Joanna Weng

Parameterisation of electromagnetic showers and search for large extra dimensions in *pp* collisions at √s = 14 TeV

universitätsverlag karlsruhe

Joanna Weng

Parameterisation of electromagnetic showers and search for large extra dimensions in *pp* collisions at √s = 14 TeV

Parameterisation of electromagnetic showers and search for large extra dimensions in *pp* collisions at √s = 14 TeV

by
Joanna Weng

Impressum

The reasonable man adapts himself to the world;
the unreasonable one persists in trying
to adapt the world to himself.
Therefore all progress depends on the unreasonable man.
George Bernard Shaw

Contents

Introduction

It has always been the wish of mankind to understand the foundations of the apparently extremely complex world around us. In search for a theoretical description of the constituents of matter, the discovery of electrons, protons, neutrons and many other new subatomic particles in the last century, has led to the development of the Standard Model (SM) of particle physics. In a simple and elegant way it states that everything in our universe is built up from a handful of elementary particles which can interact with each other according to a few basic rules. The SM is not only an arrangement of the particles, but a theory in which their properties are largely determined by principles involving mathematical symmetries. It is considered as one of the greatest achievements in particle physics and was honoured by the Nobel price in 1979. The predictions of the Standard Model were tested against experimental observations down to length scales of 10^{-15} mm. Nevertheless, there has been at least one severe problem in the Standard Model at the beginning: in contrast to our everyday's experience, all particles had zero mass. In order to rescue the theory, the Higgs field has been postulated in 1964 by Peter W. Higgs [1, 2], Brout and Engelbert [3] and added to the SM. This mechanism generates the mass of the W and Z bosons by interacting with a new scalar particle - the Higgs boson - and requires the existence of one or more of these Higgs bosons. The discovery of this crucial ingredient of the Standard Model has been one prime goal of high energy experiments ever since. There are also other limitations of the SM, for example the bothering fact that all attempts to integrate gravity - the force we are most familiar with - into the SM had not been successful. In summary, one can say that the Standard Model is considered to be an impressively accurate but incomplete description of particle physics phenomena and is probably an effective theory, i.e. a low energy limit of a more fundamental theory not yet known. The hunt for new phenomena "Beyond the Standard Model" (BSM) has been therefore - beside the Higgs boson hunt - the focus of attention in high energy physics.

History teaches us that big jumps in human innovation come mainly as a basic result of pure curiosity, and that the primary force for innovation is fundamental research. Today we are in the privileged situation to experience such a innovation in particle physics: the start-up of the most powerful particle accelerator ever in 2008, the Large Hadron Collider (LHC) at CERN. The LHC will break new ground: it

has the potential to look for the Higgs boson in the entire mass region allowed by theory and to probe the TeV energy scale, where new phenomena are expected in many models; it promises to be the most exciting moment in particle physics since many years.

This thesis has been written during the preparation time before the first LHC run in the scope of the Compact Muon Solenoid (CMS) experiment, one of the two general purpose detectors at the LHC. It is divided into two parts: In the first part the experimental environment in which this thesis is set is described in general, focusing on the software sector, which has been the main working field. As the basis of the project to experimentally explore energy scales in the TeV-region and to search for New Physics, the Large Hadron Collider and the CMS detector are introduced. Its various technical challenges and most striking features are described in chapter 1. The enormous complexity of this flag ship project of particle physics is also reflected in in the complex software. This does not only include the programs to control the accelerator and the detector, but also the software to perform the simulation, reconstruction and physics analysis. A validated, well documented, portable and flexible software is crucial for a successful commissioning of the CMS experiment and a very important milestone of the preparation phase. First, the CMS software is briefly sketched, then the variety of technical contributions, tools and projects that were worked on in this thesis and are currently used in the CMS community is presented, focusing on the projects to which significant contributions were made. Three examples are discussed in more detail (chapter 2).

The next chapter, chapter 3, focuses on one particular component of the CMS software: the simulation of the detector response using the Geant4 toolkit. The detector simulation is by far the most time consuming step when simulating Monte Carlo events. Physics studies on large samples of simulated data are essential in order to understand the detector and the real data, once available, and has been a major activity in the CMS collaboration in the preparation time before the LHC start-up. A significant amount of the computing time is spent in the calorimeters when simulating electromagnetic showers. One goal of this thesis has been to speed up the simulation by using the concept of parameterised shower profiles instead of fully simulating every single particle. Once the shower parameterisation and its physics performance is tested and validated, it serves as an essential ingredient to speed-up significantly large scale Monte Carlo production in CMS. It is shown in this thesis that with shower parameterisation the full CMS simulation of one LHC event is up to 4 times faster, depending on the event topology. It therefore has the potential to save computing resources and is of high importance and benefit for the CMS collaboration. After a brief review of the physics of electromagnetic showers and a short introduction to calorimetry, the implementation of the shower parameterisation in the Geant4 framework and CMS is discussed in detail and the obtained results are presented. The results have been published in [4, 5] and in the

CMS Physics Technical Design Report volume 1 [6].

The second part of the thesis deals with physics analysis, where a particular New Physics model which can be probed by the LHC, is introduced and examined using the tools developed in part one. The focus lies here on so called models with extra dimensions. The possibility that spacetime is extended beyond the familiar 3+1-dimensions captivates the imagination and has intrigued physicists for the last century, inspired as well by some concepts from String theory. The consequences of a dimensionally richer spacetime would be indeed profound. Recently, new theories with higher dimensional spacetimes have been developed which offer some answers to the current problems in the SM, for example they resolve the hierarchy problem in particle physics. These scenarios make distinct predictions which allow for the LHC experiments to probe the existence of extra dimensions in new ways. At the beginning, a short review of the limitations of the Standard Model is given, followed by a short review of possible extensions and scenarios beyond the Standard Model. The model actually analysed, the Arkani-Hamed, Dvali and Dimopolous (ADD) Large Extra Dimension model [7], is discussed in more detail (chapter 4). Finally, the analysis of the CMS discovery potential within this model in the photon + missing transverse energy ($\not{E}_T$) final state as well as a method to control and calibrate the main background for this channel are presented (chapter 5). Large parts of this analysis has been published in a CMS analysis note [8] and included in the CMS Physics Technical Design Report volume 2 in 2006 [9]. The results show that in this channel, depending on the model parameters, a $5\,\sigma$ discovery at CMS can even be made with an integrated luminosity of less than 1 fb^{-1} of data.

Chapter 1
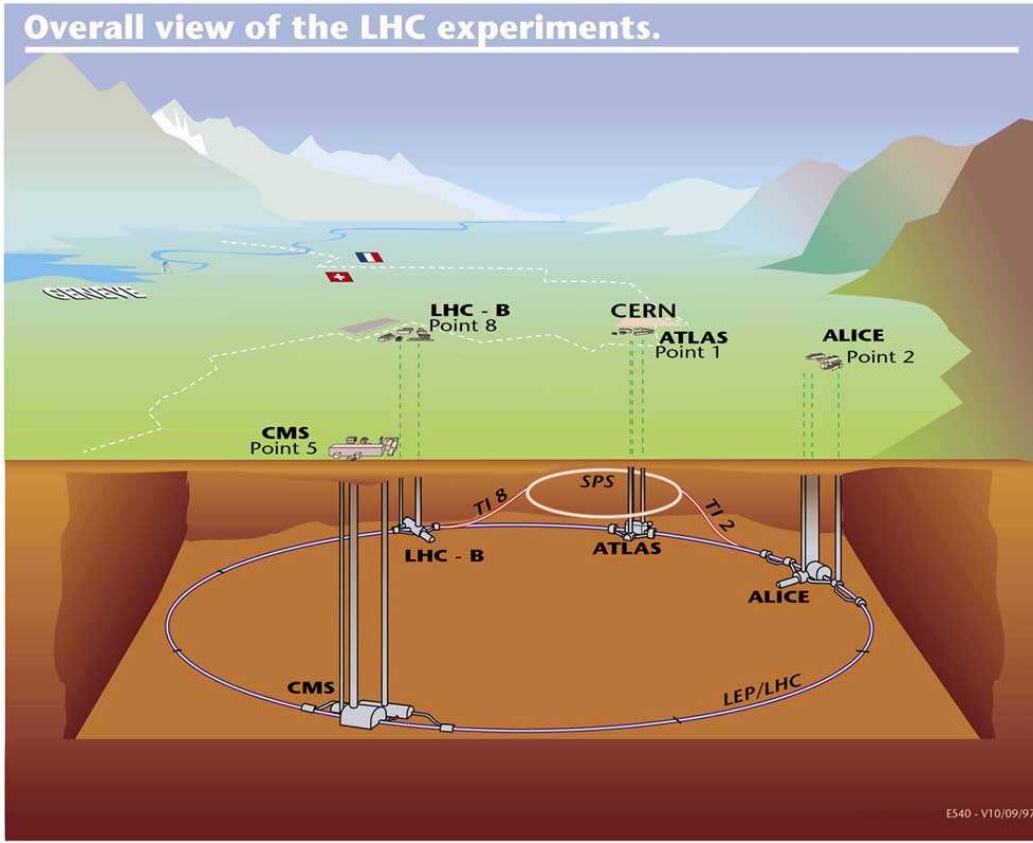
The CMS detector at the Large Hadron Collider

A number of fundamental questions of particle physics are expected to be answered by the new, exciting research instrument in this field: the Large Hadron Collider (LHC), currently under construction at CERN. Its primary goals are to explore the nature of electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible and to search for new physics Beyond the Standard Model (BSM, see chapter 4). In the following section the LHC and the Compact Muon Solenoid (CMS) are introduced. This description is mainly based on the CMS Physics Technical Design Report (PTDR) volume 1 [6] and contains the most notable and impressive facts.

1.1   The Large Hadron Collider

The Large Hadron Collider which is currently under construction at CERN is supposed to start the production of proton-proton collisions with a centre-of-mass energy of 14 TeV in 2008. This challenging new collider will offer the possibility to study a multitude of new physics topics due to the seven-fold increase in energy and a hundred-fold increase in integrated luminosity (see Equation 1.3) over the current hadron collider experiments. Currently the highest energy particle collider in the world is the Tevatron at the Fermi National Accelerator Laboratory in Batavia, Illinois (FERMILAB). Protons and antiprotons are collided here with a centre-of-mass energy of 1.96 TeV, the integrated luminosity is around 1.5 fb^{-1}. In contrast, at CERN two proton beams are accelerated to the high energy of 7 TeV each in the superconducting accelerator ring, which is installed in the 26.7 km long LEP/LHC tunnel shown in Figure 1.1.

At four interaction points huge detectors are foreseen : ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty experiment). ATLAS

Figure 1.1: Layout of the LEP/LHC tunnel with the four detectors ALICE, ATLAS, LHCb and CMS.

and CMS are designed as general purpose experiments; their main objective is to perform detailed measurements of Standard Model physics, for example investigate QCD in multi-jet and top events, as well as observe and study new physics effects not described by the SM. LHCb will be specifically dedicated to b-physics i.e. mainly to the study of CP violation in the B-sector in order to obtain precise measurements of the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [1]. Finally, the fourth detector, ALICE, is focused on the properties of quark-gluon plasma in heavy ion collisions. This phase of matter is interesting, since it is believed to have existed during the first 20 or 30 microseconds after the Big Bang. The B-physics programme will be carried out at the low luminosity running of the LHC to avoid too many overlapping events. For the heavy ion programme, the proton beams in the LHC accelerator will be replaced by ion beams. In Pb-Pb collisions, the nucleon-nucleon centre-of-mass energy reaches 5.5 TeV, which is significantly higher than what is obtained at the currently working heavy ion colliders. The geographical location of the detectors in the LHC tunnel can be seen as well in Figure 1.1.

---

[1] The CKM matrix describes the probability of a transition from a quark q to another quark q'.

The costs for these four detectors are already immense, but the LHC itself will be still many times more expensive, mainly due to the superconducting magnet system. In order to produce the magnetic field which is necessary to focus and bend the highly energetic beams into the right trajectories (up to 8.36 T), new types of superconducting niobium-titanium magnets had to be developed, which operate at liquid helium temperatures ($\sim$ 1.9 Kelvin). In view of the size of the accelerator - there are 1232 dipole magnets foreseen - this poses a major challenge for the cryogenics. Moreover, one has to carefully avoid that a magnet accidentally leaves its superconducting phase ("quenches") since this would lead to an enormous energy deposit in the magnet ($= RI^2$ with $I \approx 13000$ A). A picture of the LHC tunnel and schematic view of a LHC magnet are shown in Figure 1.2.

Since protons are about 2000 times heavier than electrons and the energy loss is

Figure 1.2: On the left: Schematic view of a LHC dipole magnet. On the right: the LHC tunnel with the installed magnets.

$\sim 1/m^4$, that means $(2000)^4 \sim 10^{13}$ times smaller, a proton collider can reach much higher energies than an e^+e^- collider. Proton colliders have an additional complication compared to lepton colliders: protons are not elementary particles as leptons are but composite objects, made of quarks and gluons which carry only a fraction of the protons momentum. This is the reason why the pp centre-of-mass energy of the collider should be in the multi- TeV range in order to be able to produce for example heavy Higgs bosons with a sufficiently high rate. The event rate dN/dt is given by the following equation and depends on one of the most important characteristic values of every collider experiment:

$$\frac{dN}{dt} = \mathcal{L}\sigma \tag{1.1}$$

with N being the number of interactions per time t. Equation 1.1 relates the observed event rate dN/dt with the corresponding cross section σ. The machine dependent proportional factor $\mathcal{L}$ is called *luminosity* and has to be measured e.g. via comparison to a theoretically well-known reaction. In order to gather as many events of a certain kind as possible, one would like to have a large luminosity. The luminosity of an accelerator which collides bunches containing n_1 and n_2 particles at a frequency f is given by:

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} \tag{1.2}$$

where σ_x and σ_y are the transverse beam profiles, approximated by Gaussian functions.
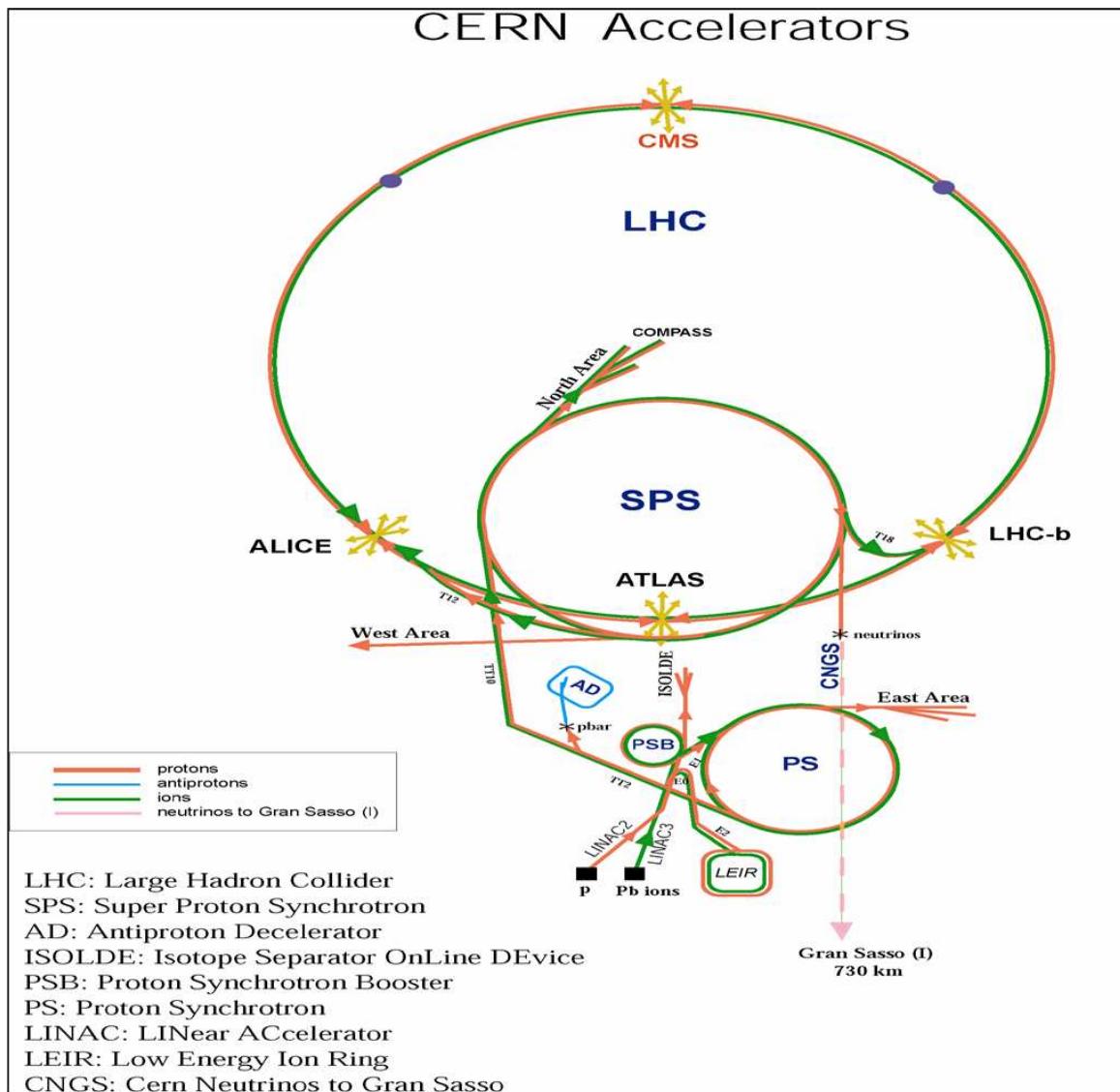
Figure 1.3: The CERN accelerator complex.

If the expected events for e.g. one year LHC are considered, Equation 1.1 has to be integrated over this time period and it follows

$$N = \mathcal{L}_{int}\sigma \qquad (1.3)$$

where $\mathcal{L}_{int}$ is the integrated luminosity and N is the number of expected events. The design luminosity for the LHC is $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1} (commonly referred to as "high luminosity" running). However, during the first three years the LHC will operate at a reduced luminosity of $\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1} (referred to as "low luminosity" running). The high luminosity running leads to around 1 billion proton-proton interactions per second, since the total proton-proton cross-section at 7 TeV is approximately 110 mb. This total cross-section can be broken down in contributions from inelastic scattering (∼60 mb), elastic scattering (∼40 mb) and single diffractive events (∼12 mb); it is only the inelastic scattering that give rise to particles at high angles with respect to the beam axis (high p_T values). Reaching the LHC design luminosity requires a small transverse beam profile, a high bunch collision frequency and a large number of particles per bunch. The luminosity is reached by filling each of the two rings with 2808 bunches of around 10^{11} protons. Because of this large number, the average number of inelastic pp collisions ("minimum bias" events) per bunch crossing is as high as 20 for the design luminosity. This leads to increasingly more difficult experimental conditions, since the rare interesting events that may occur in a bunch crossing are superimposed ('piled-up') on top of these 20 minimum bias events. The LHC bunch crossing rate will be 40 MHz, which means that a bunch crossing will occur every 25 ns. Such a high frequency, however, imposes stringent requirements on the response times of the LHC detectors.
Contrary to most hadron colliders in the past, the LHC will produce pp collisions instead of the more traditional $p\bar{p}$ collisions. The reason for this is that it is very difficult to produce sufficient amounts of antiprotons needed to achieve the LHC design luminosity. Since at LHC energies the most active components of the protons in the production of new particles are gluons rather than quarks and the distribution of gluons in protons and antiprotons is the same, it was decided to produce pp collisions rather than $p\bar{p}$. This decision has important consequences for the design of the collider. The LEP collider could use a single beam pipe with an elliptic transverse shape where electrons and positrons could circulate next to each other, since for a given electromagnetic field configuration the e^- and e^+ move in opposite directions. Because the LHC uses protons which carry the same sign, two beam pipes and two different (opposite) magnetic field configurations are needed. Before entering the LHC, the protons are first accelerated by the Linac and the Booster to an energy of 1.4 GeV. The bunches are then formed in the 25 GeV Proton Synchrotron (PS) with the correct 25 ns spacing. The beam is subsequently accelerated to 450 GeV in the Super Proton Synchrotron (SPS) and finally transferred to the LHC. The total energy stored in the beams at the maximum energy is 362 MJ, equal to about 77.4

kg TNT. The energy stored in the LHC magnets (11 GJ) would be enough to melt 50 tons of copper. The chain of accelerators is shown in Figure 1.3. After the filling procedure, which takes about seven minutes, the lifetime of the beam in the LHC is about 15 hours. Data taking is restricted to the first ten hours because after that time, due to the collisions, the luminosity has decreased too much to be useful. The most interesting LHC parameters are summed-up in Table 1.1.

Table 1.1: The LHC machine parameters, relevant for the detectors.
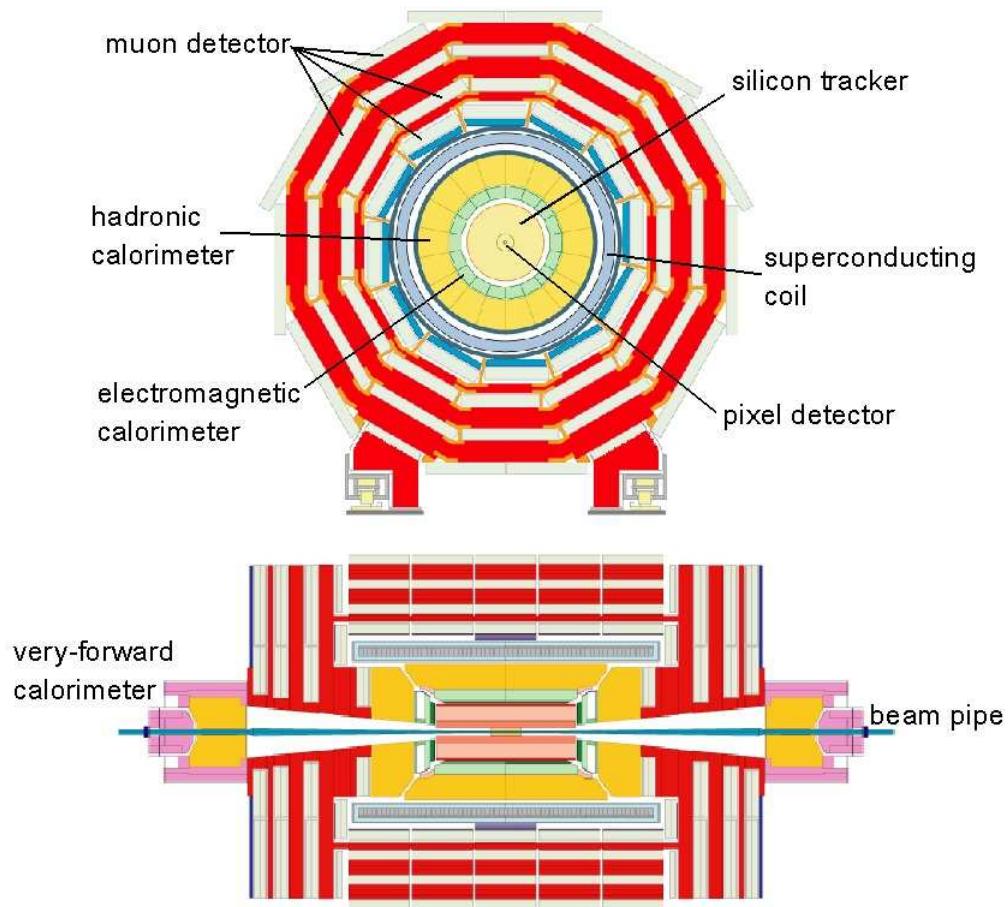
		pp	**HI**	
Energy per nucleon	E	7	2.76	TeV
Dipole field at 7 TeV	B	8.33	8.33	T
Design Luminosity*	$\mathcal{L}$	10^{34}	10^{27}	$cm^{-2}\,s^{-1}$
Bunch separation		25	100	ns
No. of bunches	k_B	2808	592	
No. particles per bunch	N_p	1.15×10^{11}	7.0×10^7	
RMS beam radius at IP	σ^*	16.7	15.9	μm
Luminosity lifetime	τ_L	15	6	hr
Number of collisions/crossing	n_c	≈ 20	–	

* For heavy-ion (HI) operation the design luminosity for Pb-Pb collisions is given.

The commissioning of the LHC machine with beams is expected to start in November 2007 with a 450 GeV calibration run. The aim is to establish first beam collisions at a centre-of-mass energy of 0.9 TeV. This run will consist of machine development periods interleaved with data-taking runs used for the commissioning of the detectors. Full commissioning to 7 TeV will take place in 2008 starting at 75 ns and subsequently 25 ns bunch spacing. The 75 ns operation is considered an important step in the commissioning of the LHC and the experiments. During the first full year of physics running, the LHC should reach a peak luminosity of $\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1}. However, the integrated luminosity will most likely be limited by the time taken to master LHC operation. The integrated luminosity is likely to be about 1 fb^{-1} in the first year. It may well be lower, as prolonged machine development periods may be required and higher than foreseen inefficiencies encountered. In summary, the following integrated luminosities are forseen: 1 fb^{-1} during the initial operation, 10–30 fb^{-1} in the "low luminosity" phase ($\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1}) and 100–300 fb^{-1} over several years of operation at design or "high luminosity" ($\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}). The first run with heavy-ion beams could be in 2008, after the first pp physics run. In the following year, depending on the machine performance, the luminosity is expected to reach the nominal 10^{27} cm^{-2}s^{-1} that corresponds to a minimum-bias interaction rate of 8 kHz.

1.2   The CMS detector

The CMS detector is a general purpose detector with a strong emphasis on the sensitivity to signatures expected from the Higgs boson as well as possible signatures from new physics or particles at the LHC. It is built by an international collaboration consisting of 2030 scientists from 174 institutions from 38 countries (status October 2006). From a conceptual point of view, a general purpose collider detector ideally should be designed as a perfect sphere around the collision point in order to detect all particles produced in the collision. For technical reasons, however, a cylindrical shape has been adopted, essentially driven by the solenoidal magnet shape, but still allowing for an almost 4π coverage. CMS is a large, technologically advanced detector made up of several sub-detectors, which are positioned in concentric layers around each other ('onion' structure) and each of them dedicated to different and complementary types of measurements. A general view of the CMS detector is shown in Figure 1.4.

Figure 1.4: Different views of the CMS detector.

Coordinate conventions

The coordinate system adopted by CMS has the origin centred at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the centre of the LHC. Thus, the z-axis points along the beam direction toward the Jura mountains from LHC Point 5. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$. Thus, the momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components.
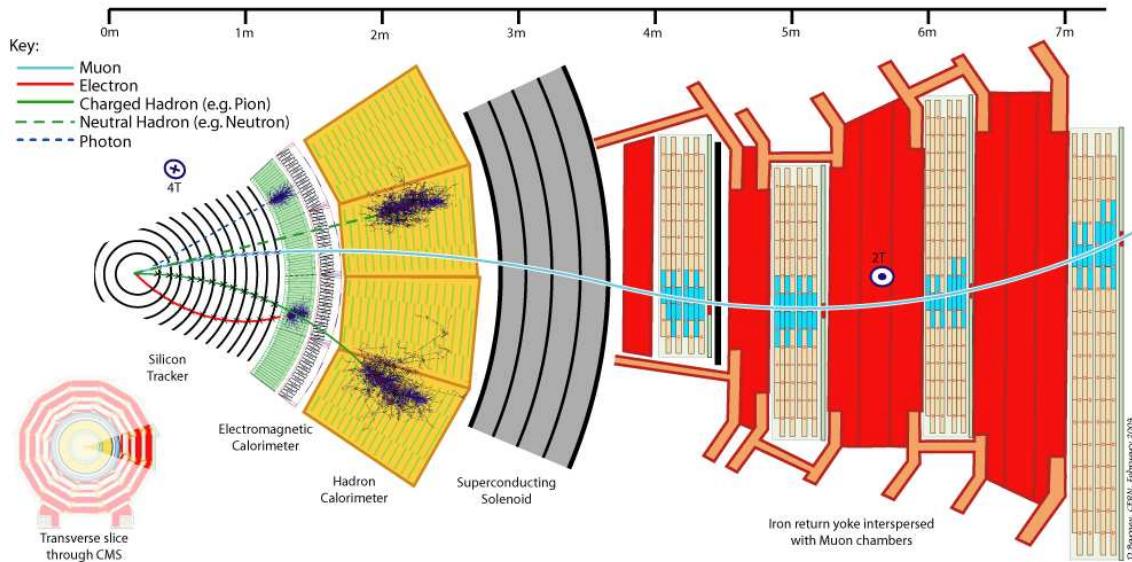
General overview and design

This paragraph gives a brief overview of the CMS detector, followed by a more detailed discussion of its components, focusing on the sub-detectors relevant for this thesis. The detector is divided into a barrel region and two endcap regions - overall it is nearly 22 m long with a width of 14.6 m, and a total weight of about 12500 tons. The heart of CMS is a 13-m-long, 5.9 m inner diameter, superconducting solenoid, generating a field of 4 Tesla which corresponds to a stored energy of 2.5 Giga Joules and is the biggest superconducting solenoid magnet with the highest field and stored energy ever. In order to achieve good momentum resolution within a compact spectrometer a high magnetic field was chosen. The return field is large enough to saturate 1.5 m of iron, allowing 4 muon "stations" to be integrated to ensure robustness and full geometric coverage. The emphasis on muons (in CMS even in the detector name) originates from the fact that muons are particles with a large penetrating power, which makes them very suitable for experimental observation (efficient detection and precise reconstruction even at very high luminosities). Furthermore, muons are also important signatures for Higgs boson decays or many types of new physics. Crucial to the detection and measurement of all charged particles is the choice of the magnetic field configuration. In the CMS detector the strong solenoidal magnet produces a magnetic field along the beam axis, thus bending charged particles in the transverse plane. The precision with which momenta then can be reconstructed depends on the strength of the magnetic field and the size of the detector. For a charged particle moving in a magnetic field, the following equation holds in the transverse plane (from balancing the centrifugal force with the Lorentz force):

$$p_\mathrm{T} = 0.3\,B r \tag{1.4}$$

where p_T is the transverse momentum of the particle (in GeV), B is the magnetic field (in Tesla) and r is the bending radius of the particle (in meter). The factor 0.3 is a conversion factor from SI units to GeV. A particle emerging from the collision and travelling outwards will first encounter the tracking system, which measures precisely the positions of passing charged particles, allowing to reconstruct the tracks.

Charged particles will follow spiralling paths in the CMS magnetic field and the curvature of their paths will reveal their momenta, as discussed above. This is a crucial point, since an important part of the CMS physics programme relies on the capability of the detector to reconstruct charged particle tracks and measure with high resolution their momenta and their vertex of origin. Experience has shown that robust tracking within a strong magnetic field is a very powerful tool for the identification and accurate reconstruction of muons, electrons, photons and jets. Moreover, accurate vertex reconstruction is expected to play an important role in jet flavour tagging, especially for b- and τ-jets. Momentum measurement of tracks in the 1-5 GeV range is also very important to define "isolated objects" (e, μ, γ, τ, etc.). The energies of the particles are measured in the next layer, the calorimeters. Electrons (e^-, e^+), photons and particle jets will be showering here, allowing their energy to be determined. The first calorimeter (ECAL) is designed to measure the energy of electrons and photons with high precision- since these particles interact electromagnetically, it is called electromagnetic calorimeter. Particles which interact by the strong interaction, hadrons, deposit most of their energy in the next layer, the hadronic calorimeter (HCAL). The only particles to penetrate beyond the HCAL are muons and neutrinos. The muons are measured in the muon chambers, their momentum is estimated from the bending of their paths in the CMS magnetic field. Neutrinos hardly interact and their presence can only be seen indirectly by adding up all transverse momenta of the detected particles ("Missing E_T or p_T"). An illustration of the particles tracks mentioned above can be seen in Figure 1.5.

Figure 1.5: Overview over particle tracks in the CMS detector.

Tracking system

The tracking system consists of a silicon pixel detector and a silicon strip detector. Three pixel layers and ten silicon strip layers are installed in the barrel. In each endcap, two pixel layers, three inner and nine outer forward disks of silicon detectors will be placed. The outer radius of the CMS Tracker extends up to 107-110 cm, the total length is approximately 540 cm. The pixel detector has an excellent position resolution ($\sim$ 15 μm) allowing impact parameter determination and vertex reconstruction with high precision. The whole silicon strip detector has a total active surface of 170 m^2 (225 m^2 if the double contribution of stereo detectors is included) instrumented with about 10^7 channels. Signals are read out by a charge sensitive amplifier, whose output voltage is sampled at the beam crossing rate (40 MHz). The silicon strip covers a pseudorapidity region of $|\eta| < 2.4$. A complete description can be found in the Tracker TDR [10].

Electromagnetic calorimeter

The electromagnetic calorimeter consists of $\sim$80000 lead tungstate (PbWO$_4$) crystals to perform the accurate measurement of electron and photon energies and their directions of flight. PbWO$_4$ crystals are chosen mainly because of their short radiation length ($X_0 = 0.89$ cm, due to the high density 8.2 g/cm^3) and a small Molière radius ($R_M = 2.2$ cm) [2], allowing for a compact ECAL design with narrow showers. The crystals are about 23 cm long (corresponding to almost 26 X_0), thereby containing more than 99% of the shower energy. A second advantage of using PbWO$_4$ is that the scintillating process is fast: 80% of the light is emitted within 20 ns, matching the LHC bunch crossing time of 25 ns. A third reason for using lead tungstate is that the material is intrinsically radiation hard (up to 10 Mrad), which is important due to the high radiation level inside CMS. Negative aspects are the low light yield (30 photons per MeV of incident energy) which requires a read-out through photodetectors with gain. Also, the light yield has quite a strong dependence on temperature, posing stringent requirements on the thermal stabilisation of the calorimeter. The lateral granularity of the ECAL is $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$, corresponding to a crystal front face of about 22 x 22 mm^2. The fine lateral size is required because of the need for a good π^0 rejection, to avoid that two photons from energetic π^0, which are emitted close to each other are reconstructed as a single photon. All the crystals are mounted in a projective geometry with a 3 degree tilt in η and ϕ with respect to the mean position of the primary interaction vertex in order to limit the effects of the inter-crystal gaps. The barrel section (EB) has an inner radius of 129 cm. It is structured as 36 identical "supermodules," each covering half the barrel length and corresponding to a pseudorapidity interval of

---

[2]see Equation 3.35 and Equation 3.6 for the definition of radiation length X_0 and the definition of the Molière radius R_M.

$0 < |\eta| < 1.479$. The endcaps (EE) are located at a distance of 314 cm from the vertex and are covering a pseudorapidity range of $1.479 < |\eta| < 3.0$. Further details can be found in the ECAL TDR [11].

