened spectrum.

8 RADIATION DOSES IN THE CALORIMETERS

The bulk of the calorimeters also suffers from radiation, although less than most semiconductor devices. The main concern is the degradation of light yield and transmission in the ECAL crystals and the plastic scintillators of the HCAL [29, 30]. These can lead to spatial and temporal variations in the response. After the HCAL was extended from $|\eta|=2.6$ to $|\eta|=3.0$ to cover the imposed by the combined muon chamber shielding and forward HCAL support [3], its edge was exposed to radiation doses of 30 kGy/year in the averaged material [3]. In order to shield the HCAL edge from the electromagnetic shower it has been proposed to extend the more radiation hard ECAL also to $|\eta|=3$. The dose rates for such a configuration are shown in Fig. 9. The pronounced spikes in the HCAL are due to the copper/scintillator sandwich. According to the Bethe-Bloch formula the dose of minimum ionizing particles in the scintillator should be only 50% higher than in copper. The difference by a factor of 2–4 arises from the low energy components of the cascade, in particular from neutrons which transfer more energy to the hydrogen in the plastic than to the copper atoms. This kind of signal amplification by neutrons in hydrogenated active layers is well known from the theory of compensating calorimetry [31, 32].

9 PARTICLE FLUXES IN THE MUON SYSTEM

The forward muon system of CMS is based on CSC and RPC detectors. The RPCs are likely to saturate at a signal rate of about 1 kHz/cm$^2$ [33]. A comparable limit is posed by the occupancy of the CSC detectors. Thus each muon station has to be shielded so that the total signal rate at peak luminosity stays below 1 kHz/cm$^2$.

Even at its maximum – in the low edge of MF1 – the muon rate is only of the order of 50 Hz/cm$^2$. Most of the signal rate is due to other background, which is composed of charged hadrons, photons and neutrons. Charged particles lead directly to signals by ionizing the gas. For neutrons two principally different ways to induce signals are possible. Upon elastic scattering neutrons transfer momentum to the atoms of the gas. If this energy is high enough the recoil ion may lead to a visible signal. By simple kinematics hydrogen recoils are the most likely to produce signals and thus hydrogenated gases will give higher signal rates in a neutron environment than non-hydrogenated ones. Usually more important, however, is neutron capture, which is most likely to occur at thermal energies. With a few rare exceptions a nucleus, after having captured a neutron, emits one or several photons. These, like all other photons, can release electrons by the photoelectric effect, Compton scattering or pair pro-
The hadronic production models have to provide reliable cascades that can traverse significant distances in most materials. If they reach the active volume these electrons give visible signals in the detectors.

Detailed simulations have shown that the sensitivity factor for a neutron to produce a signal by the recoil mechanism in the muon chambers of CMS is of the order $1 - 3 \times 10^{-4}$ to be compared with a sensitivity factor of $1 - 2 \times 10^{-2}$ for photons. Including the contribution due to capture gamma emission the total neutron sensitivity factor – which is the experimentally observable one – rises to $2 - 6 \times 10^{-5}$. Obviously these sensitivity factors depend on the granularity of the detector. The values quoted above are obtained by treating energy depositions initiated by two different tracks as separate if they were more that 1 mm apart from each other. The artificiality of this cut reflects the uncertainty in the flux to occupancy conversion. It has to be understood that this depends on the detector type (pixels, strips or pads), the granularity, the thickness of the active layer and the signal threshold required.

Since shielding of photon fluxes is relatively easy, and quite sufficient shielding is provided already by the CMS iron yoke, most of the neutral particle problems are connected with neutrons. These are more penetrating than photons and can traverse significant distances in most materials. If a neutron enters a muon chamber volume it can produce a capture gamma at a position where there would otherwise be no background photons.

Thus the signal rate is mainly due to the photons which are generated by the neutrons inside or on the boundary of the chamber volume.

The simulation of the signal rate in the muon system of an LHC experiment is one of the most demanding tasks for a simulation code. The code first has to be able to transport reliably the high energy particles in a complicated geometry in the presence of very general magnetic fields. Then the hadronic production models have to provide reliable cascade development and neutron emission at all angles. Until recently it was here that high energy physics codes stopped – and some still do. But modern codes like FLUKA have integrated neutron transport modules which perform the transport of the neutrons down to thermal energies. Care has to be taken that also the important capture gammas are properly generated in all materials. Finally the produced photons have to be transported.