Hadronic calorimeter

The hadron calorimeter surrounds the electromagnetic calorimeter and is used in conjunction with the latter to measure the energies and directions of particle jets and to provide hermetic coverage for measuring missing transverse energy. It also helps in the identification of electrons, photons and muons. The active elements of the barrel and endcap hadronic calorimeter consist of plastic scintillator tiles with wavelength shifting fibre readout. Layers of these tiles alternate with layers of brass absorber to form the sampling calorimeter structure. The tiles are arranged in projective towers with fine granularity to provide good di-jet separation and mass resolution. The pseudorapidity range $|\eta| < 3.0$ is covered by the barrel and endcap hadron calorimeters which are located inside the magnetic field of the CMS solenoid. The HCAL is subdivided into a barrel part (HB) of 9 m length and between 2 and 3 metres in diameter, covering a pseudorapidity region of $|\eta| < 1.48$, and two endcaps (HE) about 1.8 m thick with an inner radius of 40 cm and an outer radius of about 3 m, covering a pseudorapidity range of $1.48 < |\eta| < 3$. Outside the volume of the CMS detector, 6 m downstream each endcap, two forward calorimeters (HF) extend the hermeticity of the hadron calorimeter up to $|\eta| = 5$. The HF has an inner radius of 12.5 cm, and outer radius of 1.5 m and a length of about 3 m. Apart from improving the E_T^{miss} measurement, it is useful to tag or veto high p_T jets in the forward direction. At $|\eta| = 0$ the calorimeter is 79 cm thick, corresponding to about 5 nuclear interaction lengths (λ). This thickness is not sufficient for full hadronic shower containment, resulting in a low-energy tail in the hadron distributions. Such tails are one of the main sources of fake E_T^{miss} and thus need to be avoided. Therefore, in order to increase the sampling depth for $|\eta| < 1.4$, a hadron outer calorimeter (HO) is placed outside the solenoid, consisting of one scintillator layer. Its location enables the HO to exploit the solenoid magnet coil and the first muon absorber plate as additional absorbers. With this additional detector, which increases the total depth of the calorimeter to at least 10 λ in the entire η range, it is possible to identify and quantify the contribution from late starting showers. More details can be found in the HCAL TDR [12].

The superconducting magnet

One of the basic design features of CMS is a strong 4 Tesla magnetic field, which is provided by a superconducting solenoid of 13 m length and an inner diameter of 5.9 m. The tracking and calorimetry subsystems are completely enclosed within the field. The magnetic flux is returned via a 1.8 m thick saturated iron yoke which is

instrumented with muon stations. In the return yoke, the field is about 2T. The magnetic field provides the necessary bending power for efficient muon detection up to $|\eta| = 2.4$, hence making the addition of forward toroids unnecessary. Coil protection in case of a quench is considered to be a critical issue because of the very large stored energy (2.5 GJ). The conductor carries a current of 19.5 kA. A complete description is given in the Magnet TDR [13].

Muon system

The Muon system is crucial to the concept of the CMS experiment. It is composed of four muon stations interleaved with the flux return iron yoke plates and is divided in a barrel part ($|\eta| < 1.2$) and two endcaps ($0.9 < |\eta| < 2.4$). The total thickness of the absorber before the last muon station amounts to 16 interaction lengths, allowing good muon identification. The muon system uses three different technologies to detect and measure the muons: drift tubes (DT) in the barrel region, cathodes strip chambers (CSC) in the endcap region, and resistive plate chambers (RPC) in both the barrel and the endcap. The last type of detectors provide a lower spatial resolution than the others but provide a faster timing signal (time resolution $\sim$ 2-3 ns). The DTs or CSCs and the RPCs operate within the first level trigger system (see next section), providing 2 independent and complementary sources of information. The combination of these technologies therefore allows efficient triggering, identification and measurement of muons. In the initial stages of the experiment, the RPC system will cover the region $|\eta| < 1.6$. The coverage will be extended to $|\eta| < 2.1$ later. The full muon system covers $|\eta| < 2.4$. A more detailed description can be found in the Muon TDR [14].

Data acquisition and the event trigger

For a bunch crossing frequency of 40 MHz and a design luminosity of $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} approximately 20 inelastic collisions occur every 25 ns corresponding to an interaction rate of the order of 1 GHz. Not all of these interactions will produce interesting results; most of the time protons will just graze past each other. Head-on collisions will be rare, and the processes which produce new particles rarer still. The Higgs boson, for example, is expected to appear only a few times every day (< 10). Considering the $\sim 10^8$ channels of the CMS detector, the data stream coming from the detector can be estimated to be ~ 1 MegaByte per event, resulting in 100 TeraByte of data per second. This input rate of 10^9 interactions every second has to be reduced by a factor of at least 10^7 to 100 Hz and a data rate of ~ 100 MB/s in order to match the capabilities of the mass storage and offline computing systems. For this purpose, CMS plans to have a multi-level trigger system, which is designed to efficiently select events showing signatures of interesting physics processes. The first stage of this rate reduction is performed by the Level-1 (L1) trigger. The L1

trigger system is organised into three subsystems: the L1 calorimeter trigger, the L1 muon trigger, and the L1 global trigger. It decides whether to accept or reject an event within a few microseconds after a collision. During that period the full detector information is kept in the memory buffer. The maximum output rate of the L1 trigger is 100 kHz, which is determined by the speed of the detector electronics readout and the input of the data acquisition system. To account for the limited reliability of rate predictions, a safety factor of three is taken into account, therefore a maximum L1 rate of 30 kHz is foreseen in the studies. This rate is shared equally between muon and calorimeter triggers. The second stage is performed in software by the High Level Trigger (HLT) system (normally L2 and L3). The HLT stages have much longer processing times and are therefore based on computer farms. The L2 Trigger rate is designed to be about 5 KHz, the L3 Trigger rate about 100 Hz. Further offline analysis is planned to be also performed by global networks like e.g. the GRID. The quality of the selection algorithms is of high importance. A complete description of the Trigger and data acquisition (TriDAS) system is provided in the corresponding two TDR volumes, the [15], [16].

Chapter 2

The CMS software

While the CMS detector is under construction at CERN, the development of the software to simulate the detector response, perform the reconstruction and handle the large amount of data as well as to perform first physics analysis on this "toy" data, is a crucial and very important step to prepare for the commissioning and first data taking with the real detector. The software for this complex and ambitious project consists of many different parts: interfaces to common event generators as well as complete event generator libraries to generate Standard Model and new physics events at the LHC, detailed and fast simulation packages to simulate the interaction of the generated particles with the detector material, a variety of detector specific algorithms to perform the reconstruction and a collection of tools to facilitate physics analysis and visualise the results. Apart from these main ingredients, a multitude of external helper applications is needed. Since CMS computing is supposed to be carried out on a global network (GRID), a lot of effort has been invested as well in the tools to control these computing jobs, transfer huge data amount between different computing centres (called TIER centres) and make the CMS software portable enough to be (remotely) installable at different computing centres or local standalone computers. In the following section an overview of the most important CMS software components is given, focusing on the parts which are relevant to the physics analysis of this thesis. Then three software-related projects to which significant contributions were made are discussed in more detail.

2.1 Overview of CMS software components

The CMS software to perform physics analysis is a complex collection of standalone programs and toolkits; its central ingredients are the Monte Carlo simulation tools. Since the production and the decay of elementary particles in a high energy collision have a probabilistic nature, such events can be simulated using so-called Monte Carlo techniques. These are calculational techniques which make use of random numbers in order to distribute the events according to the probability densities calculated by

theory. Also the response of a detector to the passage of the particles produced in the collision through material involve random processes such as for example multiple Coulomb scattering, therefore the transport of particles through an experimental setup can also be effectively implemented using Monte Carlo techniques. Both steps, "event generation "and "detector simulation", are carried out by separate software packages in CMS. [1] The most important software packages (not complete) are listed in the following:

1. **CMKIN**: CMKIN [17] is the CMS standard way to interface physics generators such as PYTHIA [18] with the CMS detector simulation and provide the right LHC specific parameters. The interface is based on the common block HEPEVT - a HEP standard to store particle kinematics information for one event. The HEPEVT common block is converted to HBOOK [19] n-tuples for persistency. This output is then used as input for the detector simulation. In this work most of the time PYTHIA was used to generate physics events; besides the CMKIN interfaces to COMPHEP [20] and Madgraph [21] were used.

2. **COBRA** *(Coherent Object-oriented Base for simulation, Reconstruction and Analysis)*: COBRA [22] is a general framework including many tools, which is supposed to provide the functionality common to the ORCA, IGUANA, OSCAR and FAMOS packages and to account for the general run management.

3. **OSCAR** *(Object oriented Simulation for CMS Analysis and Reconstruction)*: OSCAR [23] is the full CMS detector simulation based on Geant4 [24, 25], which models the interaction with the detector material with high accuracy. It provides a description of the detector geometry and the material budget, including also information about the magnetic field. OSCAR reads the generated events and simulates the effects of energy loss, multiple scattering, showering in the detector materials etc. with Geant4. More details about Geant4 can be found in chapter 3 of this thesis.

4. **ORCA** *(Object-oriented Reconstruction for CMS Analysis)*: ORCA [26] is a C++ toolkit which performs the digitisation (simulation of the electronic response), the emulation of the Level-1 and High-Level Triggers (HLT) as well as the offline reconstruction of physics objects.

5. **FAMOS** *(FAst MOnte-Carlo Simulation)*: FAMOS [27] is the framework for the fast simulation of particle interactions in CMS. It models the detector resolution with parameterisations obtained from full simulation studies and

---

[1]In 2005 the CMS collaboration has decided to redesign its software framework and integrate everything in one software package called CMSSW. Here the old version of the CMS software is described, which has been used for this thesis.

takes the acceptance of the CMS detector into account. It is embedded in the same framework as OSCAR and ORCA and delivers the same kind of reconstructed high level objects, but can be up to a factor 1000 faster. FAMOS allows to perform fast sensitivity scans in a large parameter space as is typically necessary for models of new physics.

6. ***IGUANA****(Interactive Graphics For User Analysis)*: IGUANA [28] is a graphical tool for the inspection of simulated and reconstructed data, as well as 3-dimensional viewing of the CMS detector.

7. ***SCRAM*** *(Software Configuration, Release and Management)*: SCRAM [29] performs version management in CMS and sets up the environment for the user to create and run executables. The version management of the source code is based on CVS [30] (Concurrent Versioning System).

A typical physics analysis starts with the generation of the process that is supposed to be investigated; in most of the cases PYTHIA is used in CMS for this purpose, however the usage of any other event generator which is interfaced to CMKIN is also possible "out of the box". In the next step the events are processed through FAMOS, in case of fast simulation, or through OSCAR and ORCA, in case of full simulation. While the simulation with OSCAR can take up to several minutes, FAMOS usually needs only some seconds for one event. Once the reconstructed objects are obtained, the analysis can be performed using appropriate statistical, analysis and plotting tools.

2.2 Monte Carlo event generators in CMS

One sub-project in the scope of this thesis is focused on event generators and in particular on the object-oriented new event generator SHERPA [31], which has not been used in CMS so far. The aim was to test this new generator and develop a well defined interface to the CMS simulation framework, in order to use SHERPA for the analysis part of this thesis in chapter 5. In the following, first a short introduction into event generators in general is given, then the event generator SHERPA and its integration into the CMS software is described. An overview over all event generators used in CMS can be found in [32].

2.2.1 General overview

Monte Carlo (MC) event generators are an essential tool in all experimental analyses. In general the MC event generation process can be divided into three main phases:

1. The hard process, where the particles in the hard collision and their momenta are generated, usually according to the leading-order matrix element. Those

particles which decay before hadronisation, for example, the top quark (life time $\tau \leq 10^{-24}$ s), are decayed before the hadronisation phase and treated as a secondary hard process.
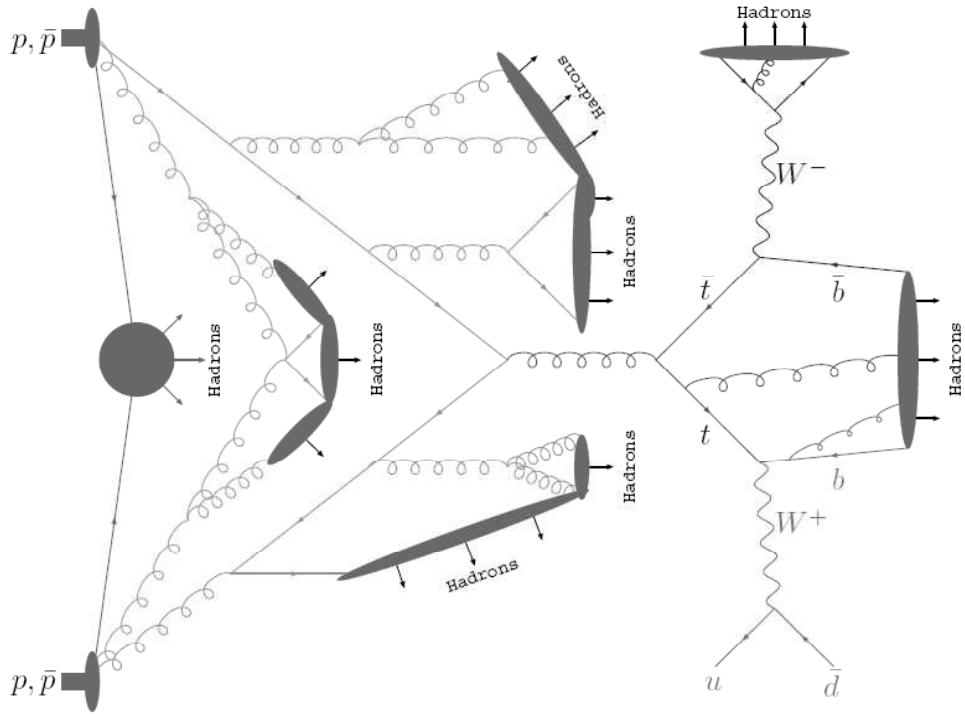
2. The parton-shower phase where gluons are emitted with high energy, leading to the creation of further $q\bar{q}$ pairs. The coloured particles in the event are perturbatively evolved from the hard scale of the collision to the infrared cut-off. The emission of electromagnetic radiation from charged-particles can be handled in the same way.

3. The quark confinement enforces the formation of colour neutral hadrons out of the colour field of the primary quarks and possibly emitted gluons. As a consequence, the final state consists of colour-neutral particles, that are grouped in jets pointing along the direction of the primary quark. This happens in the hadronisation phase in which the partons left after the perturbative evolution are formed into the observed hadrons. Afterwards, the decay of unstable leptons (τ-leptons, life time $\tau \leq 10^{-15}$ s) and hadrons is performed to form the experimentally observable final state. Hereby it has to be taken into account that long-lived unstable hadrons can possibly decay during the detection process, that is in or in between sub-detectors - this has to be taken into account properly by the detector simulation.

The different phases are schematically shown in Figure 2.1.

Most generators fall into one of two classes: general-purpose event generators aim to perform the full simulation of the event starting with the hard process and finishing with the final-state hadrons. The second class of programs perform only the hard scattering part of the simulation and rely on one of the general-purpose generators for the rest of the simulation.

General-purpose event generators

Historically, the main general-purpose generators have been HERWIG [33], ISAJET [34], and PYTHIA [18]. While the general philosophy of these programs is similar, the models used and approximations made are different. The parton-shower phase and the models used for the hadronisation phase are very different; for example, PYTHIA uses the LUND string model and HERWIG the cluster hadronisation model for the hadronisation. While these programs will continue to be used in the near future, a major programme is under way to produce a new generation of general-purpose event generators in C++. The main aim of this programme is to provide the tools needed for the LHC; work is under way to rewrite both PYTHIA and HERWIG in C++. These programs are expected to be available in the next few years.

Figure 2.1: Illustration of the different steps in the Monto Carlo generation of an event. The production of a $t\bar{t}$ pair is shown, followed by the hadronic decay of the top quarks.

Parton level programs

Many programs are available to calculate an individual hard process, or some set of hard processes. They are usually interfaced to one of the general-purpose generators, most often PYTHIA, which perform the parton shower and the hadronisation. In general, there are two important components of any such program. The first step is the calculation of the matrix element for a given momentum configuration. There are three different techniques in use to perform this step of the calculation:

1. The matrix element squared can be evaluated symbolically using traditional trace techniques.

2. A number of programs use helicity amplitude techniques to evaluate them.

3. There are techniques that can be used to evaluate the matrix element without using Feynman diagrams [35] .

The second step of the process is to integrate the matrix element. There are two main techniques in use: One approach is to use adaptive integration programs such as VEGAS [36] to perform the integration. A second approach is to use the knowledge of the matrix element to perform multi-channel phase-space integration based on its peak structure.

In practice some programs combine these two approaches. In general, adaptive programs such as VEGAS [36] are ill-suited to the integration of functions which have complex peaked structures, such as multiparticle matrix elements, and therefore for most practical applications multichannel integration techniques converge much faster.

There are a number of programs available which combine a variety of these techniques - here only three examples are given which are relevant in the context of this thesis:

AMEGIC++ [37] makes use of helicity amplitude techniques to evaluate the matrix element together with efficient multichannel phase-space integration to calculate the cross section.

CompHep [20] is an automatic program for calculation of the cross section for processes with up to eight external particles. CompHep can have up to six final-state particles for scattering processes and seven for decays. It uses the traditional trace techniques to evaluate the matrix element together with a modified adaptive integrator to compute the cross section.

MADGRAPH/MADEVENT [21] uses helicity amplitude techniques for the matrix element together with an efficient multichannel phase-space integrator to compute the cross-section.

In this work Madgraph and Comphep were used to generate the hard process (matrix elements) and then interfaced to CMKIN (PYTHIA) to perform the fragmentation. In addition, a realtively new event generator was used for which a CMKIN interface was not present: SHERPA. The setup is therefore described in more details.

2.2.2 The event generator SHERPA

SHERPA, acronym for **S**imulation for **H**igh-**E**nergy **R**eactions of **PA**rticles, is a new multipurpose event-generation framework. It is entirely written in the object-oriented programming language C++. In its current form, it is able to completely simulate electron/positron and fully hadronic collisions at high energies. SHERPAs most prominent feature is the consistent combination of matrix elements at the tree level with the parton shower. According to the algorithms proposed in [38, 39] (CKKW), this merging procedure leads to exact results to (next-to) leading logarithmic order. The key idea is to separate the phase space for parton emission into a hard region of jet production-accounted for by suitable matrix elements and the softer region of jet evolution, covered by the parton shower. Then, extra weights and vetoes are applied on the former and on the latter region, respectively, such that the overall dependence on the intersection cut is minimal. The SHERPA package is

constructed in a modular fashion, where each module encapsulates different aspects of event generation for high-energy particle reactions.

After the SHERPA package has been installed in a directory, each module is located in its own subdirectory of the same name in the `Run` subdirectory. The main steering module that initialises, controls and evaluates the different phases in the entire process of event generation is called `Sherpa`. In addition, all necessary routines for the combination of parton showers and matrix elements, which are independent of the specific parton shower are found in this module. Furthermore, this subpackage also provides an interface to the Lund String Fragmentation of PYTHIA including its hadron decay routines, which is currently used by SHERPA. Another important submodule is the `Model` package. It comprises the basic physics parameters (such as masses, mixing angles, etc.) of the simulation run. Thus it specifies the corresponding physics model. At the moment three different physics models are supported: the Standard Model (SM), its Minimal Supersymmetric extension (MSSM) [40] and the ADD model of large extra dimensions [7]. SHERPA's matrix element generator, which employs the method of helicity amplitudes [41, 42], is called `AMEGIC` [37]. It works as a generator, which generates generators: During the initialisation run the matrix elements for a set of given processes within the SM, the MSSM or the ADD model, as well as their specific phase space mappings are created by AMEGIC and stored in library files. In the initialisation of the production run, these libraries are linked to the program. They are used to calculate cross sections and to generate single weighted or unweighted events based on them. A complete description of the SHERPA modules can be found in [43].

2.2.3   SHERPA in CMS

SHERPA does not produce output in the HBOOK standard, but uses an internal event record. Thus, the strategy has been to use the HepMC interface of SHERPA and pass the events in this format to the CMS simulation software. The HepMC package [44] is an object oriented event record written in C++ for High Energy Physics Monte Carlo generators. Many extensions to HEPEVT, the Fortran HEP standard, were added: the number of entries is unlimited, spin density matrices can be stored with each vertex, flow patterns (such as colour) can be stored and traced, integers representing random number generator states can be stored as well, and an arbitrary number of event weights can be included. Particles and vertices are kept separate in a graph structure, physically similar to a physics event. The added information supports the modularisation of event generators. Event information is accessed by means of iterators supplied with the package.

In order to interface SHERPA to the CMS software, first SHERPA was downloaded (from $http://www.physik.tu-dresden.de/\ krauss$) and compiled; Hereby it is important to use the compiler flag `./configure - -enable - clhep`

and set the environment variable CLHEPDIR to the installation path of CLHEP. To run SHERPA one has to go to the SHERPA/Run directory and execute ./Sherpa PATH = `your_config_directory`, where the latter is the directory with all the control cards (∗.`dat`) required to steer the event generation for a given process. As explained above, in the first run the process specific libraries are generated, which has to be compiled in a second step by typing ./`makelibs` in `your_config_directory`. Finally, typing ./SHERPA PATH = `your_config_directory` in SHERPA/Run starts the cross-section calculation and the actual event generation. To activate the HepMC output the control flag `HEPMC_OUTPUT` = `something` has to be set in `Run.dat`; This is, however, not sufficient: the code to print the generated events to an ASCII file has to be explicitly added, for example in `Main.C`:
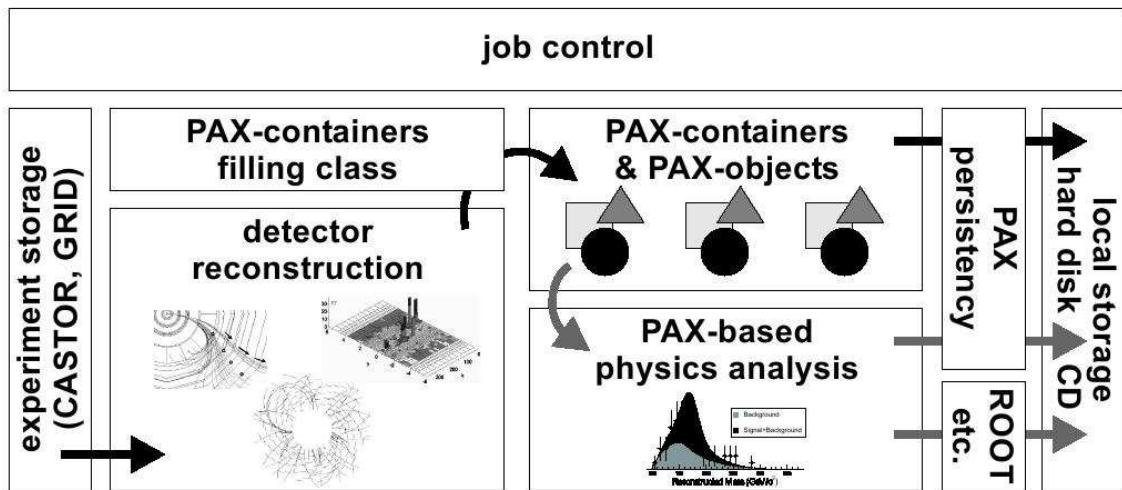
```
std::ofstream os( "SHERPAWriteHepMC.dat");
HepMC::GenEvent* evt =
Generator.GetIOHandler()->GetHepMCInterface()->GenEvent();
if (evt !=NULL) os << evt;
```

The events in the HepMC format are written to the specified output file (`SHERPAWriteHepMC.dat` in this case). The output file can then immediately be read in to OSCAR by the HepMC interface; During this work the relatively new OSCAR interface was systematically tested and improved, the HepMC interface to FAMOS was, since not available, completely newly developed and tested. To ensure the correctness of the HepMC interfaces, the same events were generated with PYTHIA and written out in (1) an HBOOK n-tuple and (2) as HepMC events. Both event formats were then processed through the complete simulation and reconstruction chain and delivered exactly the same reconstructed objects in the end, provided that all necessary random seeds were initialised with the same values.

In summary, SHERPA has been successfully set-up and interfaced to the CMS simulation chain. New HepMC interfaces were developed and included in the official CMS software; they have been carefully validated and proven to work, which opens as well the possibility to use any generator that supports HepMC output in CMS. In the new CMS software framework, CMSSW [45], HepMC was chosen by the CMS collaboration (as well as by the three other LHC experiments) as the default format to store the generator output. Starting from the interfaces in OSCAR and FAMOS, the HepMC interface to read an HepMC ASCII file and a converter to transform already produced CMKIN n-tuples into HepMC were developed as well during this thesis and contributed to CMSSW. They were of essential importance to test the detector simulation and reconstruction in the new framework and for validation purposes with respect to OSCAR and ORCA, since exactly the same input n-tuple could be used. In fact the first events ever simulated with CMSSW for test purposes using the HepMC interface has been SHERPA large extra dimensions events (see chapter 5). Both software packages (`IOMC/Input` and `IOMC/NtupleConverter`) are currently maintained by the author.

2.3 The PAX toolkit

Another software project which has been started in 2003 at the University of Karlsruhe and worked on during this thesis is the new Physics Analysis eXpert toolkit (PAX) [46]. PAX is a novel data analysis utility designed to assist physicists in the phase between detector reconstruction and physics interpretation of an event. It allows to define a level of abstraction beyond detector reconstruction by providing a general, persistent container model for HEP events. Working directly on the output of the detector reconstruction software when performing data analyses is an established habit amongst particle physicists. Nevertheless, having uniform access to all reconstructed objects (as muons, electrons, tracks, calorimeter energy depositions, jets etc.) by means of an abstract interface between detector reconstruction and physics analysis turned out to be advantageous during data analyses at LEP and HERA. The success of such a concept has been demonstrated earlier with the physics analysis packages of the ALEPH [47] and H1 [48] experiments: their use protected the physics analysis code against changes in the detector reconstruction and enabled users to comparably quickly answer physics questions.

Figure 2.2: Analysis flow with the PAX toolkit

2.3.1 Main design ideas of PAX

The design of the PAX toolkit has been developed according to the following guidelines:

1. The package is designed as a utility toolkit, i.e. the user keeps full control of every step in the program execution. The programming interface is kept simple and intuitive, minimising the need to access the manual and thereby increasing the users acceptance.

2. The package supports modular physics analysis structures and thus facilitates team work. The complexity of todays and future analyses makes efficient team work of many physicists mandatory. The physics analysis code can be used consistently among different high energy physics analyses and experiments. This results in modularity of physics analyses when consequently realised with the PAX toolkit.

The PAX toolkit provides a general, persistent event container model, in which physics objects like fourvectors, vertices and collisions can easily be stored, accessed and manipulated. Bookkeeping of relations between these objects (like decay trees, vertex and collision separation, including deep copies etc.) is fully provided by relation management tools. In addition, the event container and associated objects represent a uniform interface for algorithms and facilitate the parallel development and evaluation of different physics interpretations of individual events. An own persistency scheme, the PAX I/O (.pax) file, is also provided; the net profit here is an enormous gain of speed, because all the overhead from experiment specific software vanishes. Writing and reading PAX files is very fast, since the well approved ROOT I/O format can be used. The structure of a PAX based analysis is schematically shown in Figure 2.2.

2.3.2   PAX class structure

The basic unit in PAX is an event container representing a special view of the event, called event interpretation. This container is used to store the relevant event information in terms of collisions, vertices, fourvectors, their relations, and additional values needed in the users analysis. When the user finally deletes an instance of an event interpretation, instances of objects which have been registered with it (collisions, vertices, fourvectors, etc.) are also removed from memory. The fourvectors have the basic kinematic characteristics which they inherit from the CLHEP [49] or ROOT-packages [50], according to the users choice. In this way, they provide all capabilities contained in these well known libraries. Moreover, the fourvectors of PAX have been designed to enable direct access to all possible information in the experiments reconstruction output. A few member functions in addition to the basic kinematics have been selected to be available by default from previous experience at HERA, e.g., charge, particle-ID, etc. In addition, pointers to an arbitrary number of instances of experiment classes can be registered with PAX-objects and accessed during analysis. A dedicated, experiment specific class for filling the PAX-containers represents the interface the between detector reconstruction software and the PAX-based physics analysis. Once all relevant information is filled, the PAX-objects can be stored persistently to PAX-files for later use, and the analysis code is called. Analysis results can be managed with help of the ROOT-package, selected event interpretations additionally can be stored to PAX-files. With an analysis consistently

framed in PAX-objects, the filling class can be exchanged easily, and the identical physics analysis can be applied for instance directly to the output of a Monte Carlo event generator or a fast simulation software. Furthermore, the use of PAX-files, which provide the distilled experimental event information, allows fast analysis of the reconstructed data decoupled from the experiment-specific software and data storage environment.
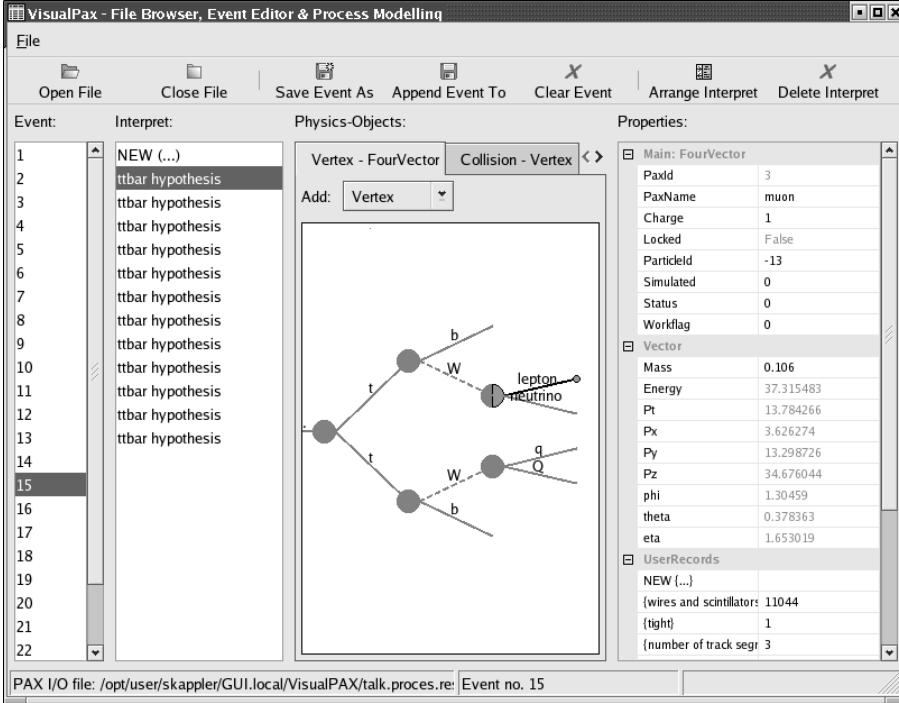
To sum up: the main purpose of PAX is to assist in the physics analysis stage of a particle physics research project. It is neither a detector reconstruction tool nor a substitute for visualisation or histogramming. It rather should be considered as a supportive tool in the step from the database to physics plot. It has the potential to add a certain amount of unification and simplification to the physics analysis that very often is a sort of black box. More details with documentation and tutorials can be found at the webpage and in [51, 52, 53].

2.3.3 PAX in the CMS software framework

During 2004, the PAX toolkit has been integrated into the CMS software framework. The PAX kernel is a continuously maintained external package in the CMS software environment and example source codes, i.e. filling classes for the CMS reconstruction software ORCA and an extensive example analysis for reconstruction, are provided and regularly revised within the CMS software repository. During this thesis the PAX toolkit has been partially used for the physics analysis in chapter 5. One use case has been the generator level comparisons, which were performed with PAX. PAX is well-suited for this kind of comparisons, since it comes with a standard CMS n-tuple interface, which fills the generator level information from the n-tuple into the `PaxEventInterpret`, a central container for PAX based analysis. This interface has been further validated and improved during this work. Furthermore, since SHERPA provides output in the HepMC format, as discussed in the previous section, a new `PAXHepMC` interface has been developed, contributed to the project and is maintained by the author. Since HepMC has a similar graph-like structure to save the particle in the event as PAX, the conversion to PAX is straightforward. Furthermore, a converter has been developed in the scope of a diploma thesis related to the analysis presented here [54], which converts the ROOT files obtained with the CMS analysis package `ExRootAnalysis` [55] to PAX files. This is realised by saving the reconstructed objects as PaxFourVectors. Then, the actual PAX based analysis can be quickly and consistently run on PAX objects.

2.3.4 Latest developments

A recently developed accessory to the PAX kernel is Visual PAX (VPAX), that allows for browsing of PAX I/O files and editing of `PaxEventInterpret` instances in a Graphical User-Interface (GUI). VisualPax is based on the wxWidgets [56]

Figure 2.3: The visual PAX graphical interface.

package. VisualPax allows to graphically display and modify event interpretations including properties and decay chains of the contained physics objects. It can assist the physicist in the development of the analysis or be used for fast visual cross checks. The VPAX GUI is shown in Figure 2.3.

2.4   CMS software installation with XCMSI

A last project which has been in the centre of attention during this thesis is XCSMI [57]. For data analysis in an international collaboration it is important to have an efficient procedure to distribute, install and update the centrally maintained software. This is even more true when not only local, but also GRID accessible resources are to be exploited. In order to achieve this aim, the development of a practical solution was started, that has been successfully employed for CMS software installations on systems ranging from physicists' notebooks up to LCG2 (LHC Computing Grid, see [58]) enabled clusters. As a member of the XCMSI team, in particular the XCMSI GUI has been developed, the installation and validation procedure tested and improved and user support provided. The product XCMSI has become an official CMS tool, was included in the CMS Physics Technical Design Report volume 1 [6] and is widely used within the CMS collaboration. A description of its central ideas and main features can be found in Appendix A. More details about the project can as well be found in [59, 60, 61].

Chapter 3

Shower Parameterisation in Geant4 and CMS

The CMS detector simulation is based on Geant4 [24, 25]. Geant4 simulates detector effects on physics events using a detailed microscopic description of the interactions between particles and matter. It is very accurate, but it can be very time consuming. In particular the simulation of electro-magnetic cascades in calorimeters is expected to account for a considerable amount of the total simulation time. Since it increases almost linearly with the energy absorbed in the detector, fully simulating one event using individual particle tracking at LHC energies may take several minutes. When a large number of simulated events are needed for physics analyses, a full simulation approach may simply be infeasible due to the computing time costs and, depending on the concrete use case, also unnecessary. For example, in calorimeter simulation different tasks can be considered: calorimeter studies, physics analysis, and feasibility studies. A detailed simulation, where all secondary particles are tracked individually down to some minimum energy is required for accurate calorimeter studies. For physics analysis and feasibility studies on the other hand a large number of Monte Carlo events may have to be produced.

Using parameterisations for electromagnetic (sub)showers can speed up the simulations considerably, without sacrificing significantly precision. The high particle multiplicity in electromagnetic showers as well as their compactness and the good understanding of the underlying physics makes their parameterisation advantageous. In the past within CMS, the stand-alone fast simulation FAMOS [27] was developed as an alternative to full simulation. However, this approach may be not be sufficiently accurate for many analyses, for example for studying pathological events that may look like a discovery. The running of past HEP experiments like CDF [62], and H1 [48] proved the importance of an intermediate level of simulation, faster than full simulation, but integrated in the same framework so that it could provide the same kind of output as full simulation on which the full reconstruction chain could be run. A set of equations, derived from the H1 parameterisation, can be used

to parameterise the electromagnetic shower development in different calorimeters. These equations were originally implemented in the GEANT3 framework [63]. The goal of this part of the thesis was to integrate this parameterisation concept to the C++ based framework of the Geant4 toolkit and the full CMS detector simulation, OSCAR [23]. In this chapter the procedure to achieve this aim is described.

First, a summary of the physics processes that give rise to electromagnetic showers is given and a simple shower model is introduced. Then, in more detail, the parameterisation Ansatz used in this work to model electromagnetic shower is discussed, followed by a brief introduction to calorimetry. The next part concentrates on technical details of the parameterisation implementation inside the Geant4 framework and the CMS simulation program OSCAR. In the last section, the benefits of using this new shower parameterisation package are presented in terms of timing and physics performance for a simple geometry and within the full CMS detector simulation.