In principle, at least, the proper treatment of all this physics makes a full radiation field simulation feasible. In practice, however, it is often problematic to obtain sufficient statistics. The methods to improve on this are provided by different variance reduction techniques and/or a stepwise simulation.

An example of stepwise simulation is provided by simulating the neutron fluxes and sensitivity factors separately: the fluxes in the middle of chambers consisting of averaged material are scored in a general simulation with relatively high energy threshold for electromagnetic particles. Simultaneously a sample is taken of all particles entering the chamber. In a second simulation only the immediate surroundings of the chamber are described. The chamber itself is described in detail. The recorded sample of particles is then injected into this geometry. All energy thresholds are lowered to the EMF minimum of 1 keV and detailed scoring of signals is performed. At the same time fluence is scored in the middle of the chamber. This fluence is then used to check for consistency and to normalize the obtained signal rate to unit flux. With the neutron sensitivity factor of $10^{-5}$ this means that good statistical significance of the neutron flux can be obtained with three orders of magnitude less CPU-time than would be needed for a direct signal scoring. In the second simulation, of course, the recorded sample is used repeatedly in order to obtain statistics for the sensitivity factor. Obviously the sample has to be large enough to be representative. One important complication is introduced by electrons, which get easily double-counted. Since, however, the CMS muon chambers are all surrounded by solid shielding only very high energy charged particles can penetrate directly to them. Thus essentially the whole electron flux is included in the photon sensitivity factors if incident photons and neutrons are used to obtain them.

In addition to making the first simulation phase faster and simpler (by removing the need for complicated scoring and adjustment of energy cuts), this method has an additional advantage: although it may be less accurate for one specific fixed geometry, it is much more general. Since the chamber itself is assumed not to affect the fluxes too much (most of the inaccuracy arises from this assumption) the first simulation provides generally valid flux values. If the chamber type is changed only the second simulation phase needs to be repeated in order to obtain updated sensitivity factors.

In the CMS simulations biasing has been mainly needed to obtain reasonable statistics for charged hadron fluxes. The main techniques used are Russian roulette at interactions and leading particle biasing in EMF combined with region importance biasing in the calorimeters, the yoke and the external shielding. Due to the complicated geometry of the detector an adjustment of the weights is a very complicated task. Therefore only moderate region importance in-
Figure 11: Neutron, photon and charged particle fluxes and estimates of resulting background signal rates in the muon stations of CMS at LHC peak luminosity. The bars indicate only the statistical errors of the simulation. Systematic errors due to accuracy of geometry description, neutron cross section and sensitivity factors can be as large as a factor of three.

Fig. 10 shows two proposed alternatives for shielding the most critical high-$\eta$ region of the muon spectrometer. In both cases the thick steel cone between $\eta=2.5$ and $\eta=3.0$ is essential. The first alternative [27] is to use an insertion of polyethylene ($\sim 10$ cm), cadmium ($\sim 0.5$ mm) and lead ($\sim 5$ cm) in the iron just below each chamber. Borated polyethylene is about equivalent to the PE/cadmium layer. The second alternative [3, 11] is to line the chamber with a layer ($\sim 2$ cm) of borated PE. While the first alternative aims at reducing the number of neutrons and photons entering the chamber the second is based on the idea that thermal neutrons get predominantly captured in the boron from which only harmless low energy photons are emitted. Results for the second option are presented in the CMS-TP [3], and for the first one in Ref. [27] and Fig. 11. As a modification to the CMS-TP [3] the thin shielding part behind the VFCAL has been changed from 10 cm steel + 30 cm heavy concrete to 30 cm steel + 10 cm borated polyethylene, 2 cm lead has been added to the wall shielding MF4 and the shield above Q1 has been made thinner to allow access to Q1. In addition the borated PE lining of the forward chambers has been changed to the PE/lead shield as shown in Fig. 10.

In general the fluxes shown in Fig. 11 are rather close to those of the CMS-TP[3]. In the critical high-$\eta$ region the photon fluxes (and thus the signal rate) are lower by a factor of about three, whereas the neutron flux is slightly higher. Both effects are due to now introduced lead, which attenuates photons very effectively but multiplies high energy neutrons. There is some controversy about the usefulness of lead, but it appears that the gain due to reduction of photon flux is larger than the loss due to the increase of neutron flux. It must be taken into account, however, that impurities in the lead could significantly degrade the performance of the lead layer. Also the idea of the borated PE lining to reduce the photon energy, is not properly accounted for if the same sensitivity factors are applied to compare the two alternatives.

10 RADIATION DOSES IN OCCUPIED AREAS

The occupied area closest to the experimental cavern is the counting room. In order to minimize cable lengths it must be situated as close to the experiment as possible. However, the radiation level produced by the inelastic $pp$-interactions alone would reach $\sim 1$ Sv/h if no shielding is provided around the collimator [34]. Even though the
shielding required by the muon system reduces this value by almost three orders of magnitude the remaining dose is far above the CERN limit (25 μSv/h) for controlled radiation areas [35].

Figure 12: Dose equivalent at the wall of the experimental hall (left) and required concrete thickness to attenuate this dose to the design level of 10 μSv (right). The upper plots show values close to the hall roof, middle ones are centered around the beam line and the lowest indicated the values just above the floor [36].

The simulation of dose attenuation in shields is quite different from the simulations aimed at determining detector performance. The accuracy requirements, in particular, are often less demanding. This allows dose equivalent to be determined from apparently very artificial estimators. Three quasi-independent methods for scoring dose equivalent have been used [36]. The fastest method, which does not require transport of electromagnetic cascades and low energy neutrons, is to score the density of high energy (> 50 MeV) hadronic interactions (stars) in concrete. The dose is then obtained from these by using a conversion factor of 4.5 × 10⁻⁵ Sv cm² star⁻¹ [37]. This factor has been obtained from fits to simulation results and experimental data and thus includes all contributions. Its major drawback is that it depends on a quantity called star, which is not well defined and different codes produce slightly different amounts of them [38].

The second method is to use energy deposition, which is a well defined quantity and provides the best estimator for dose (in Gy) but requires the time consuming transport of low energy neutrons and electromagnetic particles. Dose equivalent (in Sv) is related to this dose by an artificial quality factor Q, which is defined by authorities rather than by physics. The value of this factor for a complicated hadron spectrum is not obvious and a value of 5 is often assumed to be a good average. The third method, which in principle is the most exact one, consists of flux scoring and weighting with energy and particle dependent flux to dose equivalent conversion factors. These factors are experimental and the major uncertainties arise from the experimental methods and the treatment of low energy protons and electromagnetic particles which are partly included in the conversion factors of neutron flux. It has been shown, however, that in LHC shielding problems the purely electromagnetic dose amounts only to some 10% of the total and can thus be safely neglected.

Regardless of the estimator used the determination of shielding thickness is a deep penetration problem where heavy biasing is mandatory in order to obtain any results at all. Since the required attenuation amounts to roughly five orders of magnitude on average 10⁸ particles are needed to get one through the shielding in an analog simulation. Region importance biasing or weight windows can be used to turn this exponential increase of the required number of primaries into a linear behaviour. If, starting from the source, the region importance is increased by a factor of e per attenuation length, particles are splitted at region boundaries so that the average number of them in each region is the same. The statistical weights of the particles then contain the information about the number of traversed attenuation lengths.

Fig. 12 shows the dose equivalent at the inner surface of the CMS cavern. From the strong z-dependence it can be seen that the collimators act as the main radiation source and that the iron yoke provides good shielding. The variation of the dose equivalent as a function of vertical distance from the beam line is only due to geometrical dilution and thus almost insignificant. The concrete thickness required to reduce the dose equivalent to the design limit of 10 μSv/h is shown on the right hand plots of Fig. 12. The z-dependence of the doses is directly translated into the required shield thickness.

Fig. 13 shows the attenuation characteristics of the concrete wall and provides a comparison between the three dose estimators discussed above. It is evident that a sim-
ple exponential attenuation coupled with geometrical dilution gives a good description of the attenuation. A single effective attenuation length of 50 cm indicates that the contributions from less (electromagnetic and low energy neutrons) or more (muons) penetrating radiation are negligible. It is striking that the flux and star density estimators are in good mutual agreement whereas the energy deposition estimator with $Q=5$ is lower by a factor of two. This would indicate that the quality factor is close to 10, which is not completely unreasonable for a radiation field dominated by neutrons with energies between 10 keV and 20 MeV.