3.1   Electromagnetic shower

When a high energetic charged particle travels through matter, it can start an electromagnetic shower by emitting bremsstrahlung photons. Bremsstrahlung is the radiation of a real photon in the Coulomb field of the nucleus of the absorber medium. The bremsstrahlung photons can produce e^+e^- pairs (pair production) with lower energy which can themselves again radiate and, in consequence, a particle cascade is created. In most cases electromagnetic showers are initiated by electrons, positions or photons. A more quantitative description of the processes that takes place inside an electromagnetic shower is given in the following.

3.1.1   Energy loss by electrons and positrons
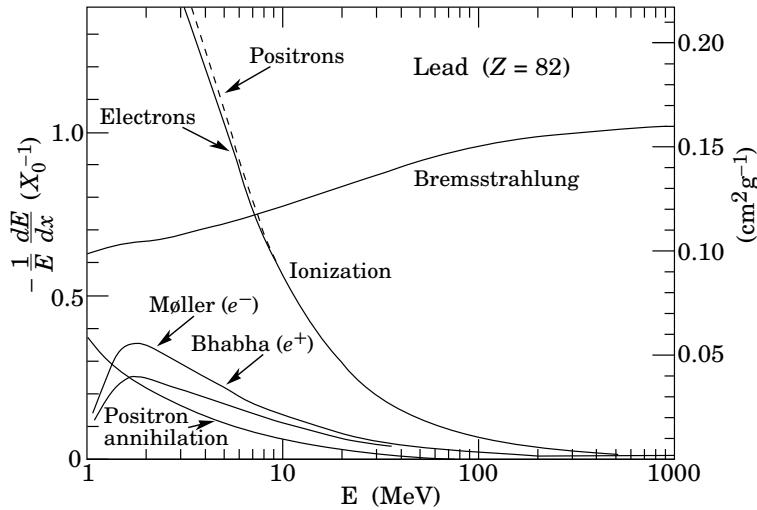
Electrons and positrons have identical electromagnetic interactions in matter (but an opposite sign) and lose energy mostly by two processes:

- Ionisation

- Radiation

In the low energy region primarily ionisation contributes, while for energies above around 100 MeV the radiation of bremsstrahlung becomes the dominant process. This can be seen in Figure 3.1, which shows the relative energy loss $-\frac{1}{E}\frac{dE}{dx}$ of electrons and positrons.

Ionisation

Generally, nuclear particles traversing a medium continuously transfer energy to the constituent atoms of that medium via the ionisation process (formation of ion pairs)

Figure 3.1: Relative energy loss per radiation length in lead as a function of the electron or positron energy. For energies above around 100 MeV the (shower initiating) radiation of bremsstrahlung in the field of the atoms is the dominant process [64].

or excitation of the atoms. In the case of electrons and positrons, ionisation is defined as the scattering of electrons or positrons with atomic electrons and an energy transfer smaller than 0.255 MeV per collision. Collisions with a higher energy transfer are called Möller-(Bhabha-) scattering and are less important. Higher energetic electrons can themselves produce ion-electron pairs while traversing the medium (secondary ionisation). The mean rate of energy loss dE/dx (or stopping power) for moderately relativistic charged particles other than the electron is given by the Bethe-Bloch equation (for integer spin) [64]:

$$ - \frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2 m_e c^2 \gamma^2 \beta^2 T_{max}}{I^2} \right) - \beta^2 - \frac{\delta}{2} \right] , \qquad (3.1) $$

with

N_A	:	Avogadro's number in $1/mol$
r_e	:	Classical electron radius in cm
m_e	:	Electron mass in MeV
z	:	Charge of incident particle
Z	:	Atomic number of absorber
A	:	Atomic number of absorber g/mol
β	:	$\frac{v}{c}$, i.e. velocity measured in c units
γ	:	$1/\sqrt{1 - \beta^2}$
I	:	Mean excitation energy in MeV
δ	:	Density effect correction to ionisation energy loss
T_{max}	:	maximum kinetic energy which can be imparted to a free electron in a single collision

Ionisation loss by electrons and positron differs from the loss by heavy particles because of the kinematics, spin, and the identity of the incident electron with the electrons which it ionises. In analogy to the Bethe-Bloch Equation 3.1 for heavy particle the ionisation energy loss of electrons is for example described by [65]. The energy of electrons created by an ionisation process is so low that the probability to initiate a shower is very small.

Bremsstrahlung

If a charged particle is decelerated in the Coulomb field of a nucleus a fraction of its kinetic energy will be emitted in the form of real photons. The radiation produced by charge particles passing through a medium is known as Bremsstrahlung (the German word for braking radiation). In particular, the term "external bremsstrahlung" is used for radiation caused by decelerations when passing through the field of atomic nuclei. The term "internall bremsstrahlung" is used to describe the radiation of non-virtual quanta, i.e. photons or gluons, by particles during an interaction. Radiation emitted by a charged particle moving in a magnetic field is called synchrotron radiation. The energy loss by bremsstrahlung for high energies can be described by [66]:

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = 4\alpha N_\mathrm{A}\frac{Z^2}{A} \cdot z^2 r^2 E \ln\frac{183}{Z^{1/3}} \quad , \tag{3.2}$$

where $r = \frac{1}{4\pi\varepsilon_0} \cdot \frac{e^2}{mc^2}$. Equation 3.2 can be rewritten for electrons as

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{E}{X_0} \quad \implies E = E_0 e^{-x/X_0} \tag{3.3}$$

where

$$X_0 = \frac{A}{4\alpha N_\mathrm{A} Z(Z+1)r_e^2 \ln(183\,Z^{-1/3})} \tag{3.4}$$

is called the radiation length. One radiation length is the length after which an electron loses $(1 - 1/e)$ of its energy by bremsstrahlung, i.e. after it has $\approx 37\%$ of its initial energy.

In order to describe the longitudinal shower development independent of the material, the longitudinal coordinate x is often measured in units of X_0 - the material dependence is already included there: [1]

$$\boxed{t := \frac{x}{X_0}} \tag{3.5}$$

According to [64] one can approximate X_0 for $Z > 2$ on a 2.5% level by: [2]

$$\boxed{X_0 = \frac{716.4\,A}{Z(Z+1)\ln(287/\sqrt{Z})} \quad [\mathrm{g/cm^2}]} \tag{3.6}$$

---

[1]Definitions in boxes are used for the shower parameterisation.

[2] To get X_0 as a length in cm one has to divide through the density of the material.

For a mixture or compound materials this becomes:

$$\frac{1}{X_0} = \sum w_j X_j \tag{3.7}$$

where w_j and X_j are the fraction by weight and the radiation length for the j$^{\text{th}}$ element. Bremsstrahlung dominates the energy loss of electrons above the critical energy. The critical energy (E_c) of an electron is the energy at which the main energy loss mechanism changes from radiation losses to ionisation losses:

$$\left.\frac{\mathrm{d}E_{\mathrm{c}}}{\mathrm{d}x}\right|_{\text{ion}} = \left.\frac{\mathrm{d}E_{\mathrm{c}}}{\mathrm{d}x}\right|_{\text{brems}} \tag{3.8}$$

For a solid or liquid absorber the critical energy can be approximated by: [66]

$$E_{\mathrm{c}} = \frac{610\,\text{MeV}}{Z + 1.24} \tag{3.9}$$

For example, the critical energy for electrons in **lead** is about 7.6 MeV. For heavy particles, bremsstrahlung plays a role at much higher energies than for electrons. At high energies also radiation from heavier particles becomes important and consequently a critical energy for these particles can be defined. Since

$$\left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\text{brems}} \propto \frac{1}{m^2} \tag{3.10}$$

the critical energy e.g. for muons in **iron** is

$$E_{\mathrm{c}} = \frac{610\,\text{MeV}}{Z + 1.24} \cdot \left(\frac{m_\mu}{m_e}\right)^2 = 960\,\text{GeV} \tag{3.11}$$

As a consequence already the muon, the "heavy brother" of the electron, does not produce a shower in matter (below 1 TeV), since its bremsstrahlung is surpressed by the factor:

$$\left(\frac{m_e}{m_\mu}\right)^2 \approx 10^{-5} \tag{3.12}$$

In the following, the formula chosen by Grindhammer and Peters for the critical energy E_c is used [67].

$$\boxed{E_c = 2.66 \cdot \left(X_0 \frac{Z}{A}\right)^{1.1} \text{MeV}} \tag{3.13}$$

3.1.2   Energy loss of photons

There are three main processes by which photons interact with matter, namely:

- Photoelectric effect

- Compton effect

- Pair production

The intensity of a photon beam varies in matter according to

$$I = I_0 \, e^{-\mu x} \quad , \tag{3.14}$$

where μ is the mass attenuation coefficient. μ is related to the photon cross sections of each of these processes σ_i by

$$\mu = \frac{N_A}{A} \sum_{i=1}^{3} \sigma_i \tag{3.15}$$

The photoelectric effect has a significant contribution in the low energy region ($E_{gamma} \sim$ keV), while in the intermediate energy region (MeV) Compton scattering dominates. For high energetic photons (MeV/GeV region) pair production becomes the relevant proccess. As an example, the cross-sections for the different processes in carbon and lead are shown in Figure 3.1.2.

Photoelectric effect

The photoelectric effect can be considered as an interaction between the photon and the atom as a whole. Incident photons whose binding energy exceeds the binding energy of an electron may be absorbed and consequently the atomic electron may be emitted.

$$\gamma + \text{atom} \rightarrow \text{atom}^+ + e^- \tag{3.16}$$

The cross-section for the absorption of a photon of the energy E_γ is particularly large in the K-shell (80% of the total cross section) and can be calculated for high energies as
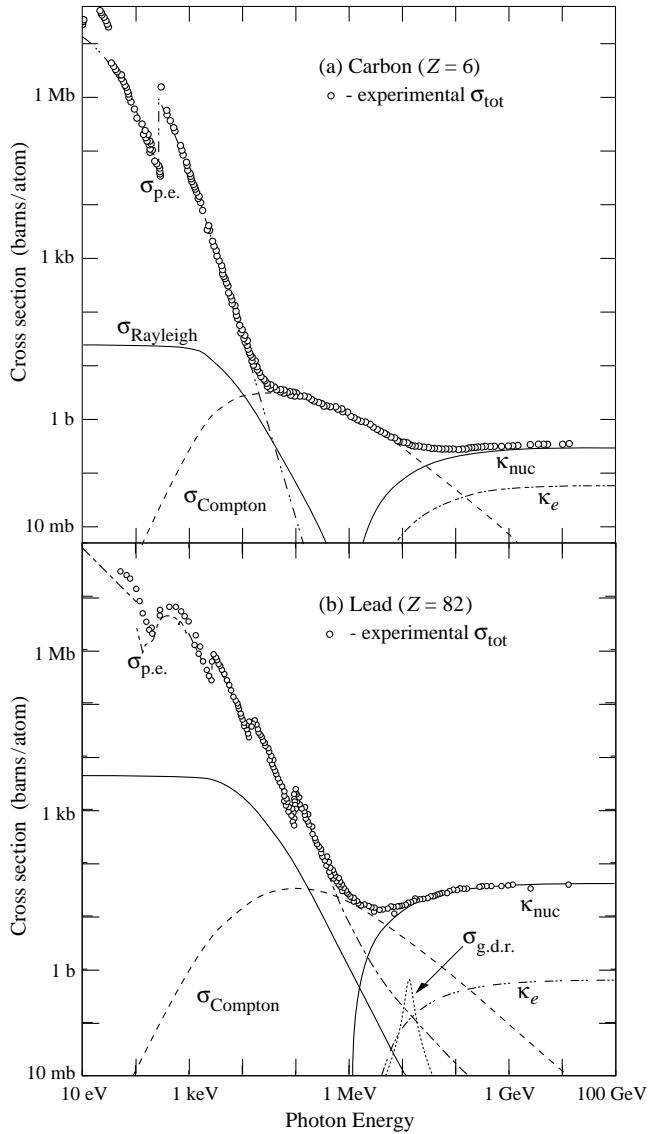
$$\sigma_{\text{Photo}}^K = 4\pi r_e^2 Z^5 \alpha^4 \cdot \frac{1}{\varepsilon} \tag{3.17}$$

(with $\varepsilon = E_\gamma/m_e c^2$, Z being the number of electrons in the target atom). It shows a strong Z^5 dependence. The photoelectric cross section has sharp discontinuities when E_γ coincides with the binding energy of atomic shells. As a consequence of a photoabsorption in the K-shell, characteristic X-rays or Auger electrons are emitted [66].

Compton scattering

The Compton effect describes the scattering of photons with quasi-free atomic electrons

$$\gamma + e \rightarrow \gamma' + e' \quad . \tag{3.18}$$

Figure 3.2: Contributions to the photon cross-section of different interaction processes in light elements (carbon) and a heavy element (lead). At low energies it can be seen that the photoelectric effect dominates, although Compton scattering, Rayleigh scattering and photonuclear absorption also contribute. The photoelectric cross-section is characterised by discontinuities (absorption edges) as thresholds for photon ionisation of various atomic levels are reached [64]. Definitions: τ: Atomic photoelectric effect, σ_{COH}: Rayleigh Coherent scattering (atom neither ionised nor excited), σ_{INCOH} : Compton incoherent scattering, κ_n: Pair production, nuclear field, κ_e: Pair production, electron field.

which results in an increase of the wavelength of the photon. The cross-section for this process, given by the Klein–Nishina formula, can be approximated at high energies by

$$\sigma_c = \pi r_e^2 \frac{Z}{\varepsilon}(\frac{1}{2} + \ln(2\varepsilon)) \tag{3.19}$$

which is $\propto \frac{\ln\varepsilon}{\varepsilon}Z$.

From energy and momentum conservation one can derive the ratio of scattered (E_γ') to incident photon energy (E_γ)

$$\frac{E_\gamma'}{E_\gamma} = \frac{1}{1 + \varepsilon(1 - \cos\Theta_\gamma)} \quad , \tag{3.20}$$

where Θ_γ is the scattering angle of the photon with respect to its original direction. The energy transfer to the electron E_{kin} reaches a maximum value for backscattering ($\Theta_\gamma = \pi$)

$$E_{\text{kin}}^{\text{max}} = \frac{2\varepsilon^2}{1 + 2\varepsilon}m_e c^2 \quad , \tag{3.21}$$

which, in the extreme case ($\varepsilon \gg 1$), equals E_γ. At accelerators and in astrophysics also the process of inverse Compton scattering is of importance.

Pair production

For high photon energies $E_{gamma} \geq$ GeV the production of an electron–positron pair in the Coulomb field of a nucleus becomes the dominant process. It requires a certain minimum energy

$$E_\gamma \geq 2m_e c^2 + \frac{2m_e^2 c^2}{m_{\text{nucleus}}} \quad . \tag{3.22}$$

Since for all practical cases $m_{\text{nucleus}} \gg m_e$, one has effectively $E_\gamma \geq 2m_e c^2$.

The total cross section at reasonably high energies ($E_\gamma \gg 20\,\text{MeV}$), is

$$\sigma_{\text{pair}} = \frac{7}{9}\frac{A}{N_A} \cdot \frac{1}{X_0} \quad , \tag{3.23}$$

X_0 being the interaction length for the material defined in Equation 3.6. More explicitly this means that after having crossed a material block with thinkness $\frac{9}{7}X_0$ there are $\frac{1}{e}$ ($\sim 36\%$) photons left, the rest being converted into e^+e^- pairs. Or, in other words, that one X_0 is $\frac{7}{9}$ of the mean free path for pair production by a high-energy photon. The partition of the energy to the electron and positron is symmetric at low energies ($E_\gamma \ll 50\,\text{MeV}$) and increasingly asymmetric at high energies ($E_\gamma > 1\,\text{GeV}$) [66].

3.1.3 Simple shower models

The development of cascades induced by electrons, positrons or photons is governed by bremsstrahlung of electrons and pair production of photons. Secondary particle production continues until the energies of photons fall below the pair production threshold and energy losses of electrons other than bremsstrahlung start to dominate: the number of shower particles decays exponentially.

Already a very simple model can describe the main features of particle multiplication in electromagnetic cascades: a photon of energy E_0 starts the cascade by producing an e^+e^- pair after one radiation length. Assuming that the energy is shared symmetrically between the particles at each multiplication step, one gets at the depth t

$$N(t) = 2^t \tag{3.24}$$

particles with energy

$$E(t) = E_0 \cdot 2^{-t} \quad . \tag{3.25}$$

The multiplication continues until the electrons fall below the critical energy E_c

$$E_c = E_0 \cdot 2^{-t_{\max}} \quad . \tag{3.26}$$

Thereafter $(t > t_{\max})$ the shower particles are only absorbed. The position of the shower maximum is obtained from (Equation 3.26)

$$t_{\max} = \frac{\ln E_0/E_c}{\ln 2} \propto \ln E_0 \quad . \tag{3.27}$$

The total number of shower particles is

$$\begin{aligned}
S &= \sum_{t=0}^{t_{\max}} N(t) = \sum 2^t = 2^{t_{\max}+1} - 1 \approx 2^{t_{\max}+1} \\
&= 2 \cdot 2^{t_{\max}} = 2 \cdot \frac{E_0}{E_c} \propto E_0 \quad .
\end{aligned} \tag{3.28}$$

If the shower particles are sampled in steps t measured in units of X_0, the total track length obtained is

$$S^* = \frac{S}{t} = 2\frac{E_0}{E_c} \cdot \frac{1}{t} \quad , \tag{3.29}$$

which leads to an energy resolution of

$$\frac{\sigma}{E_0} = \frac{\sqrt{S^*}}{S^*} = \frac{\sqrt{t}}{\sqrt{2E_0/E_c}} \propto \frac{\sqrt{t}}{\sqrt{E_0}} \quad . \tag{3.30}$$

In a more realistic description the longitudinal development of the electron shower can be approximated by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathrm{const} \cdot t^a \cdot \mathrm{e}^{-bt} \quad , \tag{3.31}$$

where a, b are fit parameters.

In addition to the longitudinal shower development, there is a lateral spread of the shower, since the particles are not only slowed down but also their original direction is changed. This is mostly due to :

- the multiple scattering of electrons

- the opening angle of e^+ e^- pairs produced by photon conversion

- the p_T of Bremsstrahlung quanta

According to the theory of multiple scattering [68] the average scattering angle of an electron in the Coulomb field of a nucleus is:

$$\langle \Theta \rangle_{ms} \approx \frac{E_s}{E} \cdot \sqrt{t} \tag{3.32}$$

with

$$E_s = \sqrt{\frac{4\pi}{\alpha}} m_e \approx 21.2 MeV \ , \tag{3.33}$$

α being the fine structure constant, m_e the electron mass. For the opening angle of e^+e^- -pairs, as well as for the angle between the electron and its radiated Bremsstrahlung photon, one can approximately write: [69]

$$\left\langle \Theta_{\measuredangle(e^+,e^-)} \right\rangle \approx \frac{m_e}{E_\gamma} , \left\langle \Theta_{\measuredangle(\gamma,e)} \right\rangle \approx \frac{m_e}{E_e} \Rightarrow \frac{\langle \Theta \rangle}{\langle \Theta \rangle_{ms}} = \frac{m_e}{E_s \sqrt{t}} \approx \frac{2.4 \cdot 10^{-2}}{\sqrt{t}} \tag{3.34}$$

As can be seen above in Equation 3.34 for $\langle \Theta \rangle$ these two processes do not dominate the lateral shower spread and not long after the shower start ($t > 1$) multiple scattering becomes the crucial process for the radial shower development. This is especially true in the case of low energy electrons (since $\langle \Theta \rangle_{ms} \sim 1/E$), which are particularly present near the shower maximum.

The lateral spread of the shower is often measured in units of the Molière radius:

$$\boxed{R_{\mathrm{M}} = \frac{21 \, \mathrm{MeV}}{E_{\mathrm{c}}} X_0 \, [\mathrm{g/cm^2}]} \tag{3.35}$$

In a homogeneous material, 90% of the shower energy is contained in a cylinder of radius R_{M} around the shower axis, in the case of a $3 R_{\mathrm{M}}$ radius 98% are located in the cylinder.

While electromagnetic cascades are initiated by charged particles, charged hadrons can also initiate a hadronic shower. As explained above, the longitudinal development of electromagnetic cascades can be described in terms of the radiation length, X_0, and their lateral width is determined by multiple scattering. In contrast to this hadronic showers are governed in their longitudinal structure by the

nuclear interaction length, λ, and by the transverse momenta of the secondary particles as far as lateral width is concerned. Since for most materials $\lambda \gg X_0$, and $\langle p_{\mathrm{T}}^{\mathrm{interaction}} \rangle \gg \langle p_{\mathrm{T}}^{\mathrm{multiple\ scattering}} \rangle$, hadronic showers are longer and wider. They also may contain several electromagnetic sub shower. A detailed description of shower parameterisation techniques for hadronic showers can be found in [70].

3.2   Shower parameterisation

3.2.1   Parameterisation Ansatz

The spatial energy distribution of electromagnetic showers is given by three probability density functions (PDF):

$$dE(\vec{r}) = E\, f(t)dt\, f(r)dr\, f(\phi)d\phi, \tag{3.36}$$

describing the longitudinal, radial, and azimuthal energy distributions. Here t denotes the longitudinal shower depth in units of radiation length, r measures the radial distance from the shower axis in Molière units and ϕ is the azimuthal angle. A gamma distribution is used for the parameterisation of the longitudinal shower profile, f(t). The radial distribution, f(r), is described by a two-component Ansatz. In ϕ, it is assumed that the energy is distributed uniformly due to symmetry: $f(\phi) = 1/2\pi$. .

3.2.2   Longitudinal shower profiles in homogeneous media

It is well known that average longitudinal shower profiles can be described by a gamma distribution (with the gamma function $\Gamma(\alpha)$):

$$\left\langle \frac{1}{E}\frac{dE(t)}{dt} \right\rangle = f(t) = \frac{(\beta t)^{\alpha-1}\beta \exp(-\beta t)}{\Gamma(\alpha)} \tag{3.37}$$

The centre of gravity, $\langle t \rangle$, and the depth of the maximum, T, can be calculated from the shape parameter, α, and the scaling parameter, β, according to

$$\langle t \rangle = \frac{\alpha}{\beta} \tag{3.38}$$

$$T = \frac{\alpha - 1}{\beta} \tag{3.39}$$

According to analytical studies by Rossi [71] the longitudinal shower moments are equal in different materials, provided one measures all lengths in units of radiation length (X_0) (see Equation 3.6) and energies in units of the critical energy (E_c) (see Equation 3.13). For Example, for the depth of the shower maximum, Rossi predicts:

$$T \propto \ln y = \ln \frac{E}{E_c} \tag{3.40}$$

It is therefore desirable to use T in the parameterisation together with the second variable α:

$$T = \ln y + t_1 \tag{3.41}$$

$$\alpha = a_1 + (a_2 + a_3/Z)\ln y \tag{3.42}$$

Assuming that also individual profiles can be approximated by a gamma distribution, T and α are obtained from fitting each single Geant4-simulated shower with Equation 3.37. The values of the coefficients used for the parameterisation are given in the Appendix B, where all formulas and numbers are summarised. The strategy has been here to follow as close as possible the numbers found by Peters and Grindhammer [67], provided that the comparison with fully simulated showers does not show unacceptable discrepancies. For the parameterisation of $\langle \ln T \rangle$ and $\langle \ln \alpha \rangle$ the logarithms of Equation 3.41 and Equation 3.42 are used, since they are found to be approximately normal distributed. The y-dependence of the fluctuations can be described by:

$$\sigma = (s_1 + s_2 \ln y)^{-1} \tag{3.43}$$

The correlation between $\ln T$ and $\ln \alpha$ as function of the energy is given by :

$$\rho(\ln T, \ln \alpha) \equiv \rho = r_1 + r_2 \ln y \tag{3.44}$$

From these formulae, correlated and varying parameters α_i and β_i are generated according to:

$$\begin{pmatrix} \ln T_i \\ \ln \alpha_i \end{pmatrix} = \begin{pmatrix} \langle \ln T \rangle \\ \langle \ln \alpha \rangle \end{pmatrix} + C \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{3.45}$$

with

$$C = \begin{pmatrix} \sigma(\ln T) & 0 \\ 0 & \sigma(\ln \alpha) \end{pmatrix} \begin{pmatrix} \sqrt{\frac{1+\rho}{2}} & \sqrt{\frac{1-\rho}{2}} \\ \sqrt{\frac{1+\rho}{2}} & -\sqrt{\frac{1-\rho}{2}} \end{pmatrix}$$

and $\beta_i = (\alpha_i - 1)/T_i$. z_1 and z_2 are standard normally distributed random numbers. The longitudinal energy distribution is evaluated by integration in steps of $\Delta t = t_j - t_{j-1} = 1X_0$,

$$dE(t) = E \int_{t_{j-1}}^{t_j} \frac{(\beta_i t)^{\alpha_i - 1} \beta_i \exp(-\beta_i t)}{\Gamma(\alpha_i)} dt$$

3.2.3 Radial shower profiles in homogeneous media

For average radial energy profiles,

$$f(r) = \frac{1}{dE(t)} \frac{dE(t, r)}{dr}, \tag{3.46}$$

a variety of different functions can be found in the literature [72, 73, 74, 75]. In the following a two component Ansatz is used, in which the radial profile is divided in a core $(r < 1\,R_M)$ and tail $(r \gtrsim R_M)$ component [67]:

$$
\begin{aligned}
f(r) &= p f_C(r) + (1 - p) f_T(r) \\
&= p \frac{2r R_C^2}{(r^2 + R_C^2)^2} + (1 - p) \frac{2r R_T^2}{(r^2 + R_T^2)^2}
\end{aligned}
\tag{3.47}
$$

with

$$0 \leq p \leq 1.$$

R_C and R_T describe respectively the median of the core and the tail component of the radial profile while p is the relative weight of the core component. The radial profile has a distinct maximum in the core component ($r \lesssim 1R_M$), which vanishes with increasing shower depth. In the tail component ($r \gtrsim 1R_M$), the distribution looks nearly flat at the beginning of the shower ($1 - 2X_0$), becomes steeper at moderate depths ($5 - 6X_0$, $13 - 14X_0$) and becomes flat again ($22 - 23X_0$). The variable $\tau = t/T$ is defined to measure the shower depth in units of the depth of the shower maximum, in order to generalise the radial profiles. This makes the parameterisation more convenient and separates the energy and material dependence of various parameters. The median of the core distribution, R_C, increases linearly with τ. The weight of the core, p, is maximal around the shower maximum and the width of the tail, R_T, is minimal at $\tau \approx 1$. The following formulae are used to parameterise the radial energy density distribution for a given energy and material:

$$R_{C,hom}(\tau) = z_1 + z_2\tau \tag{3.48}$$

$$R_{T,hom}(\tau) = k_1\{\exp(k_3(\tau - k_2)) + \exp(k_4(\tau - k_2))\} \tag{3.49}$$

$$p_{hom}(\tau) = p_1 \exp\left\{ \frac{p_2 - \tau}{p_3} - \exp\left(\frac{p_2 - \tau}{p_3} \right) \right\} \tag{3.50}$$

The parameters $z_1 \cdots p_3$ are either constant or simple functions of $\ln E$ or Z (see Appendix B for details). The complicated evolution of R_T and p with the shower depth and the dependence on the material can be explained mainly with the production of low energetic photons. Radial shape fluctuations have to be considered with some care. Even if no fluctuations of $f(r)$ are simulated explicitly, the radial energy profile at a given shower depth will fluctuate, because the shower maximum T and thus τ varies from shower to shower. The energy content of a longitudinal interval of length X_0, $dE(t)$, is calculated from the actual longitudinal energy density distribution as described in the previous section. This energy is divided into $N_S(t)$ discrete spots of energy $E_S = dE(t)/N_S(t)$, which are distributed radially according to $f(r)$ using a Monte Carlo method. This can be done easily since the PDFs, $f_C(r)$ and $f_T(r)$, can be integrated and inverted:

$$F(r) = \int_0^r \frac{2r'R^2}{(r'^2 + R^2)^2}dr' = \frac{r^2}{r^2 + R^2} \tag{3.51}$$

$$F^{-1}(u) = R\sqrt{\frac{u}{1 - u}}. \tag{3.52}$$

Random radii are generated according to $f(r)$ using two normally distributed random numbers v_i and w_i:

$$r_i = \begin{cases} R_C\sqrt{\frac{v_i}{1-v_i}}, & \text{if } p < w_i \\ R_T\sqrt{\frac{v_i}{1-v_i}}, & \text{otherwise.} \end{cases}$$

More details on the procedure can be found in [67] and [76].

A similar parameterisation can also be set up for sampling calorimeters consisting of a complicated but repetitive sampling structure, which are usually described by one single effective medium. The inhomogeneous material of sampling calorimeters influences the shower shapes. The shower maximum occurs earlier than in a homogeneous calorimeter with the same effective material properties. The sampling fluctuations, the scaling of the deposited energy to the visible energy using an appropriate sampling fraction, and the effects of the sampling structure have to be considered in parameterised simulations explicitly. The parameterisation of the longitudinal shape as given in subsection 3.2.2 for homogeneous media can therefore not be used for sampling calorimeters directly. For example, in Equation 3.41 and Equation 3.42 $\ln(T)$ and $\ln(\alpha)$ parametrise also as a function of the sampling frequency and the value of e/MIP averaged over the whole shower. Instead, the parameterisation for homogeneous media may be understood as a first approximation to which geometry dependent corrections have to be added for a sampling calorimeter. Following the procedure and functional forms of [67] and for the radial profile [75] a similar parameterisation procedure has been successfully implemented for the ATLAS liquid argon sampling calorimeter within the ATLAS Athena simulation framework, which is also Geant 4 based. Details on this work can be found in ATLAS documents and in [5].

3.3   Electromagnetic calorimetry

A calorimeter is a detector which uses the total absorption of particles to measure the energy of high energetic particles. In the process of absorption showers are generated by cascades of interactions and eventually most of the incident particle energy is converted into "heat", which explains the name calorimeter (calor = Latin for heat). For example, the total stored beam energy in the LHC beam is around 1×10^8 J, which would be sufficient to boil ($\Delta T = 100K$) 239 kg of water. If one considers the effect of a 1GeV single particle on one litre of water, however, the heating would be only $3.8 10^{-14}$ K, which is almost unmeasurable.

In reality no temperature is measured in detectors, but characteristic interactions with matter (e.g. atomic excitation, ionisation) are used to generate a detectable effect, that should be proportional to the energy of the incident particle. In fact, calorimetry is also the only practicable way to measure neutral hadrons among the secondaries produced in a high-energy collision.

3.3.1   Calorimeter types

Showers can be divided, according to the kind of interaction which produces the particle cascade, into *electromagnetic* and *hadronic* shower components. Therefore

there are typically two types of calorimeters: *electromagnetic* and *hadronic* calorimeter. [3] Electromagnetic calorimeters measure the energy of electrons and photons. They also assist in particle identification (specifically electron/charged-pion separation in conjunction with the tracker) and help to measure the energy of high energy hadrons. The electrons/positrons in the shower may produce either ionisation or light (or both), depending on the material in which the shower occurs. The light may be either scintillation (as in CsI and PbWO4, for example) or Cerenkov. It is then passed (via light-guides etc.) to photodetectors such as photomultiplier tubes (PMTs) or silicon photo-diodes.

In high energy physics experiments a variety of different shower media have been used depending upon the particular application and criteria like resolution, cost, space requirements etc. From the point of view of physics, the characteristics which govern the choice are as follows:

1. Radiation length, (X_0), the scale for the longitudinal distance of the shower (see Equation 3.6); About 25 X_0 of material is required in order to contain longitudinally 99 % of the shower.

2. Moliere radius, (R_M), the scale for the transverse spread of an electromagnetic shower.

3. The amount of detected light per unit of deposited energy N_{pe}.

4. The wavelength(s) of light emitted in the shower, which is important for the choice of photodetectors.

5. The scintillation emission time T_{scint} (if relevant).

In Table 3.1 some numbers for common calorimeter materials are listed. Depending on their structure one can further distinguish between two calorimeter types (for hadronic and electromagnetic calorimeter): *sampling* and *homogeneous*. In the sampling case, the calorimeter is divided into alternate sheets of dense shower media and signal light producer. For example, a common setup is to use layers of lead (or depleted uranium) interspersed with plastic scintillator. The shower develops in the lead layers; electrons/positrons from the shower passing through the plastic produce scintillation light, which is then detected. In a homogeneous calorimeter one substance acts as both shower medium and light producer, i.e. it is active.

---

[3]In a hadronic calorimeter one has also electromagnetic sub shower, however.

Material	X_0 (cm)	R_M	Light output	Peak λ (nm)	Emission time (ns)
NaI	2.59	4.8	1.00	410	230
CsI	1.85	3.5	0.20	315	16
CeF$_3$	1.68	2.6	0.08	340	25
PbWO$_4$	0.89	2.2	0.01	440	5-15

Table 3.1: Physical characteristics of some shower media used in electromagnetic calorimeters [77].

3.3.2 The energy resolution equation

The ultimate aim of an electromagnetic calorimeter is to measure the energy of photons/electrons as well as possible. The shower development in the calorimeter is a statistical process. This explains why the relative accuracy of energy measurements in calorimeters improves with increasing energy, $\frac{\sigma(E)}{E} \sim \frac{1}{\sqrt{E}}$. In practice the following equation is used to approximate the energy resolution:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{\sigma_N}{E} \oplus c \qquad (3.53)$$

where:

- E is the energy, usually in GeV,

- $\frac{\sigma_E}{E}$ is the energy resolution,

- a is the 'stochastic' term - mainly governed by fluctuations inherent in the development of showers,

- σ_N is the 'noise' term, covering instrumental effects such as electronics noise and pile-up,

- c is the constant term, which accounts for energy-independent effects such as calibration errors, non-uniformities and non-linearities in photomultipliers etc.

For a sampling calorimeter it is difficult to obtain a stochastic term below about 10% without demanding strict mechanical tolerances. In contrast, homogenous calorimeter have the potential to achieve stochastic terms of 2% due to much smaller sampling fluctuations. In this case, the limitation is the control of systematics building up the constant term, which is around 0.5%. The expected energy resolution of a PbWO$_4$ calorimeter, the choice of CMS [4], is [11]:

$$\frac{\sigma_E}{E} = \frac{2\%}{\sqrt{E}} \oplus \frac{\approx 200\text{MeV}}{E} \oplus 0.5\% \qquad (3.54)$$

---

[4]Since PbWO$_4$ is a very dense material that makes it possible to have a very compact calorimeter, furthermore it is a fast scintillator.

3.4   Physics simulation with the Geant4-toolkit

Geant4 is the object oriented successor of GEANT3 and widely used to describe particle interaction with matter, not only in particle physics, but also in medical applications, space simulations etc. It is a powerful package and offers a great magnitude of predefined interaction types, but leaves still a lot of freedom and flexibility to the user. So it is possible, for example, to define new interactions or particles, if one wants to simulate New Physics (example use case from CMS: interactions of R-hadrons). Several terms which are not well-defined in general usage have a specific meaning within Geant4:

- **process** - a C++ class which describes how and when a specific kind of physical interaction takes place along a particle track. A given particle type typically has several processes assigned to it.

- **model** - a C++ class whose methods implement the details of an interaction, such as its kinematics, the formula or parameterisation on which the model class is based. One or more models may be assigned to each process.

In our case the equation set 'GFLASH' is a model of the parameterisation process. For describing the integration of GFLASH into Geant4 (avaliable since Geant4 7.0) it is useful to first explain in more details the way of simulating physics processes in Geant4, focusing on the parameterisation process. A detailed description of Geant4, including a physics and a software reference manual, is available at the Geant4 website [25].