Fig. 12 gives directly an indication where cable penetrations and passage ways should be placed. Clearly the best positions are in the center of the hall in the “shadow” of the iron yoke. Except for some cryogenics supplies the CMS cable paths will start below the detector and go with one bend into the counting room. These tunnels have a cross sectional area of 9 m$^2$ and both legs are about 8 m long. Universal tunnel transmission curves [39] then predict an attenuation by three orders of magnitude. Even when starting from the maximum value of Fig. 12 the dose equivalent at the tunnel exit would not exceed 2 $\mu$Sv/h. Thus a factor of five safety margin is obtained even under the most pessimistic assumptions and detailed simulations are not motivated.

11 DOSE RATES DUE TO INDUCED RADIOACTIVITY

In a high energy hadron environment materials get activated mainly by spallation reactions, although neutron activation of some specific elements like gold, tungsten and cobalt cannot always be neglected.

After an inelastic hadronic interactions the residual nuclei is often left in a radioactive state. Some of these decay within very short time scales and are of no importance. Excluding cumulative effects the activity concentration of a sample irradiated in a flux $\Phi$ for a time $t_i$ is given as a function of cooling time $t_c$

$$A(t_c) = \rho \frac{N_A}{M} \sum_j \sigma_j \Phi \left[ 1 - e^{-\lambda_{ij} t_i} \right] e^{-\lambda_{ij} t_c},$$

where $\sigma_j$ is the production cross section of isotope $j$, which has a decay constant $\lambda_{ij}$. $N_A$ is the Avogadro number, $M$ the molecular mass and $\rho$ the density of the sample.

Since well established decay data of radionuclides is available the dose rates due to induced activity can be determined from the radioisotope distributions. Unfortunately the latter are not so easily obtained, since the complete tabulations [40] of production cross sections $\sigma_j$ are available only in some relatively rare cases, in particular for aluminium and carbon, which are commonly used for hadron fluence monitoring.

In lack of explicit cross sections the common approach is to use average activation rates. These are often expressed in terms of material dependent $\omega(t_i, t_c)$-factors, which relate a simulated star density to the $\gamma$-dose rate due to induced activity. For some materials $\omega$-factors have been measured but the value depends also on the code used to simulate the star densities. Therefore there are significant uncertainties in these factors even for standard values of $t_i = 30$ days and $t_c = 1$ day [42]. Even more uncertain is the time development of the factors. Sullivan and Overton have provided an average formula [41]

$$D(t_i, t_c) \sim \ln \left( \frac{t_i + t_c}{t_c} \right),$$

which is quite accurate for medium mass nuclei but in some cases fails quite badly [42].

The usually adopted value for the $\omega(30, 1)$-factor in iron is $10^{-8} \text{ Sv h}^{-1}/(\text{ star cm}^{-2} \text{ s}^{-1})$ but it is suggested in Ref. [42] that this might be an overestimate by a factor of 2–3. The $\omega$-factors of most materials are within an order of magnitude of the iron factors [42].

As far as activation of solid materials in CMS is concerned there are two principally different cases to consider.

1. The central tracker is a very light structure with carbon, aluminium and silicon being the major elements. Since activation properties of silicon are very similar to those of the well known aluminium, fluence scoring and use of explicit cross section is feasible in this region. Thermal neutrons have to be considered also, since the Microstrip Gas Chambers (MSGC) might be equipped with gold strips.

2. Most of the structural materials, with the important exception of the ECAL, are either concrete, steel, copper or lead. For all these some $\omega$-factors are available so that star density scoring and conversion to dose rates is possible.

A special case is provided by the ECAL, which contains a significant amount of tungsten and thus is subject to thermal neutron activation. This contribution can be calculated from thermal fluxes and added on top of the dose rate due to high energy activation.

In order to have a simple tool for estimation of $\gamma$-dose rate due to activity induced in bulk objects a point kernel subroutine FIASCO\(^2\) was developed [42]. Starting from star densities obtained in a FLUKA simulation FIASCO can be used to determine the dose rate at any point in the experimental cavern. The code requires as input photon attenuation and buildup parameters for all materials in the geometry and $\omega$-factors for all activated elements. Alternatively explicit photon energies, branching ratios and activities can be given. The FLUKA geometry is divided into a three dimensional cartesian binning and from each bin a ray is started to the point where the dose is to be calculated. Transport is done with the tracking routines of FLUKA. The material thicknesses traversed by the ray are recorded and this information is used to determine the attenuation and remaining contribution to the dose rate. FIASCO can deal with an arbitrary number of photon energies and thus could start also.