3.4.1   Processes in Geant4

The basic idea of simulating an interaction of a particle with matter is the following: A particle track consists of little 'steps': Each active discrete or continuous process must propose a step length for each step, based on the interaction it describes. The smallest of these step lengths is taken (also considering geometrical limitation, range cuts etc. - see Geant4 manual for details).

Physics processes describe how particles interact with a material. Seven major categories of processes are provided by Geant4:

1. electromagnetic

2. hadronic

3. photolepton -hadron

4. decay

5. optical

6. parametrisation

7. transportation

The generalisation and abstraction of physics processes is a key issue in the design of Geant4. All physics processes are treated in the same manner from the tracking point of view. The Geant4 approach enables anyone to create a process and assign it to a particle type. This openess should allow the creation of processes for novel, domain-specific or customised purposes by individuals or groups of users. Each process has two groups of methods which play an important role in tracking, `GetPhysicalInteractionLength` (GPIL) and `DoIt`. The GPIL method gives the step length from the current space-time point to the next space-time point. It does this by calculating the probability of interaction based on the process's cross section information. At the end of this step the DoIt method should be invoked. The `DoIt` method implements the details of the interaction, changing the particle's energy, momentum, direction and position, and producing secondary tracks if required. These changes are recorded as `G4VParticleChange` objects. `G4VProcess` is the base class for all physics processes. Each physics process must implement virtual methods of `G4VProcess` which describe the interaction (DoIt) and determine when an interaction should occur (GPIL).

The mean free path of a process, λ, also called the **interaction length** and calculated by (GPIL), can be given in terms of the total cross section :

$$\lambda(E) = \left(\sum_i [n_i \cdot \sigma(Z_i, E)] \right)^{-1}$$

where $\sigma(Z_i, E)$ is the total cross section per atom of the process and $\sum_i$ sums over all elements composing the material.

$\sum_i [n_i \sigma(Z_i, E)]$ is also called the *macroscopic cross-section*. The mean free path is the inverse of the macroscopic cross-section. Cross sections per atom and mean free path values are tabulated during initialisation. For clarification, the following example is a summary of the standard electromagnetic processes available in Geant4, which are used to simulate an electromagnetic shower:

1. Photon processes

   - Compton scattering (class name `G4ComptonScattering`)
   - Gamma conversion (also called pair production, class name `G4GammaConversion`)
   - Photo-electric effect (class name `G4PhotoElectricEffect`)

2. Electron/positron processes

- Bremsstrahlung (class name `G4eBremsstrahlung`)

- Ionisation and discrete delta ray production (class name `G4eIonisation`)

- Positron annihilation (class name `G4eplusAnnihilation`)

- The energy loss process (class name `G4eEnergyLoss`) handles the continuous energy loss of particles. These continuous energy losses come from the ionisation and bremsstrahlung processes.

- Synchrotron radiation (class name `G4SynchrotronRadiation`)

3. Hadron (e.m.) processes

  - Ionisation (class name `G4hIonisation`)

  - Energy loss (class name `G4hEnergyLoss`)

4. The multiple scattering process (class name `G4MultipleScattering`) is a general process in the sense that the same process/class is used to simulate the multiple scattering of all the charged particles (i.e. it is used for e^+/e^-, muons, charged hadrons).

3.4.2 The parameterisation process

The Geant4 parameterisation facilities allow for a shortcut to detailed tracking in a given volume and for given particle type in order to provide a user implementation of the physics and of the detector response. The volume to which one binds parameterisations is called an envelope. An envelope can have a geometrical sub-structure but all points in its daughter or sub-daughter (etc.) volumes are considered to be also in the envelope. Envelopes correspond often to the volumes of sub-detectors: electromagnetic calorimeter, tracking chamber etc. With Geant4 it is also possible to define envelopes by overlaying a parallel 'ghost' geometry. Parameterisations have three main features one has to specify:
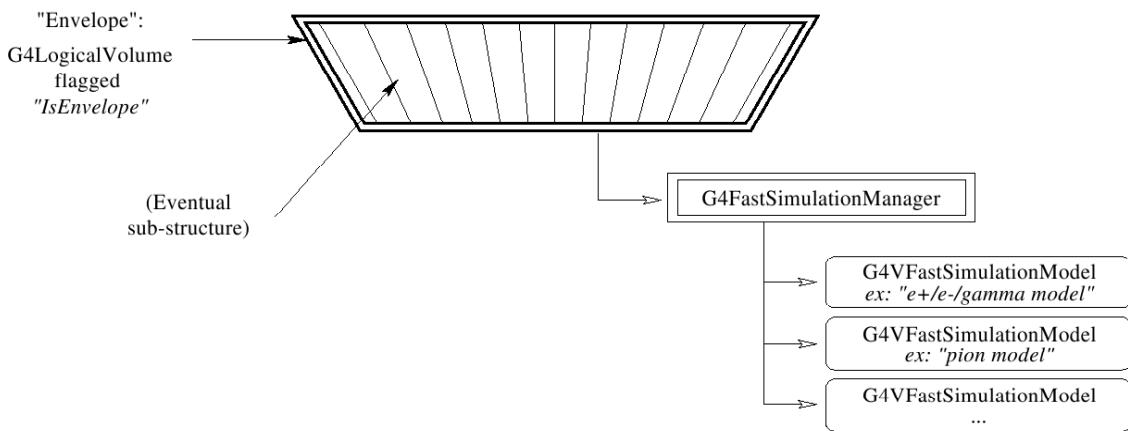
- The particle types for which the parameterisation is available.

- The dynamics conditions for which the parameterisation should be triggered.

- The parameterisation itself.

Geant4 will interrogate the parameterisations code for each step starting in the volume of the envelope. It will proceed by first asking to the parameterisations available for the current particle type if one of them (and only one) wants to invoke its parameterisation code. In this case, the tracking will not apply any other physics to the particle in the step.

3.4.3 Overview of the parameterisation components

The Geant4 components which allow the implementation and control of parameterisations are:

- **G4VFastSimulationModel**: This is the abstract class for the implementation of parameterisations - the user must inherit from it to implement a concrete parameterisation model.

- **G4FastSimulationManager**: The `G4VFastSimulationModel` objects are attached to the envelope through a `G4FastSimulationManager`. This object will manage the list of models and will interrogate them at tracking time.

- **Envelope**: An envelope in Geant4 is a `G4LogicalVolume` object which is simply flagged as being an envelope. [1] The parameterisation is bound to the envelope by setting a `G4FastSimulationManager` pointer to it. Figure 3.3 shows how the `G4VFastSimulationModel` and the `G4FastSimulationManager` objects are bound to the evelope.

- **G4FastSimulationManagerProcess**: This is a `G4VProcess`. It provides the interface between the tracking and the parameterisation and must be added to the process list of the particles the user wants to parameterise.

- **G4GlobalFastSimulationManager**: This a singleton class which provides the management of the `G4FastSimulationManager` objects and some ghost facilities.

Figure 3.3: Overview over the basic fast simulation components.

A more detailed description of the components is given in the following.

---

[1]From the Geant4 version 8.0 the envelope is defined by a `G4Region`.

The G4VFastSimulationModel class The `G4VFastSimulationModel` class has two constructors.

1. **G4VFastSimulationModel(const G4String& aName)**: here aName identifies the parameterisation model.

2. **G4VFastSimulationModel(const G4String& aName, G4LogicalVolume*, G4bool IsUnique=false)**:[1] In addition to the model name, this constructor accepts a `G4LogicalVolume` pointer. This volume will automatically become the envelope, and the needed `G4FastSimulationManager` object is constructed if necessary, giving it the `G4LogicalVolume` pointer and the boolean value. If it already exists, the model is simply added to this manager.

The `G4VFastSimulationModel` has three pure virtual methods which must be overridden in the user's concrete class:

1. **G4bool IsApplicable(const G4ParticleDefinition&)**: In the implementation of the user it must return 'true' when the model is applicable to the `G4ParticleDefinition` passed to this method. The `G4ParticleDefinition` provides all intrinsic particle information (mass, charge, spin, name ...).
For example, in a model valid for gammas only, the IsApplicable() method would take the form:

```
#include "G4Gamma.hh"
G4bool MyGammaModel::IsApplicable(const G4ParticleDefinition& partDef){
  return & partDef == G4Gamma::GammaDefinition();
}
```

2. **G4bool ModelTrigger(const G4FastTrack&)**: The trigger must return 'true' when the dynamic conditions to trigger the parameterisation are fulfilled. The `G4FastTrack` provides access to the current `G4Track` and gives simple access to envelope related features. Using these quantities and the `G4VSolid` methods, one can easily check how far the particle is from the envelope boundary.

3. **void DoIt(const G4FastTrack &, G4FastStep &)**: The details of the concrete parameterisation will be implemented in this method. The `G4FastTrack` reference provides the input information, and the final state of the particles after parameterisation must be returned through the `G4FastStep` reference. Tracking for the final state particles is requested after the parameterisation has been invoked.

---

[1]Since Genat4 8.0 : **G4VFastSimulationModel(const G4String& aName, G4Region*, G4bool IsUnique=false)**

The G4FastSimulationManagerProcess class This `G4VProcess` serves as an interface between the tracking and the parameterisation. At tracking time, it collaborates with the `G4FastSimulationManager` of the current volume, if any, to allow the models to trigger. If no manager exists or if no model issues a trigger, the tracking continues normally. In the present implementation, one must set this process in the `G4ProcessManager` of the particles which are parameterised.

The processes ordering is then:

- $(n-3)$

- $(n-2)$ `MultipleScattering`

- $(n-1)$ `G4FastSimulationManagerProcess`

- (n) `G4Transportation`

This ordering is important if one uses ghost geometries.
The `G4FastSimulationManager` must be added to the process list of the particle. The following code registers the `G4FastSimulationManagerProcess` for all the particles that are simulated:

```
void MyPhysicsList::addParameterisation(){
  G4FastSimulationManagerProcess*
    theFastSimulationManagerProcess = new G4FastSimulationManagerProcess();
  theParticleIterator->reset();
  while( (*theParticleIterator)() ){
      G4ParticleDefinition* particle = theParticleIterator->value();
      G4ProcessManager* pmanager = particle->GetProcessManager();
      pmanager->AddProcess(theFastSimulationManagerProcess, -1, 0, 0);
}
}
```

Parameterisation Using Ghost Geometries In some cases the technical setup of the tracking geometry do not allow envelopes to be defined. An interesting case involves defining an envelope which groups the electromagnetic and hadronic calorimeters of a detector into one volume. This may be useful, for example, if one wants to parameterise the interaction of charged pions. In this case one would not want electrons to be parameterised in this envelope. Geant4 provides for this use case ghost volumes, i.e. parallel geometries which can be organised by particle types. Using ghost geometries implies some extra overhead in the parameterisation mechanism for the particles sensitive to the current ghost volume, since additional navigation is provided in the ghost geometry by the `G4FastSimulationManager` process.

3.5   GFlash parameterisation in Geant4

In order to integrate a parameterisation 'a la' GFLASH into Geant4 it is on the one hand necessary to reimplement its set of equations in C++ and on the other hand to adapt the basic concepts of GFLASH to the object-oriented fast parametrisation framework. Especially the second point has to be considered with some care due to the fundamental implementation and architecture differences between GEANT3 (FORTRAN) and Geant4 (C++).

3.5.1   Basic GFlash components

As a first step, a GFlash library for homogenous calorimeter has been developed and intergrated into the Geant4 release from version 7.0.[1] Technically speaking, GFlash is a concrete parameterisation model which is based on the equations and parameters of the original GFLASH package from H1 presented in the last section and uses the 'fast simulation' facilities of Geant4 described above. Whenever a e^-/e^+ particle enters the calorimeter which is flagged to be an envelope, the parameterisation is triggered and the particle is parameterised if it has a minimum energy and the shower is expected to be contained in the calorimeter (or 'parameterisation envelope'). If these criteria are fullfilled, the particle is 'killed', as well as all its secondaries, and the energy is deposited according to the probability density functions (see section 3.2.2). Only electrons and positrons are parameterised. Photons are tracked by Geant4 until the first conversion process occurs, then the electron and the positron are parameterised. The object oriented GFlash implementation consists of the following classes:

- **GFlashShowerModel:** This is the concrete implementation of the `G4VFastSimulationModel` and acts as the interface to the full tracking. In the method `ModelTrigger` the containment of the shower inside the envelope and the dynamic conditions are checked.

- **GFlashHomoShowerParameterisation:** Here the shower profiles for homogenous calorimeters are calculated according to the original GFLASH.

- **GFlashParticleBounds:** This class checks the dynamic conditions of the particle, for example minimum energy for parameterisation etc.

- **GFlashShowerModelMessenger:** Provides the UserInterface to process commands for the interactive use of GFlash.

- **G4VGFlashSensitiveDetector:** This base class has to be used for detector volumes which serve as 'sensitive detectors', i.e. active materials which are

---

[1]Since version 8.0 sampling calorimeters are also supported, see Geant4 manual for details.

designed to process the information about the deposited energy. The users sensitive detector, which generates the hits, must be derived from this class and the G4VSensitiveDetector.

- **GFlashHitMaker:** Performs the deposition of the energy as 'hits' (=class that contains the energy and position information as well as additional user defined data) in the sensitive detector.

- **GVFlashHomoShowerTuning:** This class contains the parameter obtained from fits to full simulated showers. By instantiating an object of this class, the user has the possibility to pass his own parameters, in case a retuning is desired.

3.5.2 Usage of GFlash

To use GFlash 'out of the box' the following steps are necessary:

1. The user must add the fast simulation process to his process manager:

```
void MyPhysicsList::addParameterisation()
{
G4FastSimulationManagerProcess*
theFastSimulationManagerP = new G4FastSimulationManagerProcess();
  theParticleIterator->reset();
    while( (*theParticleIterator)() ){
       G4ParticleDefinition* particle = theParticleIterator->value();
        G4ProcessManager* pmanager = particle->GetProcessManager();
        pmanager->AddProcess(theFastSimulationManagerP, -1, 0, 0);
    }
}
```

2. The envelope in which the parameterisation should be performed must be specified (below: G4LogicalVolume m_calo_log) and the GFlashShowerModel must be assigned to this volume. Furthermore, the classes GFlashParticleBounds (which provide thresholds for the parameterisation like minimum energy etc.), GFlashHitMaker (a helper class to generate hits in the sensitive detector) and GFlashHomoShowerParamterisation (which does the computations) must be constructed and assigned to the GFlashShowerModel. This example is only valid for homogeneous calorimeters. The user must also define the material of the calorimeter, since the computation depends on the material. All necessary material dependant parameters like X_0, R_M etc. are automatically calculated by the package in a consistent way. A typical initialisation would be:

```
m_theFastShowerModel =
new GFlashShowerModel("fastShowerModel",m_calo_log);
m_theParameterisation =
new GFlashHomoShowerParameterisation(matManager->getMaterial(mat));
m_theParticleBounds  = new GFlashParticleBounds();
m_theHMaker          = new GFlashHitMaker();
m_theFastShowerModel->SetParameterisation(*m_theParameterisation);
m_theFastShowerModel->SetParticleBounds(*m_theParticleBounds) ;
m_theFastShowerModel->SetHitMaker(*m_theHMaker);
```

3. It is mandatory to use `G4VGFlashSensitiveDetector` as an (additional) base class for the sensitive detector in the user application:

   ```
   class ExGFlashSensitiveDetector: public G4VSensitiveDetector,
   public G4VGFlashSensitiveDetector
   ```

   Here it is necessary to implement a separate interface, where the energy deposits made by the `GFlashHitMaker` are processed:

   ```
   ProcessHits(G4GFlashSpot*aSpot, G4TouchableHistory* ROhist)
   ```

   The separate interface is used, because GFlash hits contain (naturally) less information than the full simulation (See next subsection 3.5.3).

Since the parameters in the GFLASH package are taken from fits to full simulations with GEANT3, some retuning might be necessary for good agreement with Geant4 showers. For experiment-specific geometries this might be necessary in any case. The GFlash framework already foresees the possibility of passing a class with user parameters, `GVFlashHomoShowerTuning` , to the `GFlashHomoShowerParamterisation` constructor. The default parameters are the original GFLASH parameters:

```
GFlashHomoShowerParamterisation(G4Material * aMat,
GVFlashHomoShowerTuning * aPar = 0);}
```

3.5.3 Implementation details and solved problems

Shower Containment: In inhomogeneous geometrical regions, i.e. calorimeter regions close to cracks or module borders where leakage of the shower is expected, fast parameterisation may not be accurate enough to simulate the detector response. GFlash assumes a uniform material and cannot take complex geometrical structures into account - therefore full simulation should be performed in these 'problematic' regions. Fast parameterisation, on the other hand, is used to speed up the simulation

in those parts of the detector that do not need an extremely accurate description for the physics process under consideration. If electrons shower before the calorimeter, each particle of the secondaries in the shower is tracked with the full simulation until it reaches the calorimeter, then each of the electrons is parameterised if it satisfies the requirements. If a particle is not fully contained in the calorimeter, the full simulation is activated. However, every fully contained secondary electron will be parameterised. The shower maximum and the lateral spread of the shower are calculated using the equations from subsection 3.2.2. An electron is considered fully contained if 90% of its energy is deposited in the calorimeter, i.e if the distance to the calorimeter bound is at least one R_M. In the worst case, the parameterisation would deposit up to 10% of the shower energy outside the calorimeter, i.e. this energy would be lost. This is, however, not a problem, since a comperable leaking energy fraction would have been lost as well in the full simulation. The minimum contained energy fraction can also be modified depending on the particular application. Technically, in the ModelTrigger method it is checked if four benchmark points on a cylinder with radius $1.5R_M$ around the longitududinal shower axis are still inside the envelope at $t = 2.5 \ X_0$.

Shower Starting Point: In GEANT3 the shower starting point was defined as the point of the first Bremsstahlung interaction of the electron or positron. In Geant4 this definition is technically difficult to implement, since after the Bremsstrahlung process is invoked, the particle has been already moved by the navigator and the secondaries have been produced. It would be necessary to shift the primary particle and kill all produced secondaries, producing a book-keeping overhead. Therefore the entry point in the calorimeter (`G4LogicalVolume`) is now defined as the shower start. The effect on the parameterisation accuracy due to this redefinition has been verified to be small.

Introduction of G4VGFlashSensitiveDetector: One problem that has been discovered during this work was that it is potentially dangerous to process energy deposits (hits) created with full tracking and fast parameterisation through the same interface (ProcessHits - a method which is mandatory for a sensitive detector), as it has been the case in the original Geant4 design. The original interface of `G4VSensitiveDetector`:

```
ProcessHits(G4Step* aStep*aStep, G4TouchableHistory* ROhist)
```

expects a `G4Step` object. However, some members of `G4Step` need information which is only available or only well defined with full tracking ('PreStepPoints', time stamps, track lengths etc.). Passing a `G4Step`, which is not correctly filled, may lead to a segmentation violation when the user tries to access this information after the parameterisation has been activated. A new clean solution,

the `G4VGFlashSensitiveDetector` base class, has been developed and included into the Geant4 release to solve this problem: the user code must now use both `G4VGFlashSensitiveDetector` and `G4SensitiveDetector` as base classes. If hit processing is performed in a parameterisation envelope, automatically the:

`ProcessHits(G4GFlashSpot*aSpot, G4TouchableHistory* ROhist)`

interface is called with a `G4GFlashSpot`, which contains only information available also in the case of fast parameterisation.

3.5.4 Example of usage in Geant4

To demonstrate the usage of GFlash an example has been developed and is now deployed in the Geant4 release in `examples/extended/parameterisations/gflash`. A simple model of the CMS calorimeter, namely a cubic 'module' with 10×10 PbWO$_4$ crystals, is used to show the usage of shower parameterisation. The crystals have a front side of 3×3 cm and a lengths of 24 cm, which is approximatively the size of the real lead tungstate crystals in CMS have. A particle gun shoots a single electron or positron with a fixed energy in the centre of the module, perpendicular to the module surface. The dimension of the module has been chosen in a way that showers at expected LHC energies are mostly fully contained in the 'parameterisation envelope' and hence the GFlash parameterisation is triggered, provided the particle is above the energy threshold for parameterisation. Geometry, sensitive detector, hits and processes are defined respectively in the example classes ExGFlashDetector-Construction, ExGFlashSensitiveDetector, ExGFlashHit and ExGFlashPhysicsList. Visualisation and an interactive user interface is also supported. The energy deposit in the central crystal, the 3×3 and 5×5 matrix around the central crystal are calculated and printed out. In addition, the simulation time is measured, so the user can see immediately the speed up using the parameterised shower. After building and running the example, the user can enter comands to the interactive prompt. For example, the command `GFlash/flag` steers whether shower parametrisation should be activated or not. The output of the example is shown below for the case of fully simulated and parameterised showers for a 10 GeV electron:
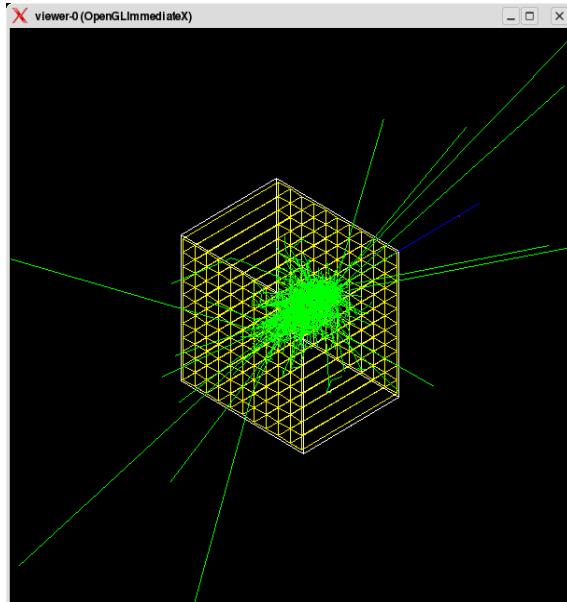
```
*****************************************
Internal Real Elapsed Time is: 0.58
Internal System Elapsed Time: 0
Internal GetUserElapsed Time: 0.56
*****************************************
 ------ ExGFlashEventAction::End of event nr. 1  -----
  20047 hits are stored in ExGFlashHitsCollection
   e1  8.34077   e3x3  9.6374   GeV  e5x5  9.75762
 Total energy deposited in the calorimeter: 9.9681011 (GeV)
```

```
     15886 trajectories stored in this event.
DRAWING 15886
number of event = 1
*****************************************
Internal Real Elapsed Time is: 0.02
Internal System Elapsed Time: 0
Internal GetUserElapsed Time: 0.02
*****************************************
 ------ ExGFlashEventAction::End of event nr. 2  -----
  2885 hits are stored in ExGFlashHitsCollection
    e1  8.31795    e3x3  9.64329    GeV  e5x5  9.77219
 Total energy deposited in the calorimeter: 9.9674102 (GeV)
     9 trajectories stored in this event.
DRAWING 9
number of event = 2
```

It can be immediately seen, that the usage of GFlash decreases the simulation time from 0.56 to 0.02 seconds, as well as the number of hits and trajectories, while the relevant energy deposits remain roughly the same. In Figure 3.4 a visual impression of the example is given showing the 10 GeV shower with full tracking.

Figure 3.4: 10 GeV electron shower with full tracking in a simplified $PbWO_4$ crystal calorimeter.

3.6   Physics and timing performance of GFlash

In order to examine the physics performance and the speed-up potential of GFlash in Geant4, as a first step, tests in pure lead tungstate were performed using the example setup described in the previous section. In this study the original GFLASH parameter from [67] were used and no tuning performed yet. Therefore, for electrons in the energy range from 1 GeV–1 TeV fully simulated showers were compared with parameterised showers. Another motivation for this procedure has been the assumption that the results obtained on this simple geometry correspond to the behaviour in the CMS calorimeter. On this basis, studies of the performance and possible corrections to the parameterisation could be preformed on this simple model, which is technically much easier to handle and less CPU consuming, and then transfered to the full CMS geometry. The thin aluminium matrix present in real life as support for the crystals is ignored in this model. It is also neglected in the material dependant calculations, where pure lead tungstate is assumed. After having defined the shower starting point as the entry point of the electron into the calorimeter module, the following quantities were compared:

1. **The longitudinal profile**: For each hit registered in the sensitive material the distance from the shower starting point is calculated and projected onto the initial particle direction.

2. **The radial profile**: For each hit registered in the sensitive material the radial component is calculated as the orthogonal distance from the initial particle direction.

3. The energy deposit in the most energetic crystal.

4. The energy deposit in the three × three matrix around the central crystal.

5. The energy deposit in the five × five matrix around the central crystal.

6. The simulation time per event.

A single electron was shot in the centre of the calorimeter perpendicular to the calorimeter surface. For each energy sample between 1 - 200 GeV, 1000 single showers were simulated, for the high energy region (500 - 1000 GeV) only a rough comparison with around 100 showers were performed due to CPU and file size limitations. The results are presented in the following. In Table 3.3 and Table 3.4 a comparison of the mean values for the longitudinal and the radial profile is presented. In Figure 3.5 to Figure 3.8 the longitudinal and radial profiles are shown, full Geant4 simulation corresponds to the histogram, GFlash simulation to the points. In order to deliver an up-to-date picture, the numbers presented in this section have been reprocessed and compared with the latest Geant4 release, 8.0, from December

2005. Since they are in agreement with the results obtained from the previous 7.0 release, this version is used in the following with the objective to be consistent with the CMS detector simulation which is also based on 7.0. Besides the profiles itself, the deposit in the crystals was compared.

This is shown in Figure 3.9, where the energy deposit in the central crystal and the 3×3 crystal matrix around the most energetic crystal is shown. Concerning the physics performance one can say in summary, that the longitudinal profiles show an acceptable agreement up to roughly 100 GeV, above 100 GeV they begin slightly to disagree. The shower maximum T occurs later in the full parameterisation. The radial profile shows a deviance as well: at the beginning of the core ($r < R_M$) component the parameterisation deposits more energy than the full simulation, in return the parameterisation deposits less energy than the full simulation in the interval between 2 and 6 R_M. The deposits in the crystals show some discrepancies as well, especially the deposit in the central crystal is too high for GFlash; this is the consequence of the disagreement in the radial profile. The 3×3 and 5×5 values show the same tendency, however, less dramatic than in the central crystal. Further tests performed on this calorimeter model have shown that the performance is independent of the angle between the electron trajectory and the calorimeter surface. Also the relative entry position in the crystal does not show an influence: the performance is equally good near the crystal border and in the centre of the crystal. Only if the particle enters exactly in the crack the performance gets worse on a percent level; this case is, however, not very likely in real life. The presence of a constant magnetic field not show a significant influence on the shower shape. Table 3.2 shows the speed-up factors for different energies. In this simple model the shower is always contained and thus the electron always parameterised. The speed-up is impressive and ranges up to a factor 100 and more for electrons with energies above 50 GeV. The same study has been performed as well with single photons. Photons trigger the parameterisation in this setup as soon as they have converted into an e^+e^- pair. The obtained results with respect to the timing and physics performance are the same as for electrons and not shown explicitly. In conclusion, this study shows that the new implementation of GFlash in Geant4 works technically and the timing performance is very good. The physics performance using the old GFLASH parameter 'out of the box' delivers acceptable results, but requires some tuning if a parameterisation in the high high energetic region (> 100 GeV) is performed or a more accurate description of the radial profile (percent level) is needed. The GFlash package is currently maintained by in Geant4 by the author and can be used in any HEP experiment or any other simulation applications based on Geant4. More details can be found in the fast parameterisation section of the G4 physics and software manual.

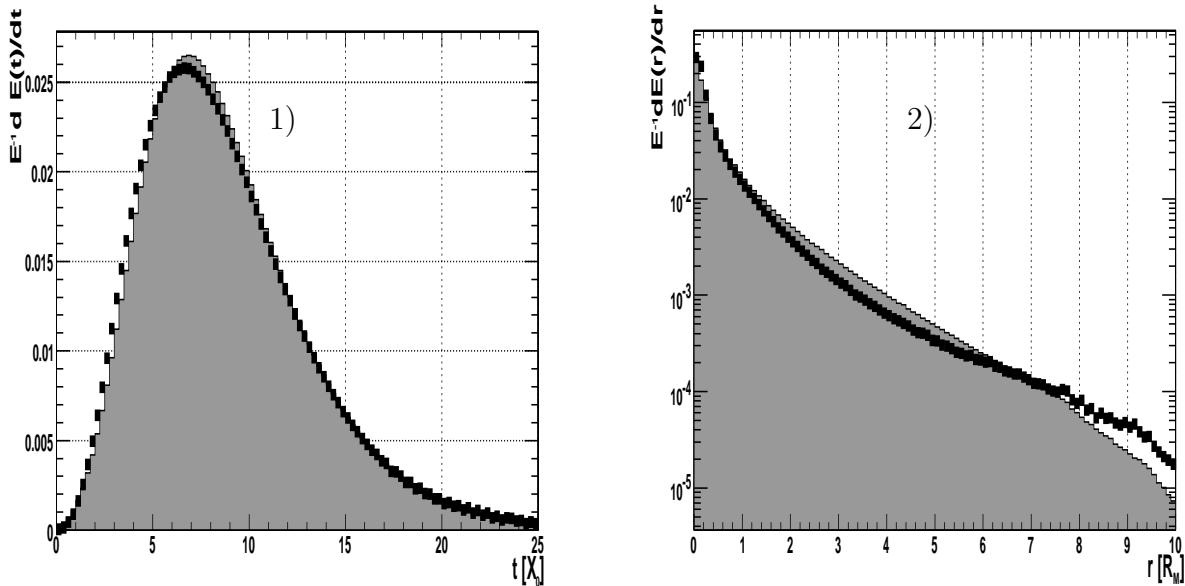| Energy	Time Geant4	Time GFlash	Speed-up factor
1 GeV	0.10	0.006	16
5 GeV	0.46	0.009	51
10 GeV	0.92	0.013	70
50 GeV	4.60	0.045	102
100 GeV	9.37	0.080	117
500 GeV	46.50	0.312	149
1000 GeV	91.75	0.566	162

Table 3.2: Simulation time and speed-up factors (rounded) of single electrons/positrons in a pure lead tungstate cube.

Energy	Mean Geant 4 $[X_0]$	Mean GFlash $[X_0]$
1 GeV	5.87	6.05
10 GeV	8.16	8.10
20 GeV	8.80	8.73
50 GeV	9.74	9.54
100GeV	10.48	10.26
200 GeV	10.99	10.84
500 GeV	12.03	11.46
1000 GeV	12.90	12.25
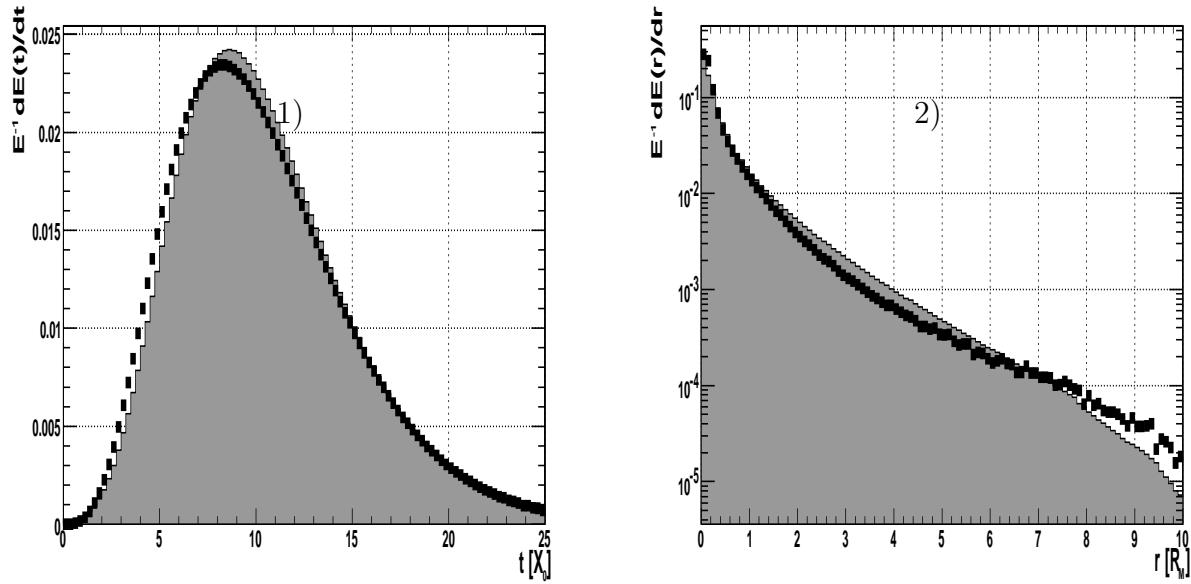
Table 3.3: Comparison between the mean value of the longitudinal profile in GFlash and Geant4 (X_0 units, X_0 calculated by Geant4).

Energy	Mean Geant 4 [R_M]	Mean GFlash [R_M]	R_M^{sim} Geant4 [R_M]	R_M^{sim} GFlash [R_M]
1 GeV	0.55	0.49	1.5	1.3
10 GeV	0.55	0.47	1.5	1.2
20 GeV	0.55	0.47	1.5	1.2
50 GeV	0.55	0.46	1.5	1.2
100GeV	0.55	0.46	1.5	1.2
200 GeV	0.55	0.47	1.5	1.2
500 GeV	0.55	0.48	1.5	1.3
1000 GeV	0.54	0.49	1.5	1.3
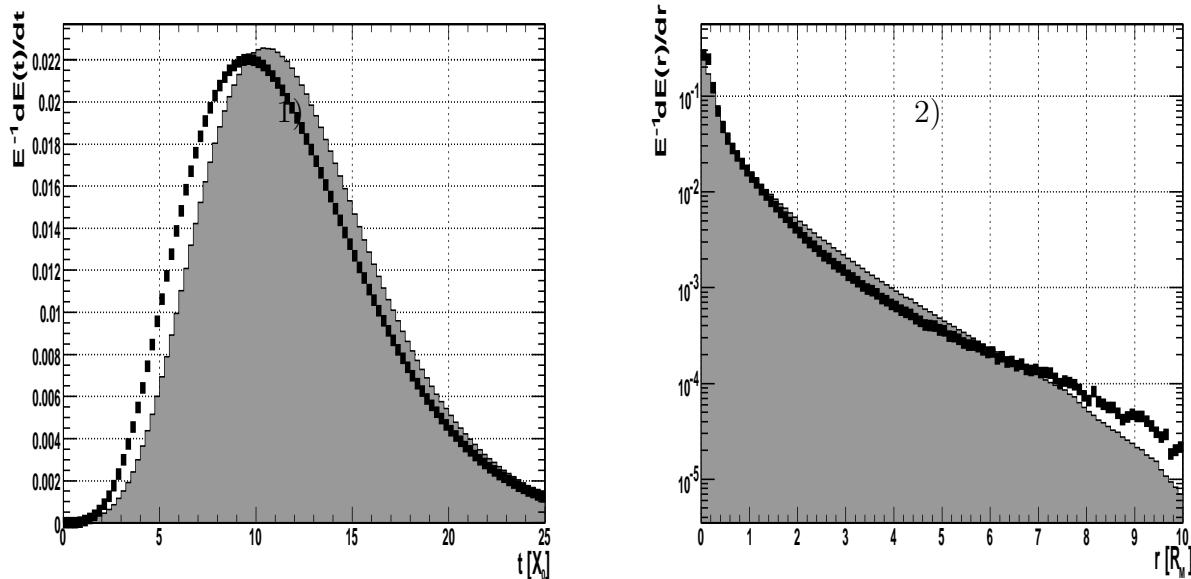
Table 3.4: Comparisons between the mean value of the radial profile in GFlash and Geant4 (R_M units, R_M calculated from Equation 3.35). In the last two columns also the Molière radius R_M^{sim} calculated from the simulation as the radius r of a cylinder around the longitudinal shower axis that contains 90% of the shower energy is compared for GFlash and Geant4.
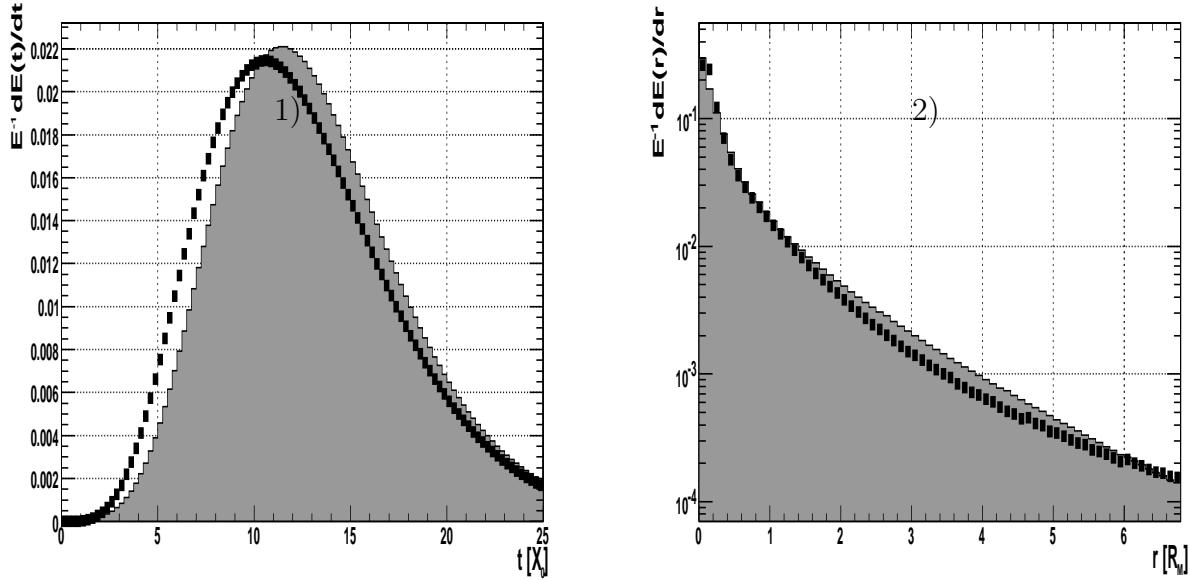
Figure 3.5: Longitudinal (1) and radial (2) shower profile for a 20 GeV electron in $PbWO_4$ (Geant4: line, GFlash: points)
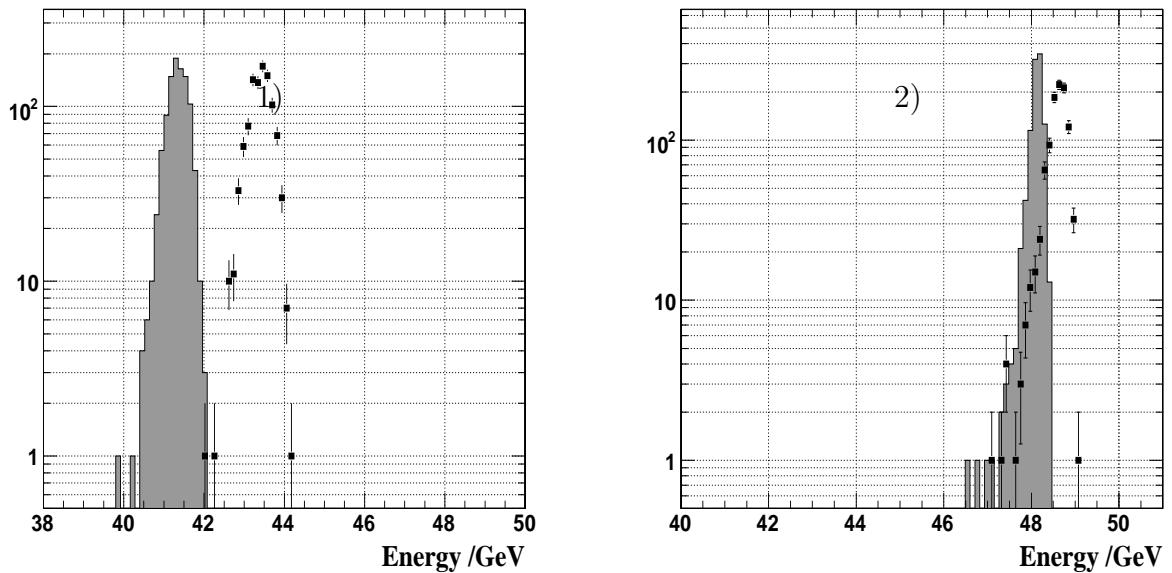
Figure 3.6:   Longitudinal (1) and radial (2) shower profile for a 100 GeV electron in
$PbWO_4$ (Geant4: line, GFlash: points).

Figure 3.7:   Longitudinal (1) and radial (2) shower profile for a 500 GeV electron in
$PbWO_4$ (Geant4: line, GFlash: points).
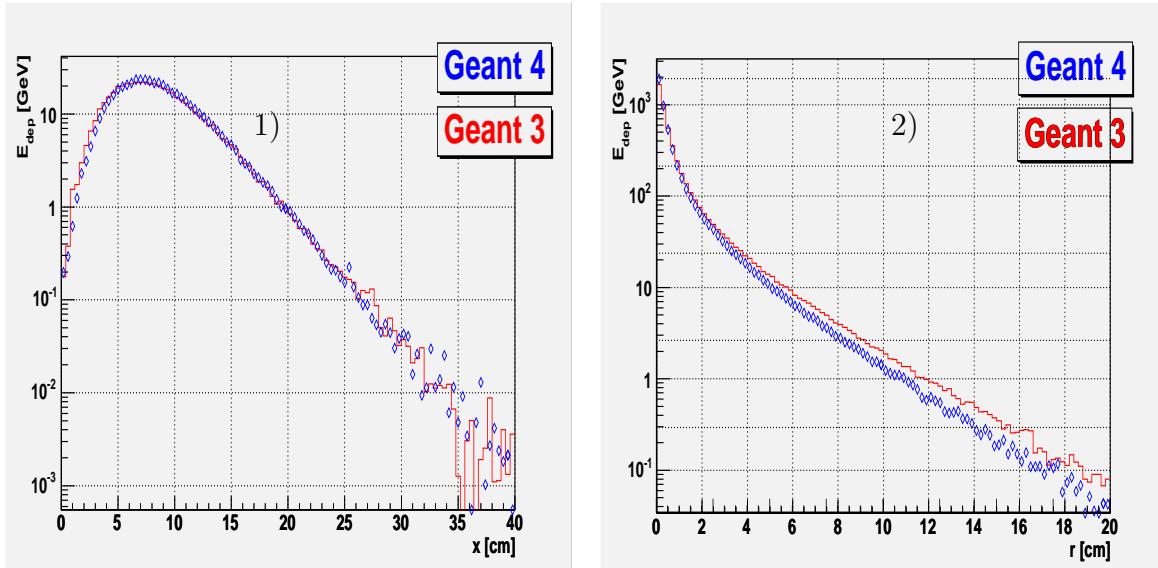
Figure 3.8: Longitudinal (1) and radial (2) shower profile for a 1000 GeV electron in $PbWO_4$ (Geant4: line, GFlash: points).
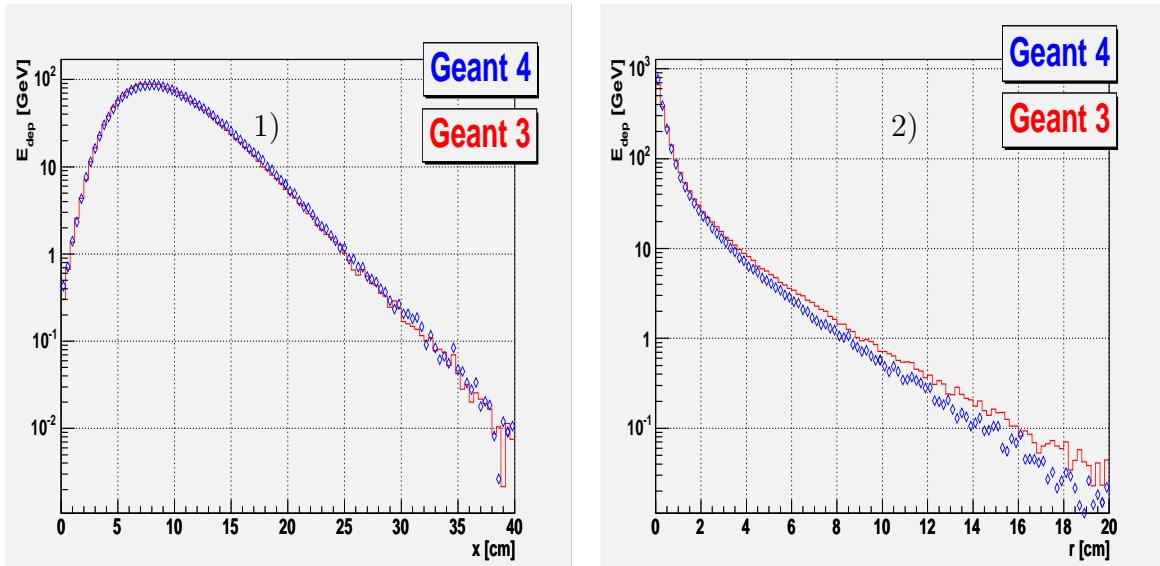
Figure 3.9: Energy deposit in the central crystal (1) and in the 3×3 matrix (2) for a 50 GeV electron. (Geant4: line, GFlash: points).

3.7   Comparison between GEANT3 and Geant4

Due to the good energy resolution in the CMS electromagnetic calorimeter, the simulation of the radial profiles with the parameters obtained with GEANT3, may not be sufficiently accurate. In order to verify that the parameter set obtained with GEANT3 can be used with Geant4, the longitudinal and radial shower profiles were compared in lead (Pb) and in the relevant material for CMS: lead tungstate ($PbWO_4$). The energy range considered was 10 to 500 GeV. The simulation setup has been similar as in the last section 3.6 with the difference that only a matter block of $PbWO_4$ has been simulated, without the crystal substructure. In the following, the comparison tables for the mean value of the longitudinal and the radial profile in ($PbWO_4$) are shown, as well as a comparison of the total energy deposit. Figure 3.10 - Figure 3.11 show example shower profiles with GEANT3 in comparison with Geant4 for 50 and 200 GeV. From this study, one can conclude that especially the radial profiles show some differences between GEANT3 and Geant4. The longitudinal profiles agree on a percent level up to 500 GeV, while for the radial profiles the Molière radii differ in some cases by about 20%, as can be seen in Table 3.6 and Table 3.7. Therefore a tuning of the parameters has to be performed if an agreement on the percent level is needed.

Figure 3.10: Longitudinal (1) and radial (2) shower profile for a 50 GeV electron in $PbWO_4$ (GEANT3: red line, Geant4: blue points).

Figure 3.11: Longitudinal (1) and radial (2) shower profile for a 200 GeV electron in $PbWO_4$ (GEANT3: red line, Geant4: blue points).

Energy [GeV]	Geant4 [GeV]	GEANT3 [GeV]
1	0.99	0.99
10	9.96	9.97
20	19.96	19.97
50	49.95	49.95
100	99.90	99.98
200	199.82	199.81
500	499.65	499. 52

Table 3.5: Comparison between deposited energy by GEANT3 and Geant4.

Energy [GeV]	Geant4 [cm]	GEANT3 [cm]
1	5.30	5.34
10	7.27	7.28
20	8.04	8.00
50	8.7	8.55
100	9.67	9.50
200	9.91	9.55
500	11.3	10.29

Table 3.6: Comparison between the mean values of the longitudinal profile simulated with GEANT3 and Geant4.

Energy [GeV]	Mean G4 [cm]	Mean G3 [cm]	Molière radius G4 [cm]	Molière radius G3 [cm]
1	0.75	0.91	2.0	2.4
10	0.80	0.95	2.2	2.6
20	0.80	0.95	2.0	2.6
50	0.80	0.94	2.2	2.6
100	0.79	0.94	2.2	2.6
200	0.8	0.95	2.2	2.6
500	0.8	0.95	2.2	2.6

Table 3.7: Comparison between the mean value of the radial profile and the Molière radii with GEANT3 and Geant4.
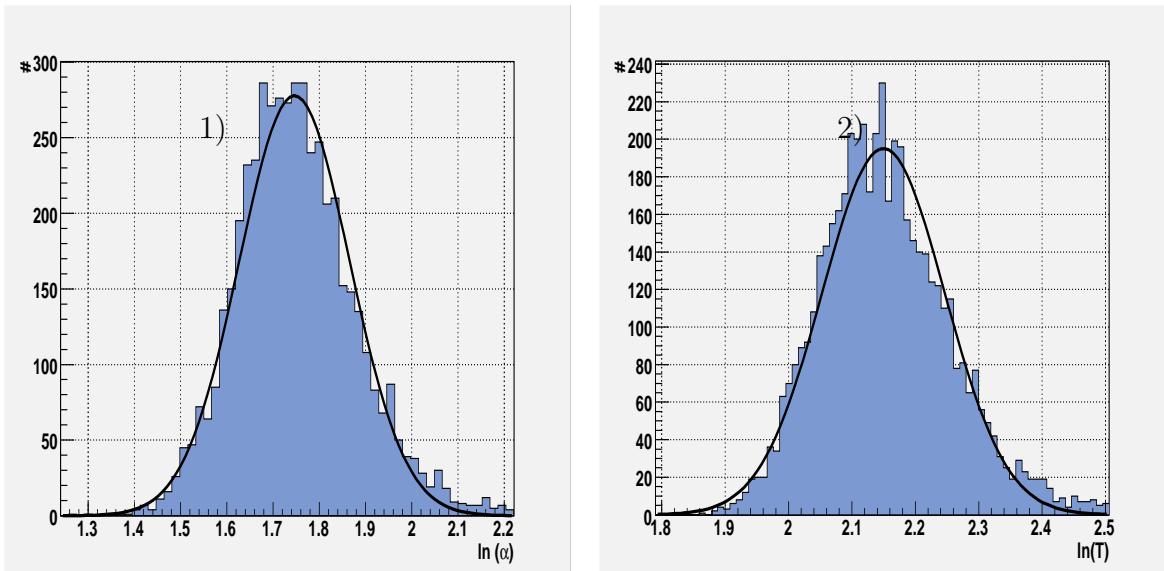
3.8   Tuning of the longitudinal profiles

The longitudinal profiles calculated by GFlash show an acceptable agreement with full simulation in the energy region up to about 100 GeV. In contrast, in the high-energy region the agreement gets worse and the centre of gravity is shifted by even up to 10%. With the objective to obtain a better description of the fully simulated Geant4 shower profiles, a retuning procedure has been set-up. In this context, 5000 shower profiles has been generated with Geant4 for single electrons with the following energies: 1 GeV, 10 GeV, 50 GeV, 100 GeV, 200 GeV, 300 GeV, 500 GeV, 1000 GeV, 1300 GeV. At each energy the shape parameter α and the scaling factor β are determined under the assumption that the longitudinal profile can be described by Equation 3.37. In this case, the following relations can be derived for the shower maximum T, the centre of gravity $< t >$ and its standard deviation $\sigma^2(t)$:

$$< t > \quad = \quad \frac{\alpha}{\beta} \tag{3.55}$$

$$\sigma^2(t) \quad = \quad \frac{\alpha}{\beta^2} \tag{3.56}$$

$$T \quad = \quad \frac{\alpha - 1}{\beta} \tag{3.57}$$

Figure 3.12:   The logarithm of the normal distributed shape parameter $\ln(\alpha)$ (1) and the shower maximum $\ln(T)$ (2) computed and histogramed for 5000 fully simulated shower.

Two strategies are usually applied to extract α and β from full simulation:

On the one hand, gamma distributions can be fitted to each single longitudinal shower profiles using a fitting tool (i.e. root) and the parameters (α and β or α and the related shower maximum T) are determined and histogrammed. On the other hand, the two parameters can be derived analytically from the first and second moment of the gamma distribution $f(t)$. In general, the moment of order n is defined as

$$Z_n = \int_0^\infty t^n f(t) dt = \frac{\Gamma(\alpha + n)}{\beta^n \Gamma(\alpha)} \tag{3.58}$$

and in particular the first moments are:

$$Z_1 = \frac{\alpha}{\beta} \; ; \; Z_2 = \frac{\alpha}{\beta^2}(\alpha + 1) \tag{3.59}$$

Hence the parameters are simply given by the equations:

$$\alpha = \frac{Z_1^2}{Z_2 - Z_1^2} \; ; \; \beta = \frac{Z_1}{Z_2 - Z_1^2} \tag{3.60}$$

In this thesis the parameters were obtained with the second method from fully simulated Geant4 reference samples. Since α and T have non-Gaussian distributions but their logarithms, on the contrary, are normal distributed, $\ln(\alpha)$ and $\ln(T)$ are used in the following rather than the parameters themselves. The advantage is that a Gaussian distribution can be entirely described by just two quantities: its mean and its width; the width of the distribution reflects the deviation of a single showers from the average shower profile. The procedure has been as follows:

1. For every fully simulated shower $\ln(\alpha)$ and $\ln(T)$ were calculated from the moments and histogramed, as shown in Figure 3.12.

2. The average values $< \ln(T) >$ and $< \ln(\alpha) >$ as well as their fluctuation $\sigma(\ln(T))$ and $\sigma(\ln(\alpha))$ were than obtained from a gaussian fit to the histograms, as is illustrated Figure 3.12.

3. Since there is correlation between α and T as shown in Figure 3.13, the correlation coefficient

$$\rho(\ln(T), \ln(\alpha)) = \frac{< \ln(T)\ln(\alpha) > - < \ln(T) >< \ln(\alpha) >}{\sigma(\ln(T))\sigma(\ln(\alpha))} \tag{3.61}$$

is also calculated and taken into account.

4. In the last step the average values $< \ln(T) >$ and $< \ln(\alpha) >$ are parameterised using the logarithms of Equation 3.41 and Equation 3.42. The parameterisation of $< \ln(T) >$ is here assumed to be not material dependent as function

of y (with $y = E/E_c$), while $< \ln(\alpha) >$ features an explicit dependence on Z. In summary, the following equations were used for the retuning:

$$
\begin{aligned}
\langle \ln T \rangle &= \ln(t_1 \ln y - t_2) \\
\langle \ln \alpha \rangle &= \ln \left(a_1 + (a_2 + a_3/Z) \ln y \right) \\
\sigma(\ln T / \ln \alpha) &= (s_1 + s_2 \ln y)^{-1} \\
\rho(\ln T, \ln \alpha) &= r_1 + r_2 \ln y
\end{aligned}
$$

For illustration, in Figure 3.14 the fitted functions for $< \ln(T) >$ and $< \ln(\alpha) >$ as function of y are shown. Thereby as error $\Delta(x)$ for their average values $< x >$ and their fluctuations $\sigma(x)$ the following terms are taken: $\Delta(< x >) = \sigma(x)/\sqrt{N}$ and $\Delta(\sigma(x)) = \sqrt{2}\sigma^2(x)/\sqrt{N}$.
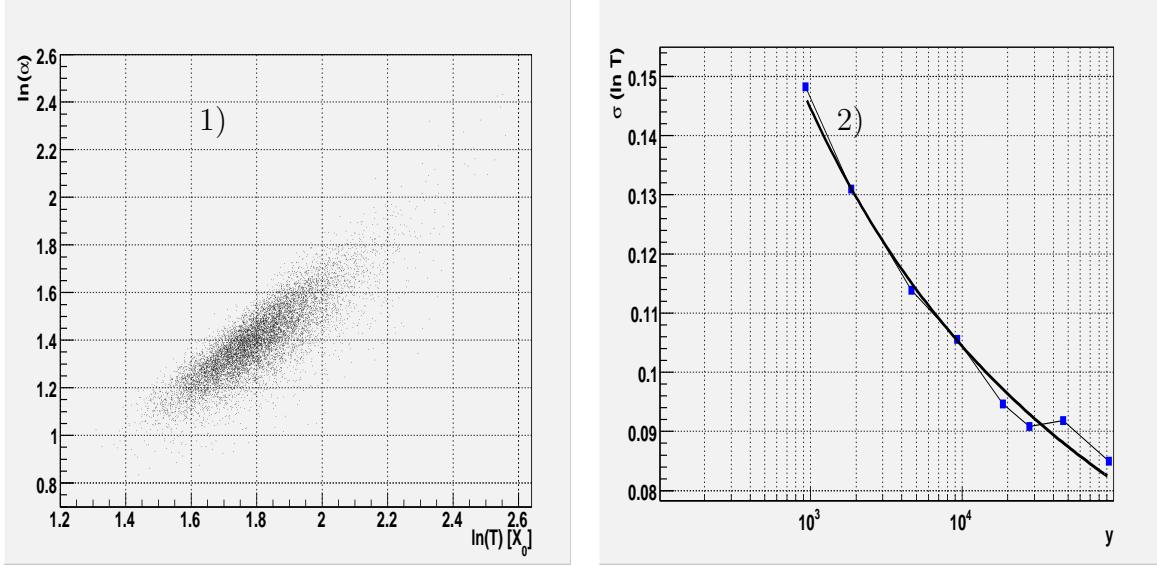
After the retuning procedure, the shower profiles for all energies were simulated with the tuned GFlash and compared to the full Geant4 simulation. To judge the effect of the retuning in a quantitative way, the following variables were looked at: the difference in the mean and the RMS values of the shower profiles and a χ^2-like quantity defined as follows:

$$
\chi^2 = \sum_{i=1}^{n} \frac{(\bar{x}_i - y_i)^2}{\sigma^2(\bar{x}_i)} \tag{3.62}
$$

Here $\bar{x}_i$ is the Geant4 average value in each bin (assumed as the truth), $\sigma^2(\bar{x}_i)$ is its statistical error and y_i the average profile value obtained using GFlash. After the tuning the χ^2 values could be improved by up to 70% in the high energetic region with respect to the χ^2 values of the original GFLASH parameter. The tuning tools setup here are fully automated and can be used 'out of the box' by executing a single script. They can be easily reused for the tuning to any new material or a new geometry. However, the ability to describe the real shower profile by GFlash in this setup is fully dependent on the accuracy of the Geant4 physics performance.

3.9   Tuning of the radial profiles

The detailed studies of the GFlash performance and the comparison with GEANT3 indicate that especially the radial profile parameterisation needs adjustment when moving to Geant4 and using a $PbWO_4$ calorimeter, a material which has not been considered explictly in [67]. Furthermore, the energy range at the LHC is much higher than the range explored by H1. From the direct comparison of the energy deposit in the crystal one can see that GFlash tends to deposit too much energy in the central crystal (see Figure 3.9). In order to stick as close as possible to the original GFlash parameter and get at the same time a better agreement for the transverse profile a correction factor, k, for the weight, p, was introduced in Equation 3.47, leading to the modified form:

Figure 3.13: Correlation between the logarithms of the shower shape parameter α and the shower maximum T (left) and example fit to the fluctuation of T, $\sigma(\ln T)$, (right).
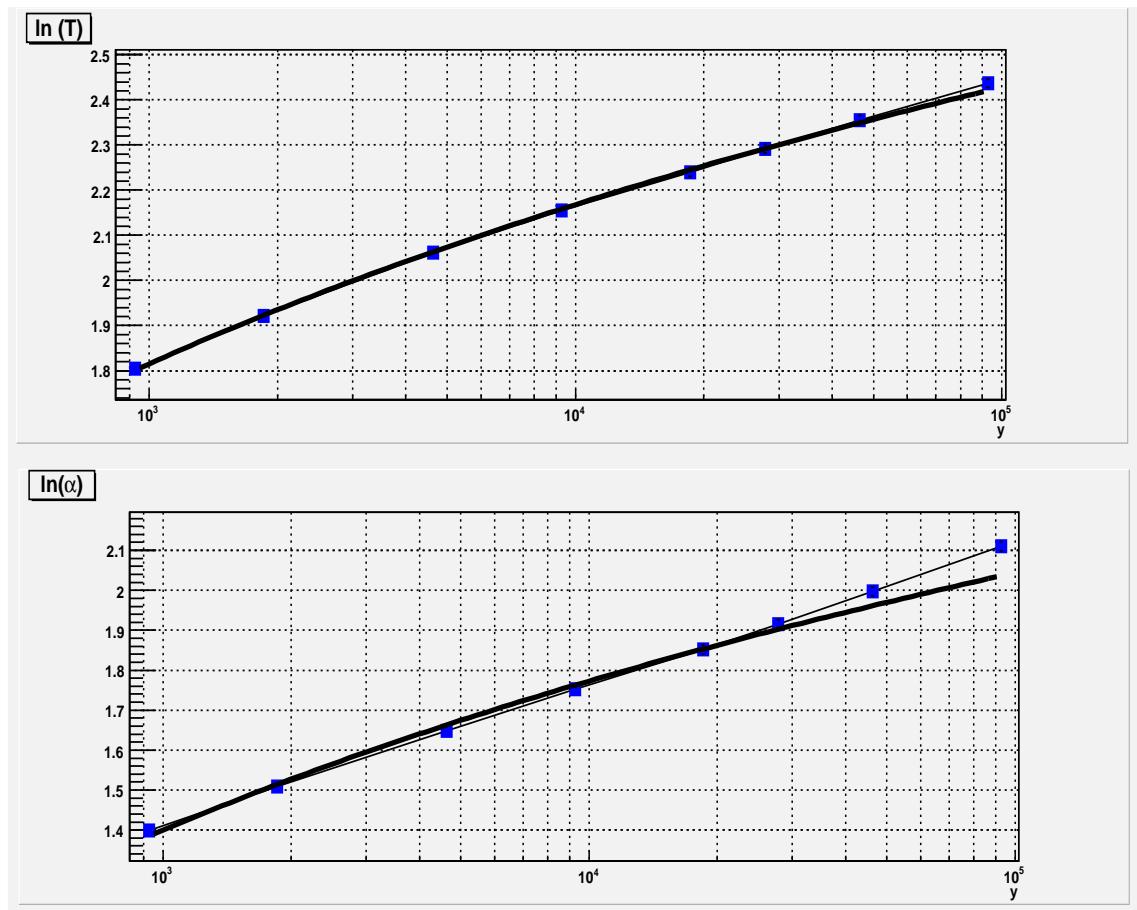
$$
\begin{aligned}
f(r) &= (p*k)f_C(r) + (1 - (p*k))f_T(r) & (3.63) \\
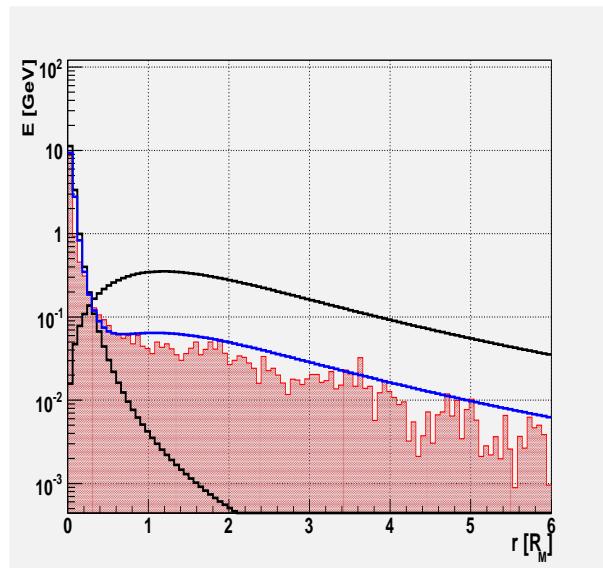&= (p*k)\frac{2rR_C^2}{(r^2 + R_C^2)^2} + (1 - (p*k))\frac{2rR_T^2}{(r^2 + R_T^2)^2} & (3.64)
\end{aligned}
$$

Thus the core and tail component of the radial profile are reweighted. In Figure 3.15 the tuning procedure, as well as the core and the tail components of the radial profile are visualised.

The procedure to tune the parameters for the radial profile with Geant4 has been as follows:

1. The radial profile is plotted in bins of the longitudinal shower profile with the size of one X_0 in the range from $X_0 < t < 20X_0$. (t being the longitudinal shower coordinate).

2. For each longitudinal interval the radial profile function from Equation 3.64 is fitted with only k as the free parameter and all others parameters fixed.

3. The obtained correction factors $k_1...k_{20}$ are now fitted as a function of the longitudinal coordinate. As a first approach, a linear function was taken for $k(t)$.
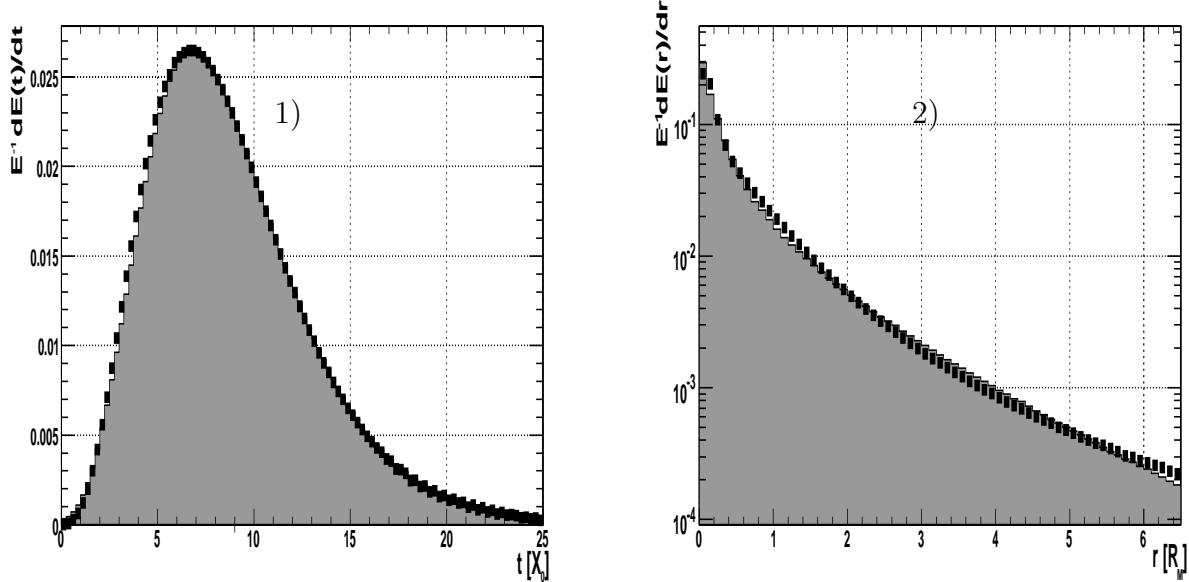
Figure 3.14: The fitted curves for $< \ln(\mathrm{T}) >$ and $< \ln(\alpha) >$ as function of $y = E/E_c$.
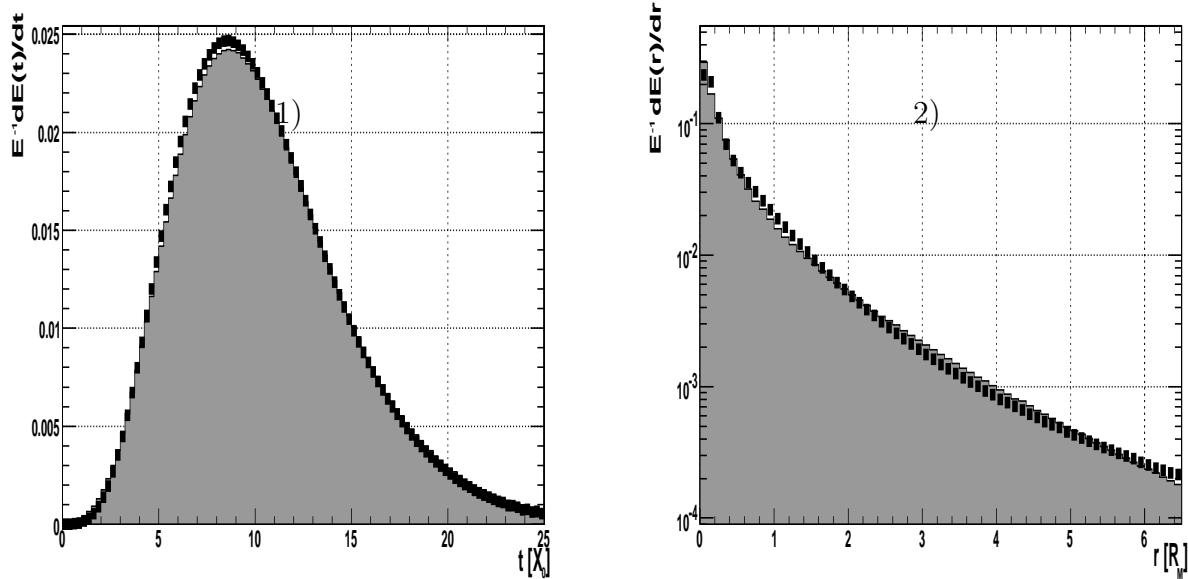
Figure 3.15: The radial function (blue, see Equation 3.64) is fitted to the fully simulated radial profile (histogram). In black the core (left) and tail (right) component of the radial function are plotted.