\(^2\)FIASCO=Fluka Induced Activity Shielding in COMbinatorial geometry
Figure 14: Induced activity dose rates (µSv/h) calculated with the FIASCO code starting from star densities in the bulk materials. Dose rates are for 60 days of irradiation at mean luminosity of $5 \times 10^{38} \text{cm}^2\text{s}^{-1}$ and 1 day of cooling. The FLUKA geometry shown is used for all simulations reported in this paper.

Dose rates obtained with the code at a few points in the CMS cavern are shown in Fig. 14. The values are throughout lower than the ones presented in the CMS-TP [3]. A factor of two arises from the lower average luminosity assumption. Most of the difference is explained by the fact that the model used in the CMS-TP [3] was constructed to stay on the safe side whereas FIASCO aims at improved accuracy. As a corollary there is no “build-in” safety factor in the values of Fig. 14.

Contrary to most bulk objects the contact dose rates of thin material layers may be dominated by $\beta$-activity. The tracker and beam pipe are the most important activated thin objects at the CMS experiment. They are exposed to relatively high fluxes and – the tracker in particular – represent large surface areas. In addition to $^{22}$Na and $^{24}$Na the only radioisotope of any significance is $^7$Be. However, compared to the large $k_\gamma$-factors and $\beta$-emission of the two sodium isotopes $^7$Be is radiologically close to negligible.

Table 2 lists the characteristics of the most important radioisotopes which can be formed in the tracker materials. In addition to high energy reactions $^{24}$Na is produced from aluminium (but not from silicon) by $(n,\alpha)$-reactions. In practice, however, $^{24}$Na production is insignificant, since an opening of the CMS endcap will take several days. By the time anybody is able to access the tracker only $^{22}$Na, $^7$Be and possibly some last remnants of $^{108}$Au will contribute to the dose rate.

The FIASCO code is not suited for thin layers and cannot handle $\beta$-activity. Using the cross sections of Table 2 the best accuracy is achieved by calculating the induced activity in the tracker materials from the particle fluxes. For this purpose the charged particle fluxes can be approximated by a $r^{-2}$-dependence, in particular since this can be made to provide safe overestimates. Like in Fig. 2 the normalization point can be taken to be 20 cm, where the average flux is roughly $2 \times 10^6 \text{cm}^{-2}\text{s}^{-1}$. The support structure of the tracker has a mass of roughly 1000 kg, which is almost uniformly distributed. To a first approximation this mass can be assumed to be half aluminium and half carbon.

Gold strips can be used to minimize aging effects of the MSGCs. The strips have a thickness of 5 µm and cover about 40% of the detector area. From simulations the thermal neutron flux in the tracking cavity is found to be of the order of $5 \times 10^6 \text{cm}^{-2}\text{s}^{-1}$ giving a $^{108}$Au production rate of 600 s$^{-1}$ cm$^{-2}$, i.e. three orders of magnitude smaller than the flux. This is fortunate, since it means that the presence of gold, which was not included in the simulations, does not

Table 2: Activation reactions, cross sections and decay data for CMS tracker materials. $\langle \sigma \rangle$ gives an approximate average over the particle spectra in the tracking cavity. The $k_\gamma$ factors are in units of fSv h$^{-1}$Bq$^{-1}$ at 1 m distance. Silicon is activated essentially like aluminium.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>$\langle \sigma \rangle$ (mb)</th>
<th>$t_{1/2}$</th>
<th>$n_\beta$/Bq</th>
<th>$\langle E_\beta \rangle$</th>
<th>$k_\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C $\rightarrow$ $^1$Be</td>
<td>10–20</td>
<td>20</td>
<td>53.6 d</td>
<td>EC</td>
<td>7.8</td>
</tr>
<tr>
<td>Al $\rightarrow$ $^7$Be</td>
<td>0.1–10</td>
<td>10</td>
<td>53.6 d</td>
<td>EC</td>
<td>7.8</td>
</tr>
<tr>
<td>Al $\rightarrow$ $^{22}$Na</td>
<td>10–40</td>
<td>20</td>
<td>2.6 y</td>
<td>0.9</td>
<td>0.215</td>
</tr>
<tr>
<td>Al $\rightarrow$ $^{24}$Na</td>
<td>5–20</td>
<td>20</td>
<td>15 h</td>
<td>1.0</td>
<td>0.553</td>
</tr>
<tr>
<td>Au $\rightarrow$ $^{198}$Au</td>
<td>$10^5$ (n$_\text{th}$)</td>
<td>2.7 d</td>
<td>$\sim 1.0$</td>
<td>0.315</td>
<td>60</td>
</tr>
</tbody>
</table>
influence the fluxes themselves.