With this procedure the agreement could be significantly improved. The energy deposit in the central crystal and the matrices after the retuning show an agreement at the level of 1%. The χ^2 values could be improved by up to 17% with respect to the χ^2 values obtained with the original GFLASH parameter set.
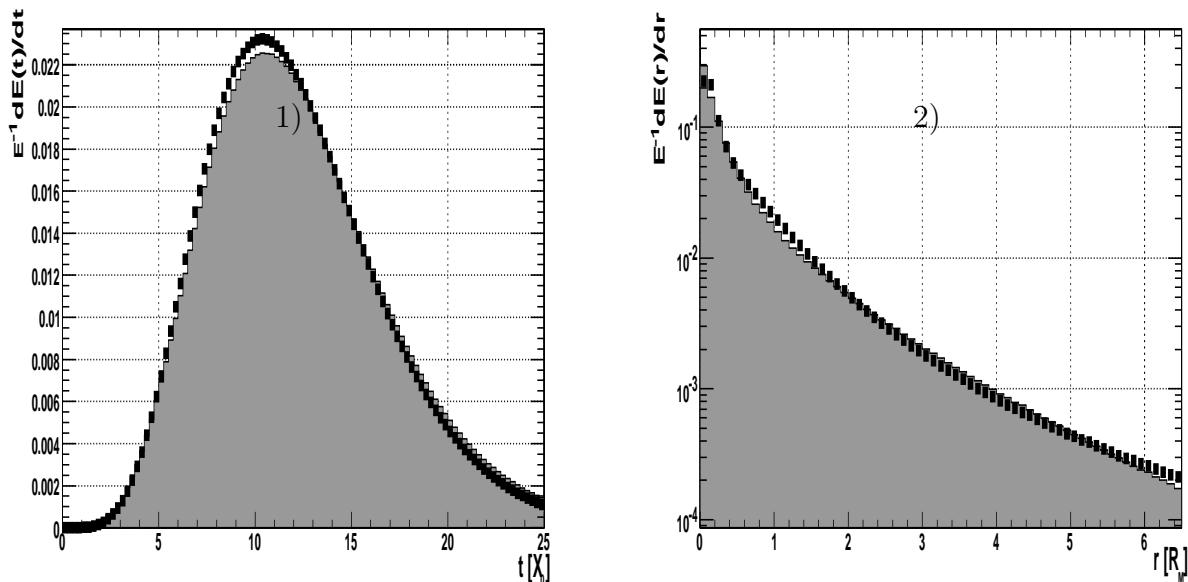
In summary, the agreement of the longitudinal and radial profiles could be significantly improved, as can be seen in Figure 3.16–Figure 3.18. The simple tuning procedure of the radial profiles improved as well the agreement in the energy distribution in the central crystal and the sourrounding matrices, as can be seen in Figure 3.20. A complete full retuning of all radial parameters can be performed, but is not necessary for the CMS use case: the energy resolution in CMS is expected to be $\sim \frac{2\%}{\sqrt{E}}$, see Equation 3.53. Since the agreement between full simulation and GFlash is already better than 1% on simulation level where no reconstruction effects has been taken into account (noise, electronics), the accuracy is sufficient.
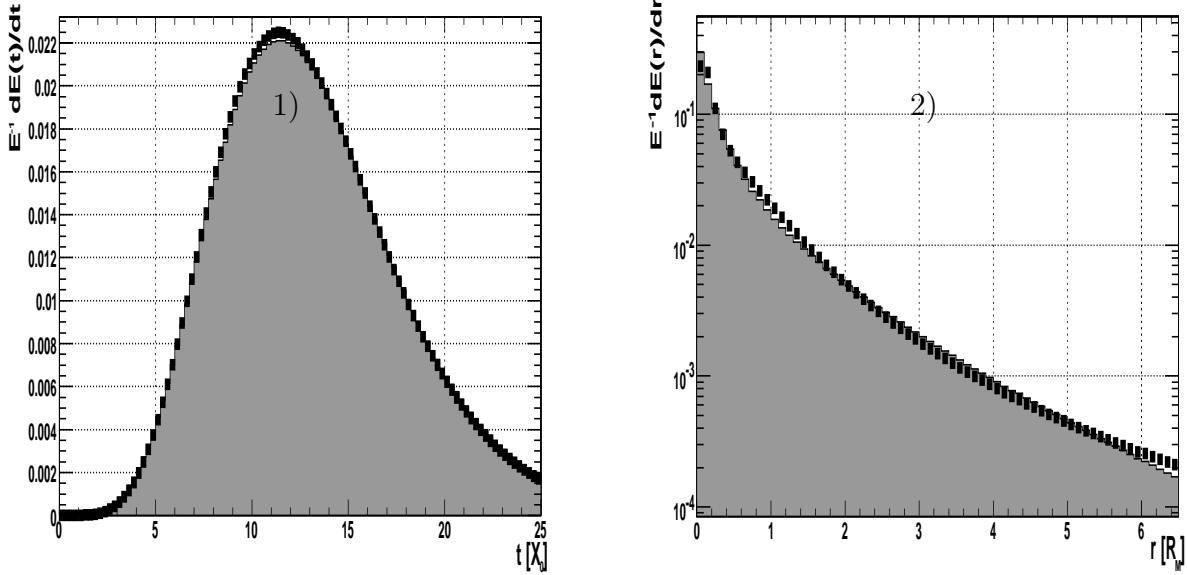
Figure 3.16: Longitudinal (1) and radial (2) shower profile for a 20 GeV electron in $PbWO_4$ after tuning (Geant4: line, GFlash: points).
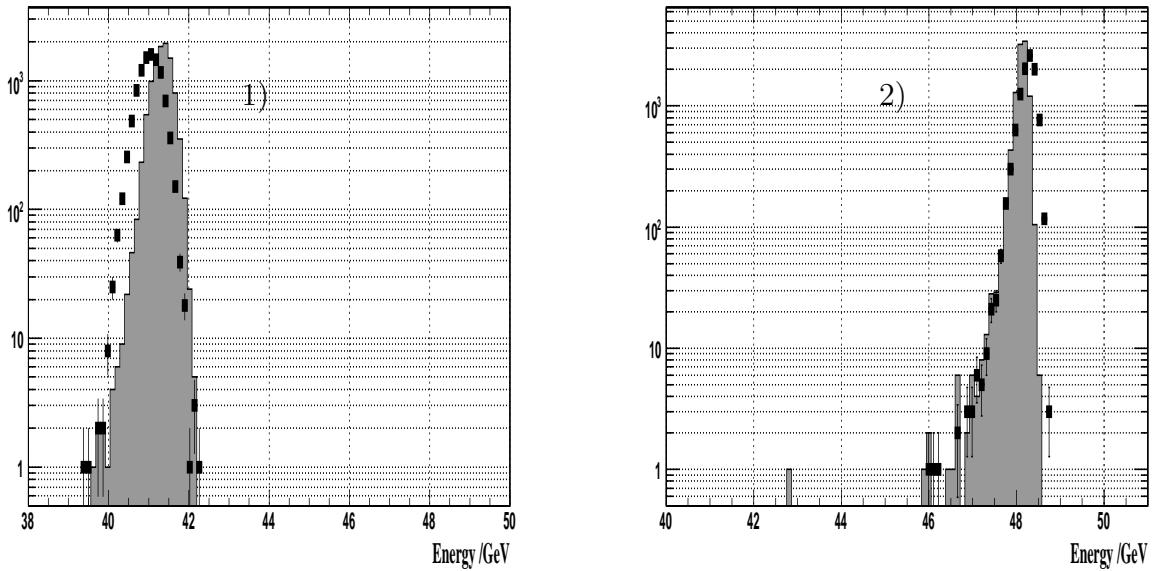
Figure 3.17: Longitudinal (1) and radial (2) shower profile for a 100 GeV electron in $PbWO_4$ after tuning (Geant4: line, GFlash: points).

Figure 3.18: Longitudinal (1) and radial (2) shower profile for a 500 GeV electron in $PbWO_4$ after tuning (Geant4: line, GFlash: points).

Figure 3.19:   Longitudinal (1) and radial (2) shower profile for a 1000 GeV electron in $PbWO_4$ after tuning (Geant4: line, GFlash: points).

Figure 3.20:   Energy deposit in the central crystal (1) and in the 3×3 matrix (2) for a 50 GeV electron after tuning. The agreement of the mean is better than one percent (Geant4: line, GFlash: points).

3.10   GFlash in the CMS detector simulation

In order to interface GFlash to the full CMS detector simulation OSCAR [23] the steps described below were performed. In addition, an example has been added to the CMS detector simulation package to illustrate the usage of shower parameterisation in CMS. (in **OscarApplication/GFlashTest**). The procedure has been the following:

1. In **PhysicsSim/GFlash**: The GFlash physics list (the physics list is the module where all physical interactions are specified) is constructed in **GFlash.cc**. It basically consists of the QGSP physics list, which is often used as default in OSCAR simulation, and, in addition, the parameterised physics, which is specified in **ParameterisedPhysics.cc** - here the FastSimulationManger is added to the processes active during simulation. Finally in **CaloModel.cc**, the parametrisation is assigned to the appropriate volume (i.e. the electromagnetic calorimeter). Technically, the XML based geometry description is read in and the correct G4LogicalVolume determined. Then the GFlash parameterisation is constructed with the correct material properties for the CMS $PbWO_4$ calorimeter and attached to this volume(s) - in this case the barrel and endcap parts of the electromagnetic calorimeter.

```
//Finding correct G4LogicalVolume for parameterisation
ConcreteG4LogicalVolumeToDDLogicalPartMapper::Vector vec =
        G4LogicalVolumeToDDLogicalPartMapper::instance()->all("volumes");
for (ConcreteG4LogicalVolumeToDDLogicalPartMapper::Vector::iterator
        tit = vec.begin(); tit != vec.end(); tit++){
if (((*tit).first)->GetName()=="ESPM"){
cout <<" GFlash added  to voulme  "<< ((*tit).first)->GetName() << endl;
barrel_log = (*tit).first;
}
if (((*tit).first)->GetName()=="ENCA"){
cout <<" GFlash added  to voulme  "<< ((*tit).first)->GetName() << endl;
ecap_log = (*tit).first;
}
}
```

2. In **CaloSim/CaloSD**: Here the separate GFlash interface had to be implemented in the hit processing method for the electromagentic calorimeter. In addition, the G4VSensitiveDetector base-class had to be added to the sensitive detector to enable the GFlash hit processing.

3. **ShowerModuleVolumes.xml**: This is the XML / DDD description of the volumes in which the parameterisation is activated.

Showers are parameterised in the endcap and barrel calorimeter region of CMS. In the η region between 1.47 and 1.50, namely the edge region between the barrel and the endcap, full simulation is performed.
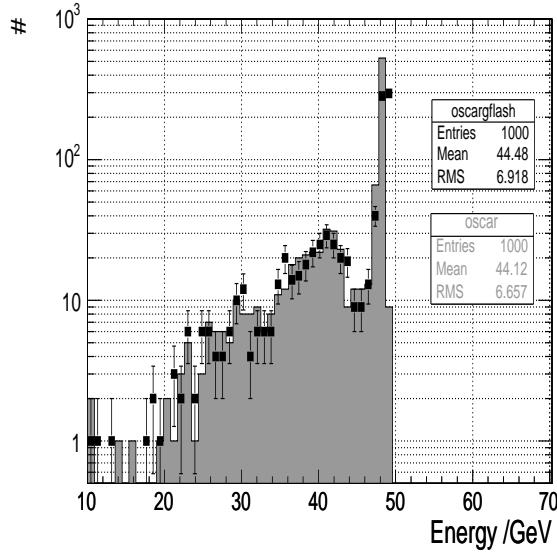
3.10.1   The GFlashTest example

This example demonstrates the usage of GFlash in OSCAR. It has been designed similar to the example in Geant4, i.e. also here the energy deposit in the central crystal, the 3×3 and the 5×5 matrix is compared for the full simulation (QGSP list) and for simulation with parametrised showers (QGSP+GFflash). This information, taking the crystal and module numbering scheme etc. correctly into account, is printed out for the energy deposits in the barrel and the endcap calorimeter and simultaneously written to a root file for further test and validation purposes. For similar reason, two files containing the expected reference output are included (Full.log and GFlash.log). In order to run the example one has to execute the command ExGFlash -c Full.orcarc for full simulation and ExGFlash -c GFlash.orcarc for parameterised showers in the electromagnetic calorimeter. A test n-tuple containing 20 events with one single electron with $p_T = 50$ GeV is also included (e50.ntpl). Finally, ShowerModelVolumes.xml contains the Geant4 volumes where the parameterisation is performed - basically the barrel and endcap volumes of the ECAL.

3.10.2   Performance for single electrons

The performance of the shower parameterisation has been tested in detail for single electrons in OSCAR on the full geometry. For this purpose, electrons with a flat distribution in rapidity were shot from the centre of the detector into the barrel and endcap region of the calorimeter. Comparisons between the GFlash-based and the full simulation in the central crystal, the $3{\times}3$ and $5{\times}5$ crystal matrices, show good agreement at the 1% level, as illustrated in Figure 3.21. The same tests were also performed with photons, which in this model are parameterised after they have converted into an electron-positron pair. The speed-up factor gained by the use of shower parameterisation is presented in Table 3.2.

The speed up is still significant. It is much smaller than in the toy setup with pure lead tungstate, since with a complex geometry and material in front of the calorimeter not all electrons are fully contained and parameterised, and produced secondaries are fully tracked, which slows down the simulation. Similar results have been obtained for photons. A problem encountered is the fact that inside a complex geometry it may be difficult to find an adequate parameterisation envelope, since besides the active material (here $PbWO_4$), also other volumes and materials may be contained in the mother volume. This possibility is neglected in the parameterisation (considering an effective material does not lead to an improvement) and has an

Figure 3.21: Energy depositions in the 5×5 crystal matrix for 50 GeV electrons. The histogram corresponds to the full OSCAR simulation and the markers to the shower parameterisation with GFlash.

Energy	Speed-up factor
50 GeV	2.0
100 GeV	3.8
500 GeV	11.8
1000 GeV	7.0

Table 3.8: Speed-up for single electrons in OSCAR 5.0.0.

impact on the agreement between full and fast simulation. If one shoots an electron directly into the CMS calorimeter starting from the crystal surface the difference between full and parameterised simulation is below 1%, well in agreement with the results obtained on the 'toy' model. If, on the other hand, the electron comes from the detector centre (without simulating the tracker) the difference increases to 2 - 3 %. This is most likely due to the non ideal shower envelope. The parameterisation deposits the energy assuming that the whole envelope consists of $PbWO_4$, which leads to energy losses if this is not the case. Full tracking handles this circumstance correctly, since the electron is basically propagated through the (usually not very dense and thin) support material and is still able to deposit energy in the active material. This problem is currently under consideration. The new region based parameterisation concept in Geant4.8 should improve the situation, since it allows for a more accurate description of the envelope.

3.10.3 Performance for full LHC events

Physics studies for the LHC require a large number of simulated proton-proton collisions. So the relevant question from the CMS physics point of view has been how much parameterised showers can speed up the complete simulation of full events, where only electromagnetic subshowers are parameterised. The speed up will depend on the event topology and the final state particles. As a sample, two event types were studied where the contribution from electromagnetic showers dominate the simulation time: the decay of a Higgs boson to 4 electrons and the production of a high energetic photon together with a graviton in the ADD Large Extra Dimension model, which is analysed in this thesis in chapter 5. The results can be seen in Table 3.9. The tests have shown that the shower parameterisation technique can significantly speed up full events. The numbers are comparable to those found in ATLAS using parameterised showers in the liquid argon sampling calorimeter. A

Physics process	Speed-up factor
Higgs $\rightarrow 4\,e$, $m_H = 300$ GeV	2.0
γ + graviton ($p_T^\gamma > 400$ GeV)	3.3

Table 3.9: Speed-up of full LHC events in OSCAR 5.0.0

better physics performance may be achieved by using the new Geant4 region concept once the CMS simulation will be ported to Geant 8.0, since then a more accurate definition of the shower envelope is possible. Concerning the timing performance, there is still some room for improvement left. One possibility is to increase the production cuts in Geant4. Doing so will result in a decreased production of low energetic photons and electrons which are at the moment fully tracked and account for a significant fraction of the total simulation time. The accuracy of the full tracking is not necessarily needed if the detector response is already approximated by a parameterised shower - this idea is currently under consideration in ATLAS and CMS.

The experience with the shower parameterisation framework in CMS were reported to the Geant4 developpers team in several Geant4 technical board meetings and on the Geant4 workshop in 2004. The results obtained so far were presented at two conferences [5, 4] and included in the CMS Physics Technical Design Report volume 1 [6].

3.11 Outlook

CMS is currently rewriting its simulation framework (CMSSW [45]). In September 2006 the GFlash package has been successfully ported to the new framework. In CMSSW a comparison with the test beam is planned as well, which would offer the

possibility to tune to real data. This has not been possible in the old framework due to technical reasons. The speed-up for full events can be further increased by using the parametrisation as well for the electromagnetic shower inside the hadronic calorimeter. This option is currently under investigation by a newly formed shower working group, which uses the current GFlash implementation as a reference and a starting point for further shower parameterisations projects. Finally, GFlash is expected to be used soon in large scale Monte Carlo productions in the new CMS framework. Here it can prove its full potential to save computing ressources and to be easily tunable to full simulation and real data.

Chapter 4

The Standard Model and Beyond

The Standard Model of Particle Physics (SM) has been established during the last few decades as our best available theory of subatomic particles and their interaction. It describes all existing accelerator data and has been tested to very high precision. In this chapter first the Standard Model is briefly outlined, followed by a discussion of its shortcomings and problems. Then, a review of possible alternatives beyond the Standard model is given, focusing on the class of models that has been analysed in this thesis.

4.1   The Standard Model

The SM asserts that all matter in our world is made up of "fermions", i.e. particle with spin 1/2, interacting through fields. The fermions are divided into two classes: six so called "leptons" and six "quarks". Particles that are made of quarks, like for example protons, are called "hadrons". The particles of matter can be sorted into three generations (corresponding to the 3 columns of Table 4.1 on the left). Each generation has the same quantum numbers, i.e. the four fermions in each generation behave almost exactly like their counterparts in the other generations; the only difference lies in their masses. For example, the electron and the muon both have half-integer spin, unit electric charge and do participate in the same interactions, but the muon is about 200 times more massive than the electron.

The two members of every family differ by one unit of electric charge: the leptons on the left in Table 4.1 carry the charge 0 in the top row and -1 in the bottom row, the quarks carry the charge 2/3 in the top row and -1/3 in the bottom row. The quarks also possess another kind of charge - the colour charge, which is relevant for the force which binds them together inside nuclear particles. All stable matter is built from the first generation of charged fermions (u,d,e); the higher-generation particles decay quickly into the first-generation ones and can only be generated for a short time in high-energy experiments. In the Standard Model there are three kinds of interactions between the particles of matter: the electromagnetic, weak and

ν_e	ν_μ	ν_τ	leptons	
e	μ	τ		particles of matter
u	c	t	quarks	
d	s	b		

γ		g	$W^\pm, Z$	
electro-magnetic		strong	weak	gauge particles

Table 4.1: Particles of matter and interaction in the Standard Model.

strong interaction.

The Standard Model of Particle Physics is a relativistic quantum field theory. In such theories, each type of interaction has a characteristic set of force carrier particles associated with quantum excitation of the force field related to that interaction. The carrier particles either appear in intermediate stages or are produced during all processes involving that type of interaction. Forces between particles can be described in terms of static force fields and exchanges of force by carrier particles, which are always bosons, i.e. particles with a spin of an integer number (e.g 0,1,2 ..).

The quanta of the electromagnetic interaction between electrically charged particles are the massless photons (γ). Since they are massless, the range of electromagnetic interactions is infinite. In contrast, the quanta of the weak interaction fields, namely the charged W^+, W^- and the neutral Z boson, are massive (~ 100 GeV) and consequently the weak interaction is short ranged ($\sim 10^{-17}$ cm). The quanta of the strong interaction which acts between colour-charged quarks are called gluons (g) and have zero mass. However, unlike the photons, they do not have an infinite range, since they carry a colour charge and interact with each other. This leads to the so called confinement which restricts the strong force to nuclear distances (R $\sim 10^{-13}$ cm) . The bosons in Table 4.1 are called "gauge" bosons, because in the Standard Model the force fields mentioned above are a consequence of how the terms describing particles (e.g. wave function) behave under certain transformations, so called "gauge" transformations. The central point is whether symmetries do exist and whether therefore physical quantities are conserved. The Lagrangian of each set of mediating bosons is invariant under a gauge transformation, that is why these mediating bosons are referred to as gauge bosons. It turns out that the gauge transformations of the gauge bosons can be exactly described using a unitary group called a "gauge group". The gauge group of the strong interaction is $SU(3)$

and the gauge group of the electroweak interaction is $SU(2) \times U(1)$. Therefore, the Standard Model is often referred to as a $SU(3) \times SU(2) \times U(1)$ group .

A particle of central importance in the context of the SM is the Higgs boson, a hypothetical massive scalar elementary particle. It is the only Standard Model particle not yet observed; it plays a key role in explaining the origin of the mass of other elementary particles, in particular the difference between the massless photon and the very heavy W and Z bosons.

The Higgs boson was first predicted in 1964 by the British physicist Peter Higgs [1, 2], Brout and Engelbert [3] and independently by others. The particle called Higgs boson is in fact the quantum of one of the components of a Higgs field. In empty space, the Higgs field acquires a non-zero value, which permeates every place in the universe at all times. This vacuum expectation value (VEV) of the Higgs field is constant and equal to 246 GeV. The acquisition of a non-zero VEV spontaneously breaks the electroweak gauge symmetry, a phenomenon known as the Higgs mechanism. Unfortunately, the Standard Model does not predict the value of the Higgs boson mass. As of 2006, no experiment has directly detected the existence of the SM Higgs and measured its mass. Within the Standard Model, the non-observation of clear signals at particle accelerators leads to an experimental lower bound for the Higgs boson mass of 114.4 GeV at 95% confidence level [78]. It is expected among physicists that the Large Hadron Collider, currently under construction at CERN, will be able to confirm or deny the existence of the Higgs boson. Precision measurements of electroweak observable indicate that the Standard Model Higgs boson mass has an upper bound of 166 GeV at the 95% confidence level as of July, 2006 [79]. A detailed review of the Standard Model and its central ideas can be found for example in [80].

4.2 Limitations of the Standard Model

The Standard Model is in remarkable agreement with particle physics data at the energy scales probed so far. In spite of this success there are, however, some problematic aspects and limitations. The Standard model has about twenty free parameters that are not fixed by the gauge principles. These are the strengths of gauge couplings, Yukawa couplings of quarks and leptons, mixing angles and the mass of the Higgs particle. The development of particle physics in the past twenty years was marked by the accurate determination of most of these parameters. Besides the fact that the Higgs boson - a central ingredient of the Standard Model - has still not been detected, the following aspects remain problematic:

- **Gravitation:** The Standard Model does not include gravitational interactions and is therefore an incomplete description of the forces we know.

- **Neutrino Masses:** Neutrinos are massless in the SM. Recent experiments

have prooven however, that the neutrino mass is indeed very small, but apparently non-zero. For instance, in 1998 Super-Kamiokande published results showing neutrino oscillation [81]. In the Standard Model, a massless neutrino cannot oscillate; some possibilities to accommodate neutrino masses in the Standard model have been discussed for example in [82].

- **Dark Matter:** Baryonic matter is not the only type of matter in the Universe. Astrophysical observations provide evidence for the existence of nonrelativistic, neutral, non baryonic dark matter. The direct evidence for the presence of dark matter are the rotation curves of galaxies. To explain these curves, one has to assume the existence of a galactic halo made of non-shining matter, which takes part in the gravitational interaction. According to the latest data, the matter content of the universe Ω is the following [83]:

$$\Omega_{\text{Vacuum}} \sim 73\% \,, \ \Omega_{\text{Dark Matter}} \sim 23\%, \Omega_{\text{Baryon}} \sim 4\% \qquad (4.1)$$

  Hence there is almost six times more dark matter than usual matter in the Universe. Since neutrinos are not massive enough to explain these observations, a new candidate for dark matter is needed.

- **Unification of Gauge Couplings:** If one believes that the three elementary interactions in the SM result from a local gauge theory with a higher fundamental symmetry, the coupling strengths of the strong, weak and electromagnetic interactions should unify at some energy scale. This idea is based on prominent examples from the past: In the 19th century experimental evidence that electric and magnetic interactions were not independent phenomena but two manifestation of a single electromagnetic interaction lead Maxwell to a common description of both observed phenomena. Also in the Standard Model the electromagnetic and the weak interaction were merged into the electroweak interaction. The coupling "constants" of the interactions are in fact not constant in the SM, but a function of the energy scale at which the interaction takes place. It turned out that the evolution of the couplings - assuming the SM particle content - does not lead to a unification in a single point. True unification can only be obtained if New Physics (NP) enters between the electroweak and the Planck scale modifying the slope of the evolution of the couplings.

- **Hierarchy Problem:** The hierarchy problem in particle physics is the question why the weak force is 10^{32} times stronger than gravity. Both of these forces involve constants of nature, Fermi's constant for the weak force and Newton's constant for gravity. More precisely, from the speed of light c, Planck's constant h and Newton's constant G one can form the Planck mass:

$$m_P = \sqrt{\frac{hc}{G}} \sim 10^{19}\,\text{GeV} \qquad (4.2)$$

This is a fundamental scale in nature, at which quantum gravitational effects become important. On the other hand, the mass scale of the electroweak symmetry breaking is set by m_W, m_Z and m_h, which is (or is expected to be in the case of the Higgs boson) around 100 GeV. Why is m_W so much smaller than m_P? This is not just an aesthetic question; It leads, more technically speaking, to the question, why the Higgs boson is so much lighter than the Planck mass, although one would expect that the large, quadratically divergent, radiative corrections to the square of the Higgs boson mass would inevitably make the mass huge, comparable to the scale at which New Physics appears unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare mass. Given this hierarchy problem with the Higgs boson mass, it is expected that New Physics should show up at an energy scales not much higher than the scale of energy required to produce the Higgs boson, and thereby provide an explanation for its small mass.

These questions - together with the origin of the symmetry breaking - have led many physicists to believe that the Standard Model is an effective theory valid up to some energy scale $\Lambda \approx$ TeV. That is, at some higher energy scale, it will be incorporated in an even more fundamental theoretical framework not yet known. The search for this transition will be a grand task of particle physics of the 21st century.

4.3 Scenarios Beyond the Standard Model (BSM)

In view of the open questions listed above, a multitude of scenarios to extend the SM and solve its problematic or uncovered aspects has been proposed in the last years. In the following the most popular ideas are listed and the model class relevant for this thesis is discussed in more detail. A detailed discussion of BSM scenarios is not possible in the scope of this thesis; it can be found in literature and in the references given below. While the most popular examples of theoretical possibilities are widely known in the particle physics community, one must keep in mind that nature may prove to be far more creative than we are and that something completely unexpected may be discovered.

4.3.1 Supersymmetry

The most popular theory- but not the only proposed - to solve the listed problems is Supersymmetry. Supersymmetry (often abbreviated SUSY) is a theory which proposes a symmetry between bosons and fermions, i.e. in supersymmetric theories, every fundamental fermion has a bosonic superpartner and vice versa. SUSY is strongly motivated by the fact that it can give true unification of the gauge interactions [84] at a high scale not far from the scale of gravity, the Planck scale. In

addition, supersymmetric extensions of the SM provide a solution of the hierarchy problem in the Higgs sector. They predict a light Higgs particle in the context of GUTs [85], in contrast with the SM. To incorporate supersymmetry into particle physics, the Standard Model must be extended to include at least twice as many particles, since there is no way to obtain a fermion-boson-symmetry between the particles in the Standard Model. The simplest possible supersymmetric model consistent with the Standard Model is the Minimal Supersymmetric Standard Model (MSSM). However, the MSSM appears to be unnatural in a number of ways, and many physicists doubt that it will be the correct theory. It yields a prediction of the Weinberg angle in agreement with present experimental measurements. Moreover, it does not exhibit any quadratic divergences, in contrast with the SM Higgs sector. The lightest supersymmetric particle (LSP) offers a proper candidate for the cold Dark Matter content of the universe, if $R - parity$ (a new quantum number which is -1 for supersymmetric partners and 1 for SM particles) is conserved. Until now, only lower mass limits have been set on supersymmetric particles. The LHC is supposed to cover a wide range of parameters of the MSSM and will be a crucial test for the MSSM and low-energy SUSY. Typical decay signatures of supersymmetric particles usually contain missing transverse momentum from the LSP escape plus multiple jets and a varying number of leptons. The LHC potential to discover SUSY is widely discussed in literature. Supersymmetry is not further discussed in this thesis.

4.3.2 Extra Dimensions

A very natural question that comes up when one thinks about the universe is why does it have three spatial and one time dimensions. In search for an answer, there has been an explosion of research activity over the last few years on theories with extra dimensions. The idea of having extra spatial dimensions is strongly motivated by String theory. The basic idea behind all string theories is that the fundamental constituents of reality are strings of extremely small scale (possibly Planck length, about 10^{-35} m) which vibrate at specific resonant frequencies. Thus, any particle should be thought of as a tiny vibrating object, rather than as a point. This object can vibrate in different modes (just like a guitar string can produce different notes), with every mode appearing as a different particle (electron, photon etc.). Strings can split and combine, which would appear as particles emitting and absorbing other particles, presumably giving rise to the known interactions between particles. The interest in String theory is driven largely by the hope that it will prove to be a consistent theory of quantum gravity or even a theory of everything. It can also naturally describe interactions similar to electromagnetism and the other forces of nature. Consistent quantisation of strings only appears to be possible with supersymmetry and with extra degrees of freedom. All String theories predict the existence of these degrees of freedom which are usually described as extra dimen-

sions. String theory is thought to include some 10, 11 or 26 dimensions, depending on the specific theory. String theory as a whole has not yet made falsifiable predictions that would allow it to be experimentally tested, though various planned observations and experiments could confirm some essential aspects of the theory, such as supersymmetry and extra dimensions. In this context it is important to keep in mind, that supersymmetry and extra dimensions are not mutually exclusive and that String theory needs them both. The most actively pursued program is superstring (or M-) theory. An overview of String theory can be for example found in [86]. Models with Extra Dimensions will be now discussed in more detail.

4.4   Models with Extra Dimensions

The idea of extra spatial dimensions is not totally new. In fact, extra dimensions were already proposed at the beginning of the 20 th century in the work of Gunnar Nordström (1914) [87], Theodor Kaluza (1921) [88] and Oscar Klein (1926) [89], who were trying to extend general relativity in order to unify gravity and electromagnetism within a common geometrical framework. They proposed that unification of the two forces occurred when spacetime was extended to a five dimensional manifold. A difficulty with the acceptance of these ideas in the 1920s was the lack of experimental implications. The early 1980s lead to a revitalisation of the concept, mainly due to the realization that a consistent string theory will necessarily include extra dimensions. An additional motivation is the fact that the behaviour of gravity has not yet been measured down to about more than a fraction of one milimeter in laboratory experiments. The hypothesis of n extra dimensions similar to the three we are familiar with would modify the law of gravity, which is:

$$F \sim \frac{1}{r^2} \tag{4.3}$$

This behaviour, in virtue of Gauss theorem, is possible if we live in three spatial dimensions. Otherwise the relation would be:
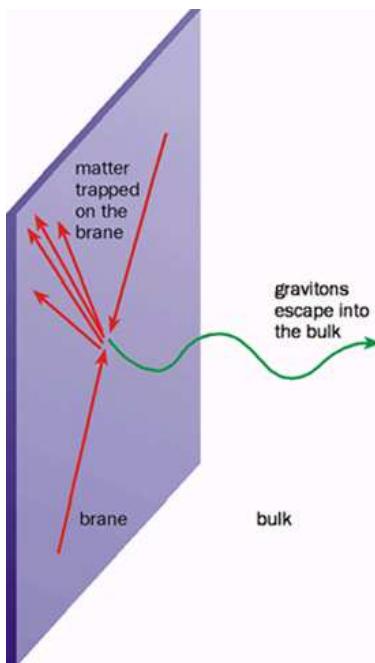
$$F \sim \frac{1}{r^{n+2}} \tag{4.4}$$

The behaviour on distances $<< 1$ mm is not known and allows for the possibility that at higher energies gravity behaves quite differently than expected. Recently, concepts developed within string theory have led to new phenomenological ideas which relate the physics of extra dimensions to observable in a variety of physics experiments and address the hierarchy problem. In the next section, first the basic concepts and terminology of Extra Dimensions models are introduced.

4.4.1   Basic theory concepts

In general, theoretical frameworks with extra dimensions have some common properties. In most scenarios, our observed 3-dimensional space is a 3-brane (sometimes called a wall), where the terminology is derived from a generalisation of a 2-dimensional membrane. This 3-brane is embedded in a higher D-dimensional spacetime, $D = 3 + n + 1$, with n extra spatial dimensions which are orthogonal to our 3-brane. The higher n-dimensional space is referred to as the "bulk". Models with Extra Dimensions can be divided into two main groups depending on the assumptions made for the metric in the D-dimensional space:

- Assuming a factorisable geometry, i.e. the metric in the four usual dimensions is independent of the position in the extra dimensions, one has the *flat compactified extra dimensions scenario.*

- Assuming a non factorisable geometry, i.e. the metric in the four usual dimensions depends on the position in the extra dimensions, one has the *warped extra dimensions scenario.*
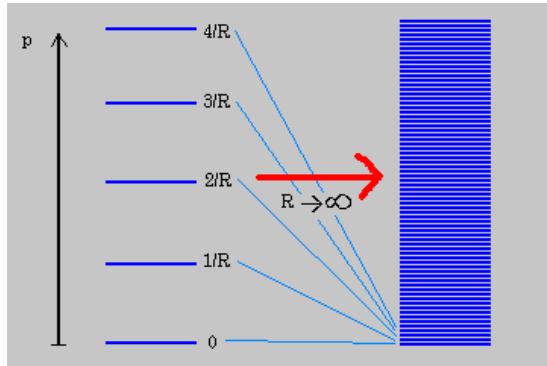
Figure 4.1: The braneworld scenario: the SM fields are trapped on a brane and only gravity spreads throughout the full $3 + n$ space.

A further distinction can be made between models in which all or some the particles are allowed to propagate into the extra dimensions and models in which the extra dimensions are accessible only to gravity. The picture is thus one where matter and

gauge forces are confined to our 3-dimensional subspace, while gravity propagates in a higher dimensional volume. In this case, the Standard Model fields maintain their usual behaviour, however, the gravitational field spreads throughout the full $3 + n$ spatial volume, as is indicated in Figure 4.1.

The additional dimensions can not be too large, since this would result in observable deviations from Newtonian gravity. The extra dimensional space is therefore often required to be *compactified*, *i.e*, made finite. However, in some alternative theories [90, 91], the extra dimensions are infinite and the gravitational deviations are suppressed by other means. The field content which is allowed to propagate in the bulk, as well as the size and geometry of the bulk itself, varies between different models. If one assumes compactification, fields propagating in the bulk expand into a series of states known as a Kaluza-Klein (KK) tower, with the individual KK excitations being labelled by mode numbers, as shown in Figure 4.2.

Figure 4.2: The Kaluza Klein tower of states: if the additional dimensions are infinite instead of being compactified ($R \longrightarrow \infty$), the n-dimensional momentum and resulting KK spectrum is continuous.

To understand this, it is instructive to make an analogy with a simple example we are familiar with: If we recall the Schrödinger Equation for a free particle moving along the x direction, the solution - since the x direction is infinite, i.e., noncompact - is just $\sim A_n e^{ipx} + B_n e^{ipx}$ where p is the particle momentum which can take on an infinite set of continuous values. In this case the momentum p is not quantised and this is due to the fact that the space is *non compact*. The situation is different if we consider a particle in a box, i.e., a situation where the potential is zero for $0 \leq y \leq L\pi$ but infinite elsewhere so that the wavefunction vanishes outside this region. Since the physical region is of a finite size it is called *compact*. The solution inside the box takes the same general form as does the case of a free particle but it must also vanish at the boundaries. These boundary conditions tell us $A\prime$ and $B\prime$ so that the solutions actually takes the form $\sim \sin(\text{ky/L})$ and that the momenta are quantised, i.e., $p = k/L$ with $k = 1, 2, ...$ These two situations are completely

analogous to having a 5th dimension which is either infinite (noncompact) or finite (compact) in size. Most extra dimension models assume that extra dimensions are compact. For a flat 5th dimension of length L the KK masses are large if the size of the extra dimension is small. In fact, the observation of KK excitations is the hallmark of EDs. It is interesting to observe that there are no solutions in the 'particle in a box' example corresponding to massless particles, i.e., those with $k = 0$, the so-called zero modes, which do exist in the KK tower. In the simplest case, a real scalar field in 5D with one extra dimension y compactified over a circle with radius R, the compactification of the extra dimension leads to one real massless scalar field field for $k = 0$ (the zero mode) and an infinite number of massive complex scalar fields (Kaluza Klein modes) with the mass values $m_k = k^2/R^2$. In the case of fifth dimensional gravity, a spin 2 field, each quanta of momentum in the compactified volume appears as a KK excited state with mass $m^2 = \vec{p}_k^2$ for an observer trapped on the brane. This builds a KK tower of states, where each state carries identical spin and gauge quantum numbers. If the additional dimensions are infinite instead of being compactified, the n-dimensional momentum and resulting KK spectrum is continuous (in analogy to the particle example above) as it is schematically drawn in Figure 4.2.

More technically, in the case where gravity propagates in a compactified bulk, one starts from a D-dimensional Einstein-Hilbert action and performs a KK expansion about the metric field of the higher dimensional spacetime. The graviton KK towers arise as a solution to the linearised equation of motion of the metric field in this background. The resulting 4-dimensional fields are again the Kaluza-Klein modes. The KK zero-mode fields are massless, while the excitation states acquire mass by 'eating' lower spin degrees of freedom (similar to the Higgs-Goldstone mechanism). This results in a single 5-component tensor KK tower of massive graviton states, $n - 1$ gauge KK towers of massive vector states and $n(n - 1)/2$ scalar towers.

As mentioned before, there are many models and scenarios with extra dimensions; a more detailed review can for example be found in [92], [93].

In the next section the most prominent, mature and 'LHC friendly' (i.e. models that predict detectable signatures at the LHC) extra dimension models are summarised.

4.4.2 TeV^{-1}-sized Extra Dimensions

In this class of models one assumes that not only gravity, but also SM fields could live in an experimentally accessible higher dimensional space [94]. This hypothesis could lead for example to unification of gauge couplings at a low scale. In contrast with the case where only gravity can probe the extra dimension, here the extra dimension must be at least as small as about TeV^{-1} in order to avoid incompatibility with experimental results. The fact that the Standard Model fields are phenomeno-

logically allowed to propagate in the bulk in these scenarios, presents a wide variety of choices for model building: (i) all, or only some, of the Standard Model gauge fields exist in the bulk; (ii) the Higgs field may lie on the brane or in the bulk; (iii) the Standard Model fermions may be confined to the brane or to specific locales in the extra dimension. The phenomenological consequences of this scenario strongly depend on the location of the fermion fields. An interesting possibility to explain the observed spectrum of quarks and lepton masses is to assume that different fermions are in different points of the extra dimension. Their different overlap with the Higgs wavefunction can generate a hierarchical structure of the Yukawa couplings [95]. The case in which all SM particles uniformly propagate in the bulk of an extra dimensional space and no branes need to be present is referred to as universal extra dimensions and is described in the next section.

4.4.3 Universal Extra Dimensions (UED)

'Universal Extra Dimensions' is the name given by Appelquist, Cheng, and Dobrescu [96] to a class of models which closely resemble the original Nordström-Kaluza-Klein scenario, but with some crucial improvements. All particles live in the full bulk, which is compactified to some kind of orbifold - branes need not to be present. The simplest case is a single extra dimension with coordinate y, compactified to a circle, which in turn is orbifolded to a line interval of length L by identifying points under $y \to -y$.

In the original Kaluza-Klein model the Kaluza-Klein mode number is conserved, as this is just conservation of (discrete) momentum in the extra dimension. However, in our simple UED example momentum conservation in the fifth dimension is replaced by a conserved parity, called "KK parity". This is enough to guarantee that the lightest massive KK mode in a UED model is stable. The situation is quite analogous to R parity in SUSY models. As with SUSY, this implies that in UED models the first massive KK modes must be produced in pairs. It also means that the lightest massive KK mode, the "LKP", is a good dark matter candidate [97, 98]. If the LKP is a major constituent of dark matter, then it is in a mass-coupling range such that it will be produced at the LHC. The other first massive KK modes will decay promptly to this LKP. Thus typical UED events at the LHC give a variety of jet and lepton signatures combined with large missing transverse energy ($\not{E}_T$). If only the first massive KK modes are produced, UED models look very much like a subset of SUSY models, in terms of their collider signatures. The crucial discriminators are the spins of the heavy partner particles. Distinguishing these spins is a very significant experimental challenge [99]. Since the KK states are allowed to be relatively light, they can produce observable effects [100], [101], [102], [103] in loop-mediated processes, such as $b \to s\gamma$, anomalous magnetic moment of the muon, and rare Higgs decays The current limits on the UED scale $1/L$ is between 300 GeV

and 600 GeV, depending on the Higgs mass [104].

4.4.4   Randall-Sundrum model (RS)

Randall-Sundrum refers to a class of scenarios, also known as warped extra dimensions models, originated by Lisa Randall and Raman Sundrum [105], [90]. In these scenarios there is one extra spatial dimension, and the five-dimensional geometry is "warped" by the presence of one or more branes. The branes extend infinitely in the usual three spatial dimensions, but are sufficiently thin in the warped direction that their profiles are well-approximated by delta functions in the energy regime of interest. Most collider physics phenomenology done with warped extra dimensions so far is based upon one very specific model, the original simple scenario called RSI. The simplest such framework comprises just one additional spatial dimension of finite size, in which gravity propagates. The geometry is that of a 5-dimensional Anti-de-Sitter space (AdS$_5$), which is a space of constant negative curvature. The extent of the 5th dimension is $y = \pi R_c$. Every slice of the 5th dimension corresponds to a 4-d Minkowski metric. Two 3-branes, with equal and opposite tension, sit at the boundaries of this slice of AdS$_5$ space. The Standard Model fields are constrained to the 3-brane located at the boundary $y = \pi R_c$, known as the TeV brane, while gravity is localised about the opposite brane at the other boundary $y = 0$. This brane is referred to as the *Planck* brane. The metric for this scenario has the form:

$$ds^2 = e^{-2ky}\eta_{\mu\nu}dx^\mu dx^\nu - dy^2\,, \tag{4.5}$$

where x^μ indicates the four usual dimensions, y the extra dimension and $\eta_{\mu\nu}$ is the Minkowski metric. This type of geometry is called *non-factorisable* because the metric of the 4D subspace is y dependent, due to the exponential function of the 5th dimensional coordinate multiplying the usual 4-dimensional Minkowski term. This exponential is known as the warp factor.

The relation

$$M_{\text{Pl}}^2 = \frac{M_5^3}{k} \tag{4.6}$$

is derived from the 5-dimensional action. Equation 4.6 indicates that in the RS philosophy all dimensionful parameters in the action have their mass scale set by $k \sim M_5 \sim M_{\text{Pl}}$, so that there is no fine-tuning and no additional hierarchies present in this model. However, the warp factor rescales them as one moves about in y so that, in particular, all masses will appear to be of order the TeV scale on the SM brane, i.e., to us. This means that if there is some mass parameter, m, in the action which is order M_{Pl}, we on the TeV brane will measure it to be reduced by the warp factor. If $kR_c \sim 11$ (a small hierarchy) this exponential suppression (warp factor) reduces a mass of order 10^{18} GeV to only 1 TeV. Thus the ratio of the weak scale to the Planck scale is explained through an exponential factor and no large ratios appear anywhere else in the model. The hierarchy is naturally established

by the non-trivial configuration of the gravitational field: the zero-mode graviton wave function is peaked around the *Planck* brane and it has an exponentially small overlap with the *TeV* brane where we live. Since $kR_c \simeq 10$ and it is assumed that $k \sim 10^{18}$ GeV, this is not a model with a large extra dimensions in contrast to the ADD model presented in the next section. Since the Standard Model fields live on the TeV brane the phenomenology of RSI - similar to the ADD model - is concerned with the effects of the massive KK modes of the graviton. The massless zero-mode of the gravitons couples with the usual $1/M_{Pl}$ strength, the coupling of the KK excitations is comparable to $1/M_{EW}$. At the LHC the KK gravitons of RSI would be seen as di-fermion or diboson resonances, since the coupling of each KK mode is only TeV suppressed [106]. The most recent experimental constraint comes from the D0 experiment at FNAL/ Tevatron, which does exclude KK gravitons up to 250 GeV for $c = 0.01$ (with $c = k/M_{Pl}$) and 785 GeV for $c = 0.1$. [107, 108]

4.4.5 Arkani-Hamed, Dvali and Dimopoulos model (ADD)

ADD is the name of the class of models which incorporate the large extra dimensions scenario of Arkani-Hamed, Dvali, and Dimopoulos [7]. These were the first extra dimensions models in which the compactified dimensions can be of macroscopic size, consistent with all current experiments and observations. For this reason they are sometimes referred to as "large extra dimensions" models. In the most basic version, n extra spatial dimensions are compactified on a torus with common circumference R, and a brane is introduced which extends only in the three infinite spatial directions. It is assumed, that the Standard Model gauge and matter fields are confined to the 3-dimensional brane that exists within the higher dimensional bulk. Gravity alone propagates in the n extra spatial dimensions.

Gauss' Law relates the Planck scale of the effective 4-dimensional low-energy theory, M_{Pl}, to the scale where gravity becomes strong in the $4 + n$-dimensional spacetime, M_D, through the volume of the compactified dimensions via:

$$M_{\mathrm{Pl}}^2 = M_D^{2+n} R^n \quad , \tag{4.7}$$

where M_{Pl}^2 is defined by Newton's constant: $M_{\mathrm{Pl}} = 1/\sqrt{G_N} = 1.2 \times 10^{19}$ GeV. M_D^{2+n} is defined as the gravitational coupling which appears in the $4+n$-dimensional version of the Einstein-Hilbert action. It is the quantum gravity scale of the higher dimensional theory, i.e. the scale where gravity becomes strong.

If M_{Planck}, M_D and $1/R$ are all of the same order, as is usually assumed in string theory, this relation is not very interesting. But there is nothing which prevents us from assuming that M_D is equal to some completely different scale. Most attractive is to take $M_D \sim 1$ TeV. M_{Pl} is then generated by the large volume of the higher

Number of EDs	Compactification Radius	Compactification scale
$n = 2$	$R \sim 0.1$ mm	$R^{-1} \sim 10^{-3}$ eV
$n = 3$	$R \sim 10^{-7}$ cm	$R^{-1} \sim 100$ eV
...	...	...
$n = 6$	$R \sim 10-12$ cm	$R^{-1} \sim 10$ MeV

Table 4.2: Size of the extra dimensions for different ADD scenarios.

dimensional space and is thus no longer a fundamental scale. The hierarchy problem is now translated to the possibly more tractable question of why the compactification scale of the extra dimensions is large. While in the Randall-Sundrum model (RSI) the hierarchy is explained by the non-trivial configuration of the gravitational field, in the ADD framework gravity is so weak, because it is diluted in a larger space $(R >> M_D^{-1})$. A complete solution would require an explanation why R is stabilised at such a large value. How large can the extra dimension be ? For $M_D \sim$ TeV, the radius R of the extra dimensions ranges from a fraction of a millimetre to ~ 10 fermi for n varying between 2 and 6. The compactification scale $(1/R)$ associated with these parameters then ranges from $\sim 10^{-3}$ eV to tens of MeV (see Table 4.2). The case of one extra dimension is excluded as the corresponding dimension (of size $R \approx 10^{11}$ m) would directly alter Newton's law at solar-system distances. Our knowledge of the electroweak and strong forces extends with great precision down to distances of order 10^{-15} mm, which corresponds to $\sim (100 \text{ GeV})^{-1}$. Thus the Standard Model fields do not feel the effects of the large extra dimensions present in this scenario and must be confined to the 3-brane. Therefore in this model only gravity probes the existence of the extra dimensions.

The ADD scenario raises the exciting possibility of observing quantum gravity at the LHC. If extra dimensions are present and quantum gravity becomes strong at the some TeV scale, then observable signatures at colliders operating at the TeV scale of the interactions of the bulk graviton with the Standard Model fields are expected. There is again a Kaluza-Klein tower of graviton modes, where the massless mode is the standard 4dimensional graviton, and the other KK modes are massive spin 2 particles which also couple to SM matter with gravitational strength. In the ADD scenario, there are $(ER)^n$ massive Kaluza-Klein modes that are kinematically accessible in a collider process with energy E. For $n = 2$ and $E = 1$ TeV, that totals 10^{30} graviton KK states which may individually contribute to a process. It is the sum over the contribution from each KK state which removes the Planck scale suppression in a process and replaces it by powers of the fundamental scale $M_D \sim$ TeV. The interactions of the massive Kaluza-Klein graviton modes can then be observed in collider experiments either through missing energy signatures or

through their virtual exchange in Standard Model processes [1]. The strategy to deduce the couplings of the gravitons to matter and the corresponding Feynman rules are catalogued in [109],[110]. If M_D is close the weak scale, black hole formation is expected as well at the LHC with a large production rate [111, 112]. For example, the production cross section of 6 TeV black holes is about 10 pb, for M_D =1.5 TeV. The produced black-hole emits thermal radiation with the Hawking temperature $T_H = (n+1)/(4\pi R_s)$, with R_s being the Schwartzschild radius. A black hole of initial mass M_{BH} completely evaporates with the lifetime $\tau \sim M_{BH}^{(n+3)/(n+1)}/M_D^{2(n+2)/(n+1)}$, which is typically $10^{-26} - 10^{-27}$ s for $M_D = 1$ TeV.

Graviton production at colliders in the ADD model

The first class of collider processes involves the real emission of Kaluza-Klein graviton states in the scattering processes $e^-e^+ \to \gamma(Z) + G$, and $p\bar{p} \to$ jet + G or $p\bar{p} \to \gamma+G$. The produced graviton behaves as if it were a massive, non-interacting, stable particle and thus appears as missing energy in the detector. The cross section is computed for the production of a single massive KK excitation and then summed over the full tower of KK states. Since the mass splittings between the KK states is so small (compared to the centre-of-mass energy at the LHC) the sum over the states may be replaced by an integral weighted by the density of KK states. The specific process kinematics cut off this integral. The expected suppression from the $M_{\rm Pl}^{-1}$ strength of the graviton KK couplings is exactly compensated by a $M_{\rm Pl}^2$ enhancement in the phase space integration.

The cross section for on-shell production of massive Kaluza Klein graviton modes then scales as simple powers of $\sqrt{s}/M_D$,

$$\sigma_{KK} \sim \frac{1}{M_{\rm Pl}^2}(\sqrt{s}R_c)^n \sim \frac{1}{M_D^2}\left(\frac{\sqrt{s}}{M_D}\right)^n . \tag{4.8}$$

The exact expression may be found in [109, 113]. It is important to note that due to integrating over the effective density of states, the radiated graviton appears to have a continuous mass distribution; this corresponds to the probability of emitting gravitons with different extra dimensional momenta. The observables for graviton production, such as the γ/Z angular and energy distributions in e^+e^- collisions, are then distinct from those of other physics processes involving for the undetectable particles.

The emission process at hadron colliders, for example, $q\bar{q} \to g + G_n$, results in a monojet plus missing transverse energy signature. For larger numbers of extra dimensions the density of the KK states increases rapidly and the KK mass distribution is shifted to higher values. This is not reflected in the missing energy distribution: although the heavier KK gravitons are more likely to carry larger

---

[1]Bremstrahhlung of ordinary gravitons is a completely negligible effect at colliders
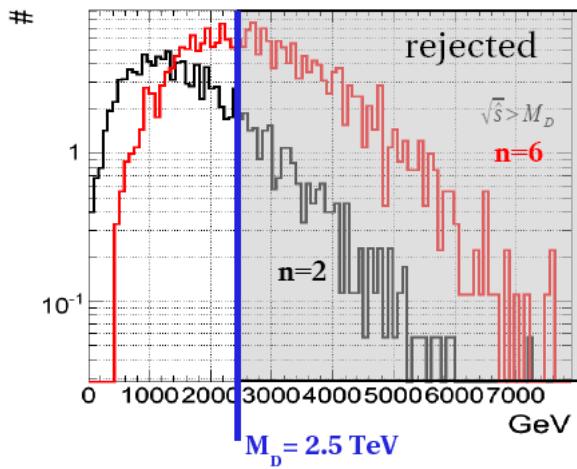
energy, they are also more likely to be produced at threshold due to the rapidly decreasing parton distribution functions. These two effects compensate each other, leaving nearly identical missing energy distributions.

The second class of collider signals for large extra dimensions is that of virtual graviton exchange [109, 114] in $2 \to 2$ scattering. This leads to deviations in cross sections and asymmetries in Standard Model processes, such as $e^+e^- \to f\bar{f}$. It may also give rise to new production processes which are not present at tree-level in the Standard Model, such as $gg \to \ell^+\ell^-$. The amplitude is proportional to the sum over the propagators for the graviton KK tower which may be converted to an integral over the density of KK states. However, in this case, there is no specific cut-off associated with the process kinematics and the integral is divergent for $n > 1$. This introduces a sensitivity to the unknown ultraviolet physics which appears at the fundamental scale. In order to regulate this integral several approaches have been proposed. The most model independent approach which does not make any assumptions as to the nature of the New Physics appearing at the fundamental scale is that of the naive cut-off. Here, the cut-off is set to $M_H \neq M_D$; the exact relationship between M_H and M_D is not calculable without knowledge of the full theory. As above, the Planck scale suppression is removed and superseded by powers of $M_H \sim$TeV. The resulting angular distributions for fermion pair production are quartic in $\cos\theta$ and thus provide a unique signal for spin-2 exchange. The $\gamma\gamma \to WW$ process has the highest sensitivity to graviton exchange. Studies to estimate the discovery potential for the ADD scenario of a Linear Collider can for example be found in [115].

Graviton production above the cutoff

At the LHC, proton–proton collisions will probe a distribution of partonic subprocess with different energies $\sqrt{\hat{s}}$. This creates a problem for the consistent analysis of missing energy signatures in the framework of ADD models. These models are simple, low energy effective theories which are only valid for $\sqrt{\hat{s}} > M_D$. This problem was first noted by the authors of [109], who suggested replacing the ADD graviton density of states $\rho(m)$ by $\rho(m)\theta(\sqrt{\hat{s}} - M_D)$, where θ is a step function. This introduces a systematic theoretical error into the analysis. The size of this error is very sensitive to the values of M_D and n. For initial LHC data sets, one would like to begin the analysis at the current $M_D \simeq 1$ TeV bounds from Tevatron and LEP. This increases the theory systematic from the cutoff for any fixed n, since the M_D value is rather low with respect to the partonic centre-of-mass energy. The effect of the cutoff is enormous for modest values of M_D and for high n values. This can be seen in Figure 4.3, showing the mass spectrum of the KK gravitons for $n = 2$ and $n = 6$: in the case of $n = 6$, more events are rejected by the cutoff than in the case of $n = 2$. For modest M_D values the theory error for the total cross section can be as large as an order of magnitude, since the shape of the graviton mass distribution does

not change with M_D; the parameter M_D only scales the production cross-section, which becomes lower for higher M_D values. The solution of this problem depends upon whether or not there is a signal in the missing energy channels. If there is a signal, the optimal procedure is to measure the observables $d^2\sigma/dp_T d\eta$ as accurately as possible, perhaps at more than one collider energy as suggested in [116, 117]. For the lower range values of M_D, the sensitivity to n suggested in [116, 117] will tend to be washed out. This is not a bad outcome, since it is a result of convolving the n dependence with the effects of strings, branes or other New Physics.

Figure 4.3: Mass spectra for KK gravitons for $n = 2, 6$; For n=6 the gravtion is heavier and the effect of the cutoff bigger.

More problematic is the case where there is no graviton signal in a given data set. Since in this case one would try to set a limit, we need an estimate of the theory systematic. The simplest possibility is to implement the cutoff defined above, and estimate the theory error by varying the cutoff. For ADD with $n \geq 6$, one expects to obtain no lower bound at all on M_D, as noted in [109]. The strategy in CMS [9] and in this analysis has been to use the step function θ as defined above. Other choices has been made as well, for example the ATLAS simulation [116] of the jet/photon and missing transverse energy signal uses the dumping factor $M_D^4/\hat{s}^2$ to suppress the events with $\sqrt{\hat{s}} > M_D$.

4.5 Present experimental status

The experiments and measurements investigating extra dimensions can be divided into three different categories: direct measurements of the gravitational interaction, astrophysical measurements and measurements at collider experiments. In the following the experimental limits for the parameter of the ADD scenario, M_D and n,

are given. At the moment there are no experimental evidences in favour of extra dimensions A more complete review of experimental results and limits can be found in [118, 119] and the references therein.

4.5.1 Constraints from direct measurements

Direct measurements of the deviation from the Newtonian behaviour at short range are mostly carried out using Cavendish-like detectors. As suggested in [120] the Newtonian gravitational potential can be replaced by a more general expression :

$$V(r) = -G\frac{m_1 m_2}{r}(1 + \alpha e^{-r/\lambda}) \qquad (4.9)$$

where α is a dimensionless strength parameter relative to the gravitational interaction and λ is a length scale, characterising the range of the new interaction. For the large extradimensional model the values of α and λ for a given mass scale M_D, depend upon the compactification scheme. For the compactification in a torus of size R this gives $\lambda = R$ and $\alpha = 2n$, where n is the number of extra dimensions. The most recent Cavendish-type experiment used a torsion pendulum and a rotating attractor. The limits on R are out of reach for $n > 2$, while $n = 1$ is already excluded for a mass scale M_D in the TeV region. For $n = 2$ the best current limits are $R < 150\,\mu m$ [121] and $R \leq 130\,\mu m$ [122] at 95% confidence level.

4.5.2 Constraints from astrophysical results

Astrophysical and cosmological considerations can impose constraints on theories of extra dimensions. The typical energy scale associated with such considerations is of the order of 100 MeV, consequently models that can produce KK states in this energy regime are restricted.

For the case of large extra dimensions of flat and toroidal form, the astrophysical bounds far surpass those from collider or short range gravity experiments for $n = 2$. If these large additional dimensions are compactified on a hyperbolic manifold instead, then the astrophysical constraints are avoided as the modified spectrum of KK graviton states admits for a first excitation mass of order several GeV. Similarly, bounds of this type are not applicable to the Randall-Sundrum scenario of warped extra dimensions with two branes since the first graviton KK state occurs at a scale $\sim$TeV. The astrophysical and cosmological considerations that restrict the scenario with large flat extra dimensions include graviton emission during the core collapse of supernovae (the emission of gravitons would accelerate the cooling), the heating of neutron stars from graviton decays, considerations of the cosmic diffuse γ-ray background, overclosure of the universe, matter dominated cooling of the universe, and reheating of the universe. The constraints from these considerations are summarised in Table 4.3 in terms of bounds on the fundamental scale M_D. It

	n			
	2	3	4	5
Supernova Cooling [123]	30	2.5		
Cosmic Diffuse γ-Rays:				
Cosmic SNe [124]	80	7		
$\nu\bar{\nu}$ Annihilation [125]	110	5		
Re-heating [126]	170	20	5	1.5
Neutron Star Halo [127]	450	30		
Overclosure of Universe [125]	$6.5/\sqrt{h}$			
Matter Dominated Early Universe [128]	85	7	1.5	
Neutron Star Heat Excess [127]	1700	60	4	1

Table 4.3: Summary of constraints on the fundamental scale M_D in TeV from astrophysical and cosmological considerations as discussed in the text [118].
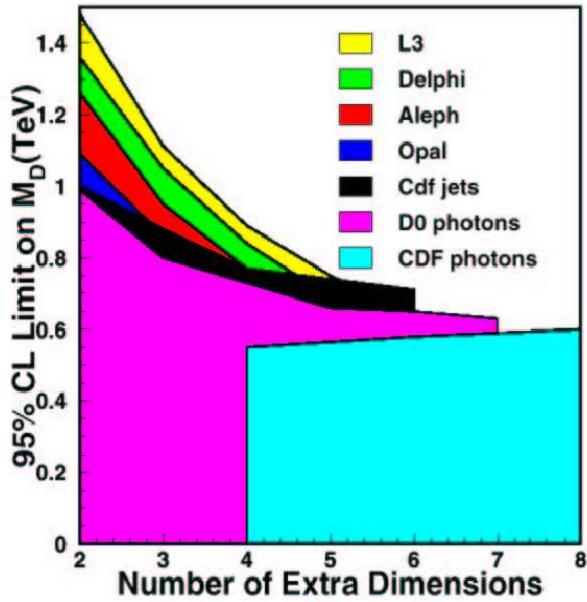
should be noted that the relation of the above constraints to M_D is tricky as numerical conventions, as well as assumptions regarding the compactification scheme, explicitly enter some of the computations; in particular, that of gravisstrahlung production during supernova collapse. In addition, all of these bounds assume that all of the additional dimensions are of the same size. The constraints in the table are thus merely indicative and should not be taken as exact. The calculation of the constraints from Table 4.3 can be found in the references given therein; An more detailed overall review of cosmological and astrophysical constraints be found in [118].

4.5.3   Constraints from collider experiments
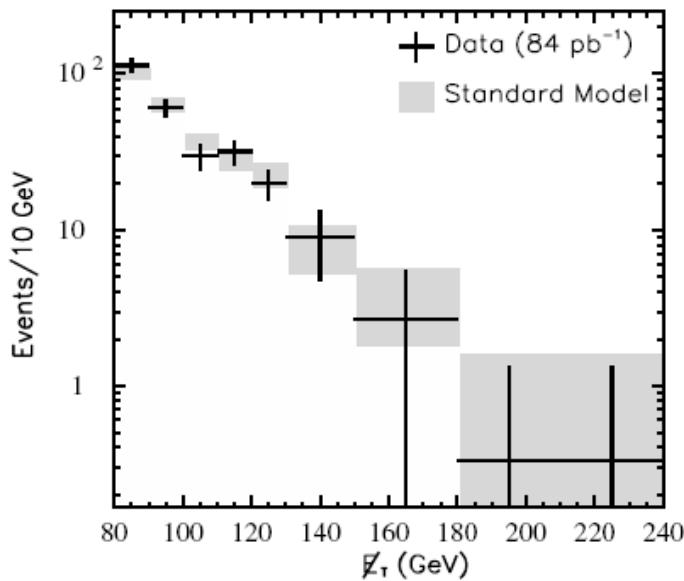
The search for extra dimensions in collider experiments can be divided into the direct search of graviton production and the search for deviations from the Standard Model predictions due to graviton virtual exchange.

Searches for direct KK graviton production were performed in the past years at the LEP e^+e^- collider at CERN in the reaction $e^+e^- \to G_n + \gamma(Z)$, using the characteristic final states of missing energy from the graviton plus single photon or Z boson. The four LEP experiments ALEPH [47], DELPHI [129], L3 [130] and OPAL[131] have excluded [132] fundamental scales up to ~ 1.60 TeV for two extra compactified dimensions and ~ 0.6 TeV for six extra dimensions, as shown in Table 4.4. These analyses use both total cross section measurements and fits to angular distributions to set a limit on the graviton production rates as a function of the number of extra dimensions.

At the Tevatron $p\bar{p}$ collider, the graviton emission is searched for by CDF [62] and D0 [133], in the channels $q\bar{q} \to \gamma G$, $q\bar{q} \to gG$, $qg \to qG$ and $gg \to gG$ corresponding

Figure 4.4: 95 % confidence limits on M_D as a function of the number of extra dimensions, for the different experiments at LEP and Tevatron [119].

Figure 4.5: Distribution of the missing transverse energy for events selected by the CDF collaboration at the Tevatron. The Standard Model predictions are also plotted as the bands [119].

to a missing transverse energy signature with a single jet or a single photon in the final state, as exemplified in Figure 4.5 . The lower limit on M_D, at 95% confidence level, from the different LEP and Tevatron experiments are summarised in Figure 4.4 for different numbers of extra dimensions.

Experiment	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
Aleph	1.26	0.95	0.77	0.65	0.57
Delphi	1.31	1.02	0.82	0.67	0.58
L3	1.50	1.14	0.91	0.76	0.65
OPAL	1.09	0.86	0.71	0.61	0.53
Combined	1.60	1.20	0.94	0.77	0.66

Table 4.4: Combined limits on M_D (in TeV) from LEP [108].

The second class of collider signals for large extra dimensions is that of virtual graviton exchange[109, 114] in $2 \rightarrow 2$ scattering. This leads to deviations in cross sections and asymmetries in Standard Model processes, such as $e^+e^- \rightarrow f\bar{f}$. It may also give rise to new production processes which are not present at tree-level in the Standard Model, such as $gg \rightarrow \ell^+\ell^-$. Using virtual Kaluza-Klein graviton exchange in reactions with di-photon, di-boson and dilepton final states, $e^+e^- \rightarrow G_n \rightarrow \gamma\gamma, VV, \ell\ell$, LEP experiments exclude $M_H \sim 0.5 - 1.0$ TeV independent of the number of extra dimensions. At the Tevatron, the combined Drell-Yan and diphoton channels exclude exchange scales up to ~ 1.1 TeV. In addition, H1 and ZEUS at HERA have both placed the bound $M_H \sim 800$ GeV [118].
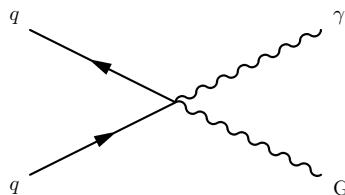
In summary, present facilities have searched for large extra dimensions and excluded their existence for fundamental scales up to $M_D \sim 1$ TeV. There are many compelling reasons to investigate further models with extra dimensions and the LHC is well suited to the task of exploring new energy domains. In the next chapter the analysis of the discovery potential of CMS for the ADD large extra dimensions scenario in the photon + $\not{E}_T$ channel is presented.

Chapter 5

Search for extra dimensions in the $\not{E}_T + \gamma$ final state

In the last years there has been an increasing interest in models that introduce extra dimensions in addition to the 3+1 dimensions from everyday's experience. As has been discussed in chapter 4, several models with extra dimensions predict detectable signatures at the LHC and are therefore attractive to experimentalists. The scenario of Arkani-Hamed, Dvali, and Dimopoulos (ADD) [7], see (subsection 4.4.5), was the first model with extra dimensions in which the compactified dimensions can be of macroscopic size, but stay consistent with all current experiments and observations; they are therefore referred to as "large extra dimensions" models.

Figure 5.1: Feynman graph of the ADD graviton production together with a photon.
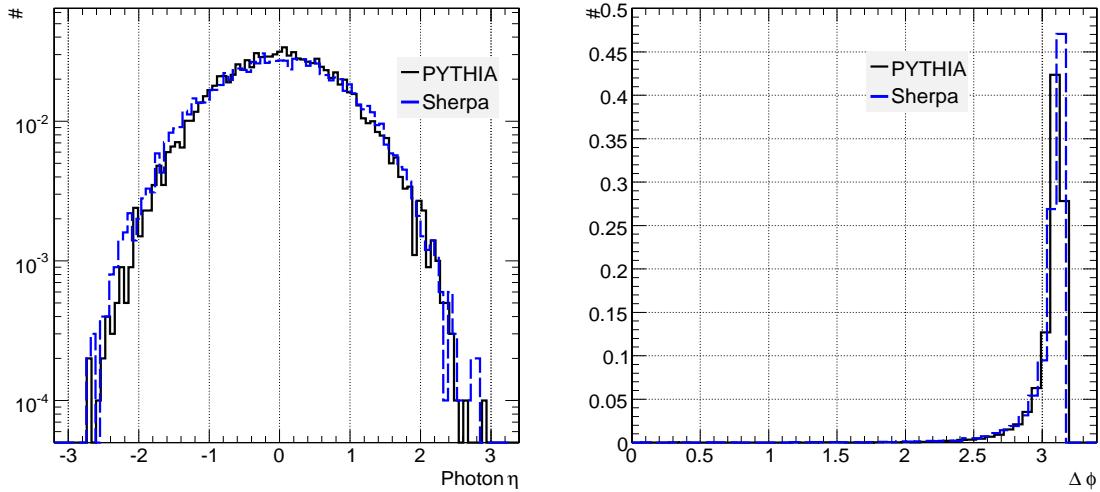
One relevant process for the LHC within the ADD framework is the graviton emission. At hadron colliders the graviton can be produced together with jets, giving rise to jets plus missing transverse energy ($\not{E}_T$) final states: $gg \rightarrow gG$, $qg \rightarrow qG$ and $q\bar{q} \rightarrow Gg$. The other significant contribution to the graviton production is the $q\bar{q} \rightarrow G\gamma$ process, which leads to an experimental signature of a photon plus $\not{E}_T$ and is studied in this chapter. The Feynman graph of the processes is show in Figure 5.1, the calculation of the cross sections can be found in [109].

5.1   Studies of the signal at generator level

The topology of the single photon + graviton event can be characterised by:

- a single high p_T photon in the central η region

- high missing p_T back to back to the photon in the azimuthal plane with a similar p_T distribution.
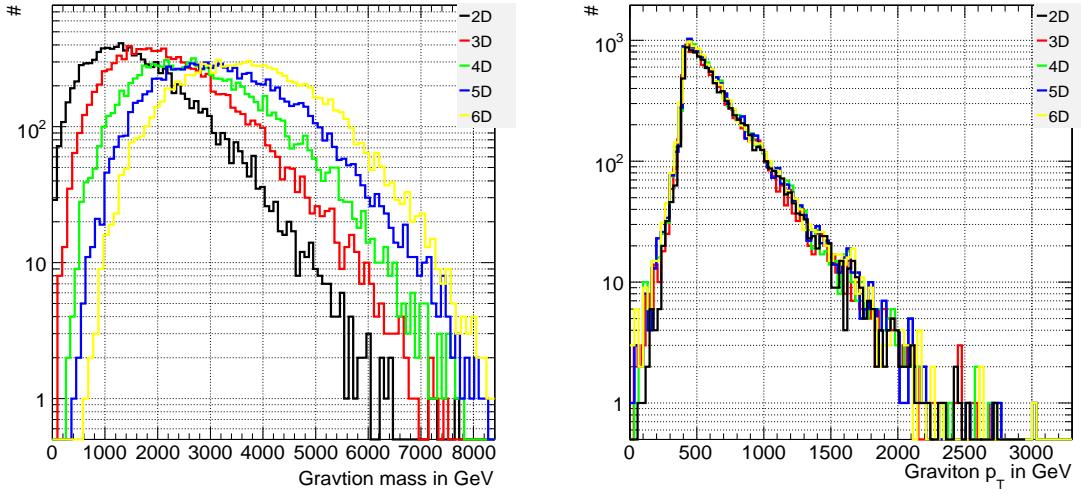
These characteristics listed above are almost independent of the parameters and shown in Figure 5.2 for an ADD scenario with two extra dimensions ($n = 2$) and a fundamental scale $M_D = 5$ TeV. In Figure 5.3 the mass of the graviton and its p_T spectrum for several number of extra dimensions are shown - the graviton gets "heavier "with increasing number of extra dimension, the p_T spectrum shows almost no dependence on this parameter. Therefore it is not possible to determine the model parameter n from the p_T spectrum, which is similar to the photon spectrum p_T^γ. Figure 5.4 also indicates that the η distribution of the photon does not offer the possibility to distinguish between the number of extra dimensions. Details of the comparisons at generator level are described in the following.
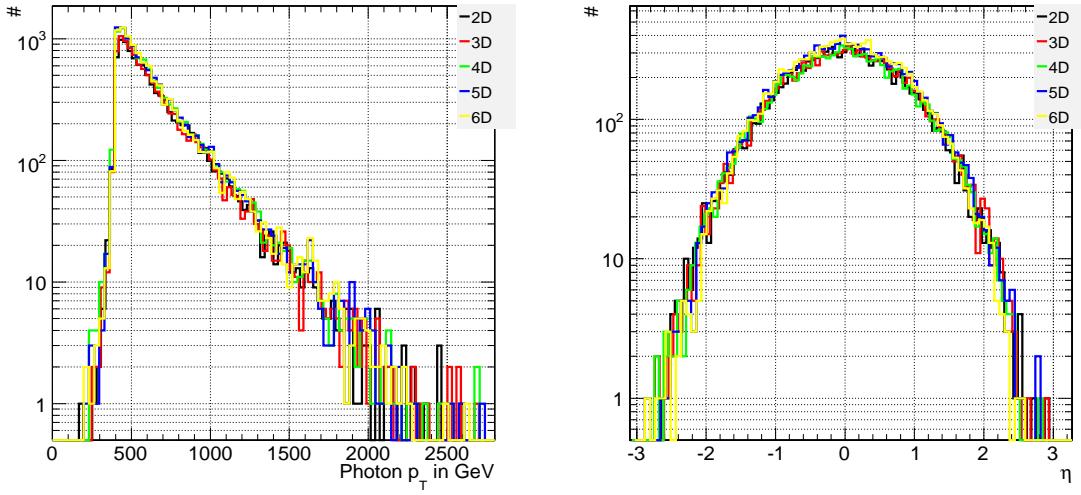
Figure 5.2: On the left: pseudorapidity η of the photon. On the right: angular difference $\Delta\phi$ in the azimuthal plane between the photon and the graviton (PYTHIA in black, SHERPA in blue (dotted), scenario with two extra dimensions ($n = 2$ and a fundamental scale $M_D = 5$ TeV).