The $\gamma$-dose in Sv/h at a distance $r$ (cm) of a point source of activity $A$ (Bq) is

$$H_\gamma = 10^{-11} \frac{A k_\gamma}{r^2}.$$  \hspace{1cm} (3)

$\beta$-particles are quickly attenuated in air but contribute mainly to the contact dose rate by the approximate formula [43]

$$H_\beta = 0.19 \times 10^{-6} \frac{A}{x_1} E_\beta [1 - \exp(-\frac{x_1}{\lambda_1} \exp(-\frac{x_2}{\lambda_2})].$$  \hspace{1cm} (4)

where $\langle E \rangle$ is the mean energy in MeV, $x_1$ is the thickness of the active material in g/cm$^2$, $\lambda_1$ is the apparent mean free path in the active material and $A$ is given in Bq/cm$^2$. The result is in Sv/h. Typical values of $\lambda$ are around 0.1 g/cm$^2$ [43]. The last exponential takes into account attenuation in some non-active secondary layer. This term is significant for the $^{198}$Au-$\beta$-s of the MSGC:s, which get attenuated in the glass and plastic substrates enclosing the gas volume.

Fig. 15 shows the $\gamma$-dose rates at the center of the whole tracker and at the center of a single wheel. Also shown is the dose rate a person would get when standing next to a single wheel. In addition contact dose rates are shown for the innermost pixel, silicon strip and MSGC modules (i.e. worst cases for each detector type). The operation schedule of Fig. 1 is clearly reflected in the dose rate. Even the short stops are visible as the large variation of the relatively short lived $^{198}$Au dose. It must of course be emphasized that the run-time dose rates in Fig. 15 are severely underestimated, since the contributions from all short lived radioisotopes ($^{11}$C, $^{10}$F and $^{24}$Na in particular) are neglected. As far as detector maintenance is concerned the results indicate that tracker wheels can be stored in an occupied area without any special shielding. Access restrictions to the immediate vicinity of the wheels may, however, be recommended. The obtained contact dose estimates show that all MSGC:s are harmless after about 2 weeks of cooling. Even at the pixel layers, where the $^{24}$Na activity reaches the highest values, the contact dose rate remains well within allowed limits for hand exposure.

### 12 PRESENT STATUS AND FORTHCOMING TASKS

FLUKA has proved its usefulness as the principal code for radiation environment simulations at hadron accelerators. Much of this success is due to the high quality physics modules and especially the recent improvements they have undergone.

A good basic understanding of all radiation aspects at LHC has been gained from the FLUKA simulations. This includes several ideas about needed and possible shielding. The main results concerning the CMS experiment, with the exception of beam loss and beam halo effects, have been presented in this paper.

An unresolved question is why FLUKA and MARS simulations disagree on some questions like charged fluxes in muon chambers [3] and neutron fluxes in the VFCAL [13]. It is likely that these differences are due to geometry and material descriptions or different cross section data but verifying the origin of all significant discrepancies is a prerequisite for the detailed design of shielding for CMS.

Such design work, based on the acquired knowledge of the radiation environment, is the next step when proceeding towards a final detector. The tentative shields used in all previous simulations have to be transformed into materials and configurations which fit into the overall detector layout. This requires detailed accounting for cost, space limitations, interferences with physics performance, mechanical feasibility and safety aspects, in particular the flammability of the important PE-moderators.

### 13 ACKNOWLEDGEMENTS

I want to express my gratitude to P.A.Aarnio for a good collaboration in CMS radiation simulations and for his valuable comments to this manuscript. The parts concerning personnel shielding and activation of materials are strongly influenced by discussions with G.R.Stevenson. The advices of A.Fassò and A.Ferrari have been invaluable for using the FLUKA code in an efficient way.
14 REFERENCES