5.1.1   Comparisons of SHERPA and PYTHIA

Two generators which provide the Arkani-Hamed, Dvali and Dimoupolos Large extra dimension model have been investigated: PYTHIA [18] and the object-oriented,
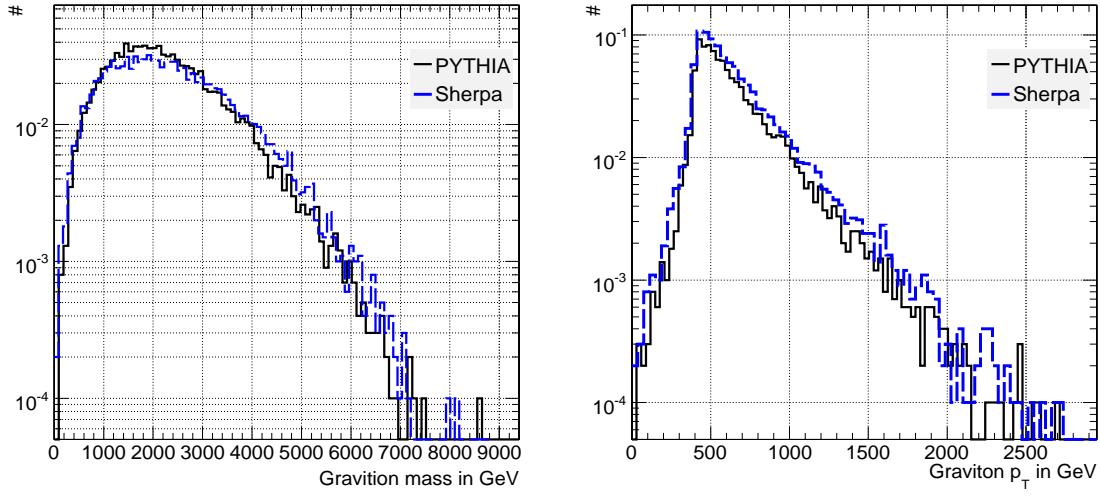
Figure 5.3: On the left: graviton mass for $M_D = 5$ and different number of extra dimensions. On the right: graviton p_T; samples generated with PYTHIA.

Figure 5.4: On the left: photon p_T for $M_D = 5$ and different number of extra dimensions. On the right: the η of the photon; samples generated with PYTHIA.

standalone event generator SHERPA [43]. The generator level studies were therefore performed using the PAX toolkit [46], which provides a standard CMS n-tuple and HepMC interface and allows for fast and efficient generator level comparisons (see subsection 2.2.3). The following versions of the generators and analysis tools have been used for the study:

- CMS generator package (CMKIN 4.4.0 [17]) containing PYTHIA 6.2.2.7

Figure 5.5: On the left: the mass of the graviton. On the right: distribution of the graviton transverse momentum p_T^G (PYTHIA in black, SHERPA in blue (dotted), scenario with two extra dimensions ($n = 2$) and a fundamental scale $M_D = 5$ TeV).

- SHERPA 1.06

- PAX toolkit version 2.00.10

- ROOT 5.08.00 [134]

The relevant distributions are generated and compared for several benchmark points with 1 TeV $\leq M_D \leq 5$ TeV and $2 \leq n \leq 6$. In both generators the CTEQ6L [135] parton distribution set was used. As will be explained later in the background section, rough estimates show that the event signature will not be detectable at the LHC in the low-p_T region, because the cross-section of the backgrounds, particularly of the irreducible $Z^0(\to \nu\bar{\nu}) + \gamma$ background, is too large. For all signal and background samples therefore a minimum p_T^γ of 400 GeV is consistently requested since the signal cross-section for the theoretically "safer" region ($M_D > 3.5$ TeV, see section 4.4.5) and the $Z^0 + \gamma$ cross-section are here of the same order of magnitude. (In an ATLAS study of this channel, a minimum $\not{E}_T$ of 500 GeV was chosen [116]). The following selection was applied:

- PYTHIA 6.2.2.7: CKIN(3) > 400 GeV (CKIN(3) is the minimum partonic centre of mass p_T, often named ($\hat{p}_T$) and corresponds roughly to p_T^γ.

- SHERPA 1.06 : $p_T^\gamma > 400$ GeV

In general one can say that the distributions from PYTHIA and SHERPA show good agreement for the benchmark points as is shown in Figure 5.2, Figure 5.5 and

M_D	1 TeV	2 TeV	3 TeV	4 TeV	5 TeV
$n = 2$	206.2 fb	12.0 fb	2.5 fb	0.8 fb	0.3 fb
$n = 3$	687 fb	21.0 fb	2.8 fb	0.6 fb	0.22 fb
$n = 4$	2.536 pb	39.0 fb	3.5 fb	0.61 fb	0.16 fb
$n = 5$	10.02 pb	78.0 fb	4.5 fb	0.611 fb	0.128 fb
$n = 6$	44.10 pb	161.0 fb	6.3 fb	0.631 fb	0.10 fb

Table 5.1: Total cross-sections for the signal for different model parameters calculated by SHERPA with a lower bound on the photon p_T of 400 GeV.

M_D	1 TeV	2 TeV	3 TeV	4 TeV	5 TeV
$n = 2$	221.8 fb	13.8 fb	2.73 fb	0.86 fb	0.35 fb
$n = 3$	753.9 fb	23.5 fb	3.10 fb	0.73 fb	0.24 fb
$n = 4$	2.69 pb	42.0 fb	3.69 fb	0.65 fb	0.17 fb
$n = 5$	10.07 pb	78.6 fb	4.6 fb	0.61 fb	0.12 fb
$n = 6$	39.18 pb	153.0 fb	5.97 fb	0.59 fb	0.10 fb
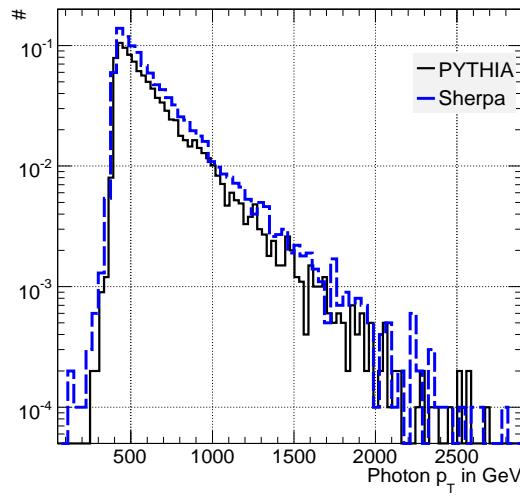
Table 5.2: Total cross-sections for the signal for different model parameters calculated by PYTHIA with a lower p_T cut of 400 GeV.

Figure 5.6. The cross section tends to be slightly smaller in SHERPA, the differences are on the level of some percent, as can be seen in Table 5.1, the Table 5.2 and in Figure 5.7.
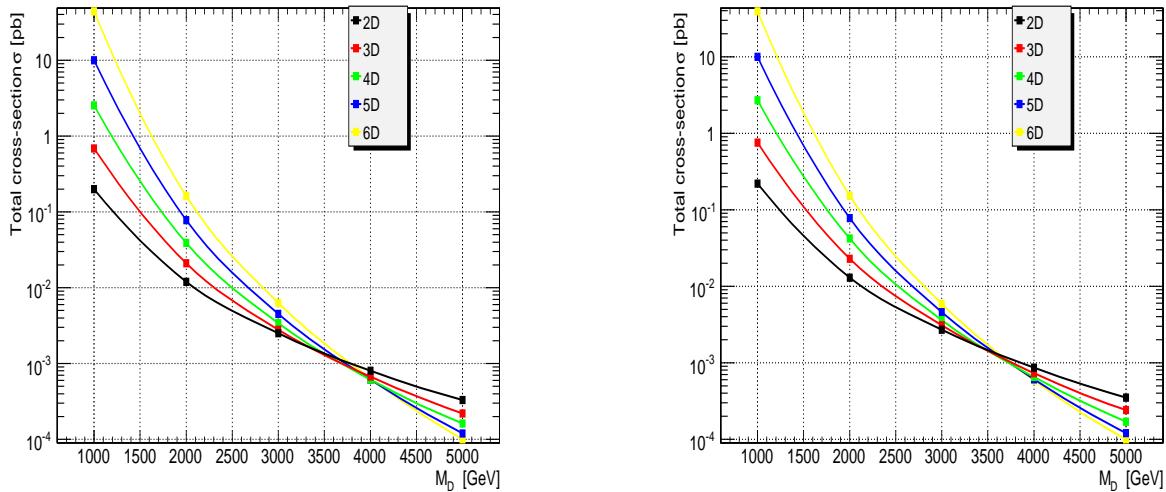
5.2 Background processes

The backgrounds considered in this analysis and their total cross-sections are listed in Table 5.3 and discussed below:

- The largest irreducible background is the di-boson production of $\gamma + Z^0 \rightarrow \nu_i \bar{\nu}_i$; the invisible decay of the Z^0 gives rise to a large $\not{E}_T$ rendering this process signal-like. This major background is studied in detail and discussed separately in section 5.5, where a normalisation method of this background from measured data is presented.

- The di-boson production $\gamma + W\pm \rightarrow e\nu$ is another background, when the electron is lost.

- A contribution is expected as well from the direct W production. The W boson decays in 10.72% into $W\pm \rightarrow e\nu$. The neutrinos show up in the detector as $\not{E}_T$, while the electron can be misidentified as a photon. This background,

Figure 5.6: Photon transverse momentum p_T^γ for an effective Planck scale $M_D = 5000$ GeV and two extra dimensions ($\delta = 2$). (PYTHIA in black, SHERPA in blue (dotted))

Figure 5.7: Total cross-section as function of the fundamental scale $M_D = 5$ for scenarios with different numbers of extra dimensions. (On the left with SHERPA, on the right with PYTHIA).

as well as all backgrounds containing highly-energetic charged particles (e, μ, jets, etc...) can be suppressed using a high-p_T track veto.

- $\gamma + W\pm \to \mu\nu$ where the muon is lost.

- $W\pm \to \tau(\to e\nu\bar{\nu})\nu$ is considered as well.

- QCD production can contribute to the background if a jet fakes a photon or is grossly miss-measured, so a dijet event can look like a $\gamma + \not{E}_T$ event.

- $\gamma+$ jets events will appear as $\gamma + \not{E}_T$ events, if the jet is not measured correctly or lost (i.e. along the beam pipe).

- $Z^0(\to \nu\bar{\nu})+$ jets is also a potential background, since it always has a natural amount of $\not{E}_T$; It can only be suppressed by a photon reconstruction with high purity and an efficient rejection of jets faking photons or non isolated photons in jets.

- Di-γ events (box and born diagram) where one γ is lost.

- Cosmic rays, where the muon undergoes a bremsstrahlung in the central electromagnetic calorimeter, have been the largest background contribution at CDF in a similar analysis (60%) [136]. The CDF detector is however situated closer to the surface. A muon may give rise to $\not{E}_T$ and/or create a bremsstrahlung photon. The same problem can occur with muons originating from the beam halo. However such events must coincide with an LHC event registered by the trigger. The study of this background class requires full detector simulation to correctly handle the time stamp information of the event, to which the cosmic or beam halo muon would contribute. The possible impact of these effects for this analysis at the CMS detector has not yet been investigated- so far only a rough estimate of the rate can be given.

Background	σ for $\hat{p}_T > 400\,\text{GeV}$
$Z^0\gamma \to \nu\bar{\nu} + \gamma$	2.16 fb
$W^\pm \to e\nu$	18.2 fb
$W^\pm \to \mu\nu$	18.2 fb
$W^\pm \to \tau\nu$	18.2 fb
$W^\pm\gamma \to e\nu+\gamma$	0.83 fb
$\gamma+$Jets	2.50 pb
QCD	2.15 nb
di-γ born	5.20 fb
di-γ box	0.14 fb
$Z^0 +$ jets	0.69 pb

Table 5.3: Total cross-sections for the Standard Model backgrounds considered.

Background	Rate for $p_T^\mu > 400\,\text{GeV}$
Cosmic muons	11 Hz
Beam Halo	1 Hz
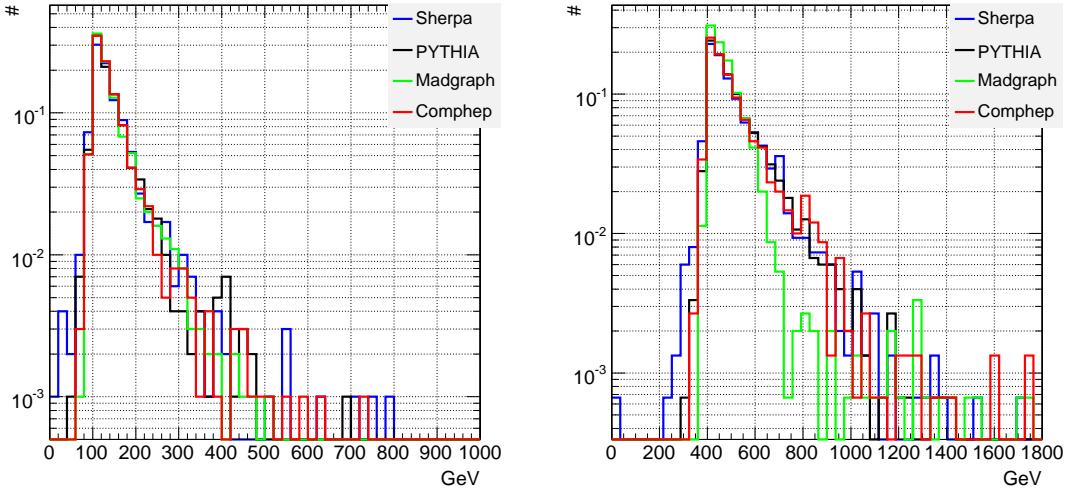
Table 5.4: Estimated rates for cosmic and beam halo muons (from CMKIN cosmic muon generator and first beam-halo studies [137, 138].)

5.2.1   Studies of the $Z^0 \to \nu\bar{\nu}$ background at generator level

The main, irreducible background $\gamma + Z^0(\to \nu_i\bar{\nu}_i)$ has been simulated and compared at generator level with four different event generators (PYTHIA, SHERPA, Comphep [20] and Madgraph [21]) in order to compare the relevant distributions and especially the tails of the p_T^γ and p_T^G spectrum. The settings for the process in the different generators has been chosen as identical as possible. For Madgraph and Comphep the CMKIN interfaces were used, the final comparison has been again performed with PAX. Good agreement between PYTHIA, SHERPA and Comphep has been found up to approximately 1 TeV, where the number of available events gets too low. However, the cross-section for the high-energetic tail is very small. An overview of the generators and cuts used for this comparison is shown in Table 5.5, the obtained distributions are shown in Figure 5.8. Only Madgraph shows a disagreement, which grows with an increasing production cut; this seems to be a bug and has been reported to the Madgraph team. The other event generators show a good agreement. For technical simplicity and consistency, PYTHIA is used in the following to generate this main background (as well as the other backgrounds). In addition, a relative normalisation method is proposed to measure this main background from data, which will be described later in section 5.5.

Event Generator	Cut [GeV]	Total cross-section σ [fb]
CompHEP 4.2p1	$p_T^{Z^0} > 100$	255
Madgraph	$p_T^{Z^0} > 100$	240
SHERPA 1.06	$p_T^{Z^0} > 100$	247
PYTHIA 6.227	CKIN(3) > 100	252
CompHEP 4.2p1	$p_T^{Z^0} > 400$	2.21
Madgraph	$p_T^{Z^0} > 400$	2.28
SHERPA 1.06	$p_T^{Z^0} > 400$	1.9
PYTHIA 6.227	CKIN(3) > 400	2.16

Table 5.5: Cross-section and settings for different event generators.

Figure 5.8: Comparison of the photon p_T^γ for the main background $\gamma + Z^0(\to \nu_i \bar{\nu}_i)$ with a lower production cut of 100 GeV on the left and 400 GeVon the right for different event generators: Comphep, Madgraph, PYTHIA and SHERPA.

5.3 Data samples and software

Due to the lack of official, fully simulated samples, CPU limitations and in order to increase the statistical precision, most samples were produced using the fast simulation [27]. However, small reference samples with the full simulation chain were produced as well in private production in order to compare the relevant physics objects to the fast simulation and to examine the performance of FAMOS for our process.

The following CMS software packages were used to perform the study:

- The generation of proton-proton collisions at 14 TeV centner of mass energy is done with CMKIN 4.4.0, based on PYTHIA with the CTEQ6L parton distribution set. The produced samples were used for generator studies, fast and full simulation.

- Most samples were simulated using the CMS fast simulation and reconstruction. All samples include pile-up with diffractive events. The fully simulated samples were produced with OSCAR [23], the Geant4-based CMS simulation package. Geant4 handles the particle propagation and simulates the interactions with the detector in detail. The simulation of the detector response as well as the reconstruction of the fully simulated events was performed with the CMS tool-kit ORCA [26].

- PAX 2.00.10 [46], a CLHEP 2.0 based toolkit for high energy physics is used

for the analysis itself.

- ROOT 5.08.00 is for histograms, statistics and fitting.

The data samples produced and used for the analysis are listed below
(all events are generated with a minimum partonic centre of mass p_T $CKIN(3) >$
400 cut):

- Signal samples: for each $n = 2-6$, $M_D = 1000-5000$ GeV 10,000 fast simulated events,

- 20,000 fully simulated signal events for comparison ($M_D = 5$ TeV, n = 2).

- 125,000 fast simulated $\gamma + Z^0 \to \nu_i \bar{\nu}_i$ events

- 60,000 fast simulated QCD events, additional study of this background in different p_T bins

- 50,000 fast simulated $\gamma +$ jets events

- 40,000 fast simulated $W^{\pm} \to e\nu/\mu\nu$ events

- 40,000 fast simulated $W^{\pm} \to \tau\nu$ events

- 40,000 fast simulated di-photon events (box and born diagram)

- 10,000 fast simulated $W^{\pm}+\gamma$ events

The following data samples have been produced in addition for the $\gamma + Z^0 \to \nu_i \bar{\nu}_i$
Candlecalibration:

- 20,000,000 generator events with $\gamma + Z^0 \to \nu\bar{\nu}$ at $CKIN(3) > 50$.

- 20,000,000 generator events with $\gamma + Z^0 \to \nu\bar{\nu}$ at $CKIN(3) > 300$.

- 135,000,000 generator events $\gamma + Z^0 \to \mu^+\mu^-$ at various energies.

- 1,000,000 fast simulated $\gamma + Z^0 \to \mu^+\mu^-$ at $CKIN(3) > 50$.

- 1,000,000 fast simulated $\gamma + Z^0 \to \mu^+\mu^-$ at $CKIN(3) > 300$.

- 30,000 fully simulated events $\gamma + Z^0 \to \mu^+\mu^-$ at $CKIN(3) > 50$.

- 30,000 fully simulated events $\gamma + Z^0 \to \mu^+\mu^-$ at $CKIN(3) > 300$.

- 1,000,000 fast simulated $\gamma + Z^0 \to e^+e^-$ at $CKIN(3) > 50$.

- 250,000 fast simulated $\gamma + Z^0 \to e^+e^-$ at $CKIN(3) > 300$.

- 30,000 fully simulated events $\gamma + Z^0 \to e^+e^-$ at $CKIN(3) > 50$.

- 30,000 fully simulated events $\gamma + Z^0 \to e^+e^-$ at $CKIN(3) > 300$.

5.4 Comparisons between CMS full and fast simulation

Most of the data samples for this study have been produced with the fast simulation FAMOS, since a large number of events was needed and full simulation, even with GFlash, would have taken too long [1] A detailed comparison with respect to the full simulation has been carried out to estimate the accuracy of the fast simulation in our case. The strategy has been as follows: first, the same generated samples were processed with both ORCA and FAMOS and the high-level objects obtained with the CMS analysis package ExRootAnalysis [55]. Then, using the PAX toolkit, the reconstructed quantities have been compared alongside. For this study the following objects of interest are investigated:

- **Photons**: the reconstructed photons have been obtained from the default of-fline photon candidates. A photon candidate is basically a supercluster in the electro-magnetic calorimeter (ECAL). A supercluster is a collection of calorimeter clusters, which consist of arrays of ECAL crystals. Photons are reconstructed using the Hybrid algorithm in the barrel and the Island algorithm in the endcap. The choice of the clustering algorithm depends on the spread of the deposited energy. For compact energy deposits, originating mainly from the single showers of unconverted photons, the best energy measurement is achieved using a 5×5 crystal array, while superclustering algorithms provide better measurement for multiple showers originating from conversions where the bending of the electron and positron tracks, and their radiation in tracker material, spreads the energy over a larger area of the calorimeter. A very convenient measure of the lateral spread of energy deposition is provided by the $R9$ variable, defined as the ratio of energy contained in a 3×3 array of crystals (centred on the crystal with the highest deposited energy) to the total supercluster energy. Values approaching unity are obtained for unconverted photons, or for photon that have converted very close to the ECAL. Smaller values are obtained for increasing distances of the conversion vertex from the ECAL. Fake photon signals due to jets can usually be rejected by looking for additional energetic particles in a cone around the reconstructed ECAL cluster (photon isolation).

- **Electrons**: the electron candidates are reconstructed and identified with the default configuration of the offline electron reconstruction algorithm. The candidate is essentially an ECAL super-cluster (as in the photon case) with a matched track. Typically, the supercluster has an angular extension in ϕ because of the emission of bremsstrahlung along the curved trajectory. The

---

[1]Furthermore, GFlash has not been fully validated and released for large scale production at the time when the analysis was performed.
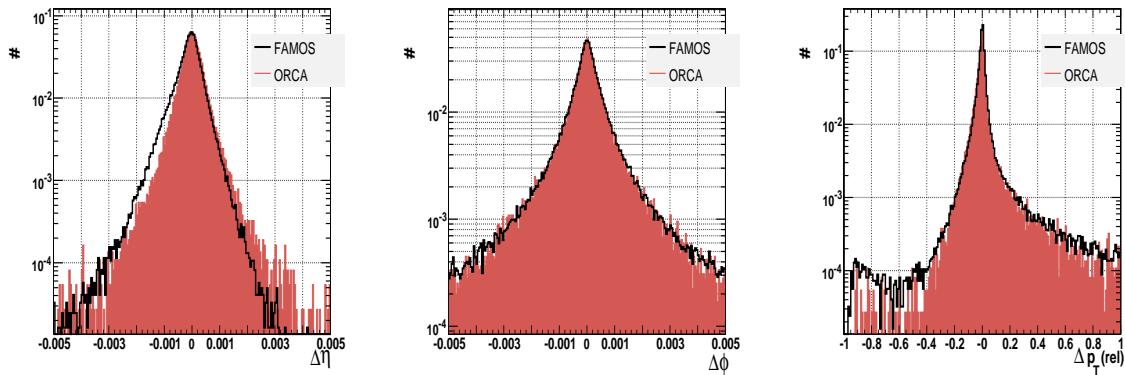
amount of radiated bremsstrahlung depends on the traversed material budget and can be very large. About 50% of the electrons radiate 50% of their energy before reaching the ECAL surface and in 10% of the electrons, more than 95% is radiated. Therefore, advanced superclustering algorithms are employed which search along the direction for energy deposits, followed by algorithmic energy corrections. The pixel hits found serve as seeds for the subsequent track reconstruction with the full tracker. The default track reconstruction method is with a Kalman filter or with a more complex nonlinear filter approach using a Gaussian Sum Filter (GSF). In addition, a likelihood for each electron candidate is calculated in this analysis based on information from the electromagnetic calorimeter and the tracker. A standard package for electron identification which computes the electron likelihood and is available in the ORCA toolkit has been used.

- **Muons**: A muon candidate is formed when a muon track is found in the standalone muon system (RPC, CSC, DT) and can be matched to a track in the central silicon tracker. Based on the information from the muon system, the position of hits in the pixel detector is predicted by backwards propagation through the magnetic field, similar to the procedure in the case of electrons.

- **Jets**: Jet reconstruction is performed starting from the electromagnetic and hadron calorimeter (HCAL) cells. In case of the HCAL, the cells are arranged in tower patterns which can be extended to also include the ECAL crystals. These "ECAL plus HCAL towers" towers serve as input to all jet and $\not{E}_T$ reconstruction algorithms. In CMS, three basic jet reconstruction algorithms are used: the Iterative Cone algorithm, the Midpoint Cone algorithm and the Inclusive k_T algorithm. In this analysis, the default uncalibrated jet candidates reconstructed with the Iterative Cone algorithm were used. Basically, a jet is formed here by casting a cone with the radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ around the highest E_T object. The objects inside the cone are used to form a proto jet. The obtained direction of this jet is used to seed a new proto jet, which is repeated until the energy does not change by more than 1% and the direction does not change by more than $R < 0.01$. The stable jet is added to the list of jets and the objects inside the cone are removed from the list of objects for the next iteration. The iteration proceeds until no objects above a seed threshold, which is a parameter of the algorithm, are available. In this study a cone of $R = 0.5$ is used.

- **Missing transverse energy**: $\not{E}_T$ reconstruction is taken as estimate of the missing p_T spectrum from the final state neutrinos. The missing energy is calculated from jets (see above) using the Iterative Cone Algorithm with activated muon and electron correction.
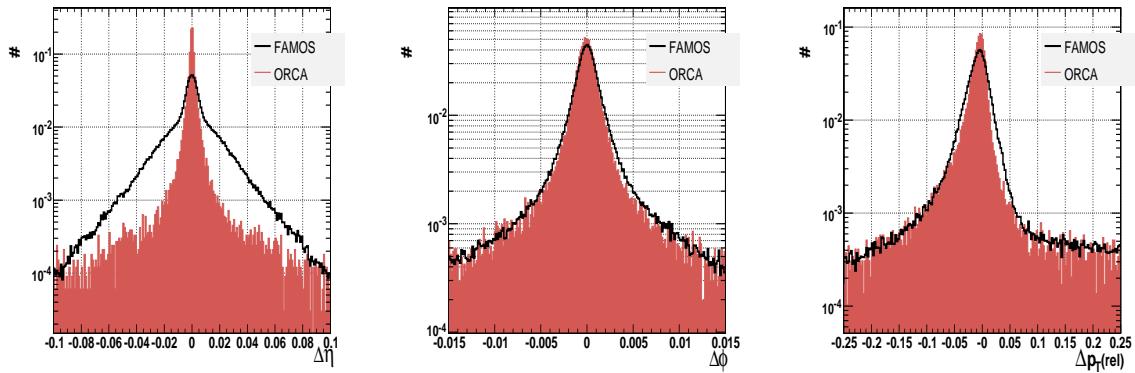
More details about the reconstruction techniques of the high level objects used in this analysis can be found in [6].

5.4.1   Resolution and efficiency studies

The reconstructed objects are matched to the corresponding generator particles with the objective to compare the resolutions, efficiencies and purities in case of fast and full simulation. The events used for this study are the same that are used later for the normalisation of the main background $\gamma + Z(\rightarrow \nu\bar{\nu})$. [2] The electrons used in the comparison are those for which the ElectronLikelihood yields a likelihood of at least 0.65.
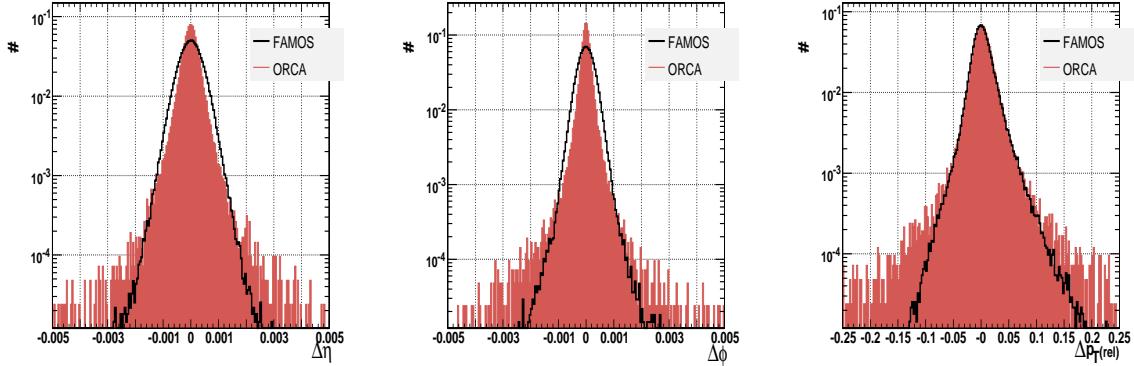
Figure 5.9: Resolution for electrons in η, ϕ and the relative p_T resolution: the agreement between ORCA and FAMOS is very good and in accordance with the design values [139].
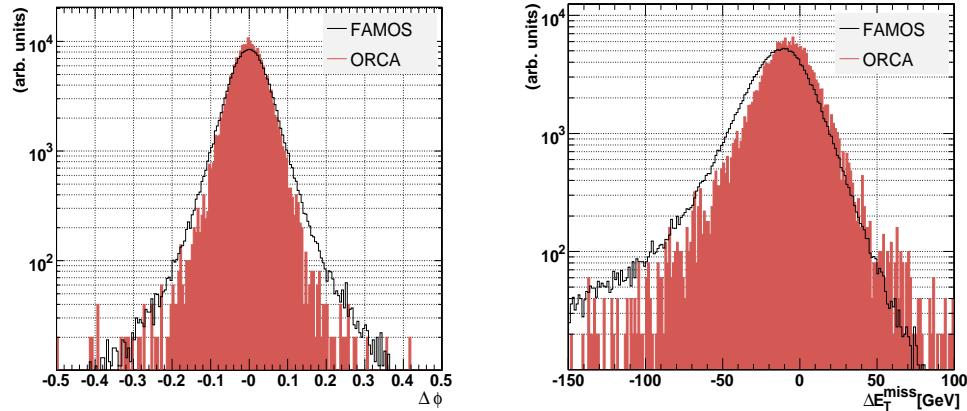
Figure 5.10: Resolution for photons in η, ϕ and the relative p_T resolution: the agreement between ORCA and FAMOS is very good for p_T and ϕ. The η value is currently not correctly determined in FAMOS and the resolution worse than in ORCA - this will be fixed in the near future.

---

[2]A comparison of signal events gives similar results.

Figure 5.11: Resolution for muons in η, ϕ and the relative p_T resolution: there is a slight difference between ORCA and FAMOS in η and ϕ, the agreement in p_T is very good.

Figure 5.12: Resolution for $\not{E}_T$ in ORCA and FAMOS in ϕ (left) and E_T (right).

The absolute resolution and the relative resolution are defined as:
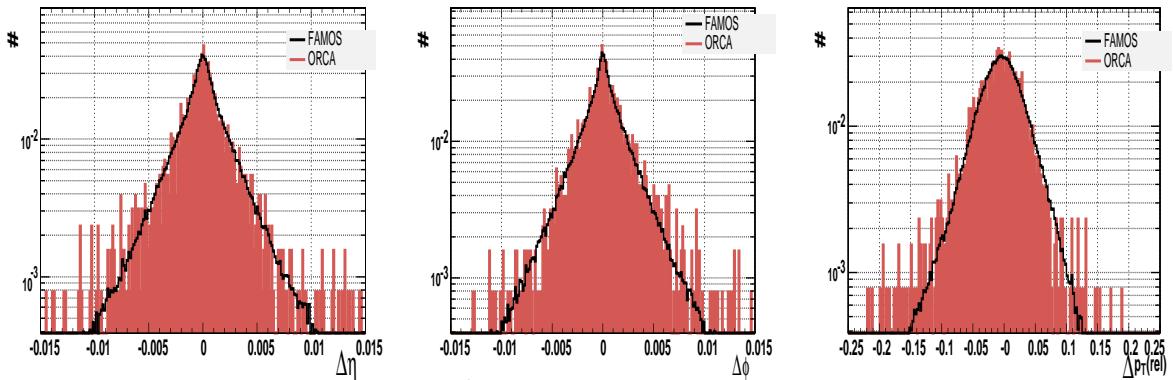
$$\Delta X = X_{rec} - X_{gen} \tag{5.1}$$

$$\Delta X_{(rel)} = \frac{X_{rec} - X_{gen}}{X_{gen}} \tag{5.2}$$
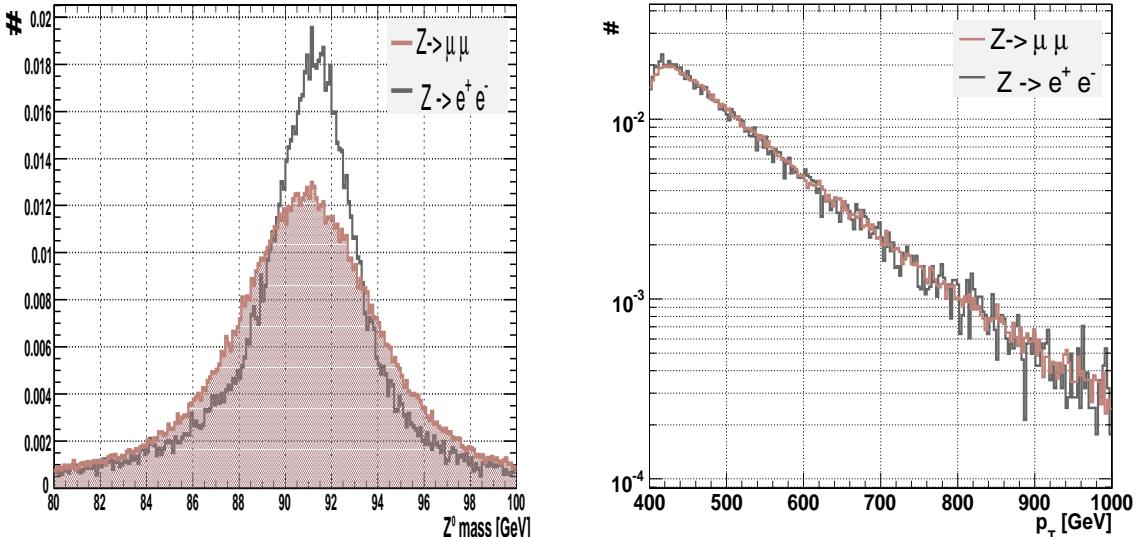
A combination $\Delta_{combined}$ of the individual resolutions $\Delta\eta$, $\Delta\phi$, $\Delta p_T(rel)$, and their respective standard deviations is used as matching criterion:

$$\Delta_{combined} = \sqrt{(\frac{\Delta\eta}{\sigma_{\Delta\eta}})^2 + (\frac{\Delta\phi}{\sigma_{\Delta\phi}})^2 + (\frac{\Delta p_{T(rel)}}{\sigma_{\Delta p_{T(rel)}}})^2} \tag{5.3}$$

A pair of a generated and a reconstructed particle is considered as matched when $\Delta_{combined} < 4$. The obtained resolutions for electrons, photons, muons and the Z^0 mass can be seen in Figure 5.9 – Figure 5.13. In general, the resolutions are consistent with the expected design values and a good agreement between ORCA and FAMOS is found. The only problem found has been the η resolution of the photon:

Figure 5.13: Resolution for the Z^0 boson, reconstructed from muons in ORCA and FAMOS; the agreement in η and ϕ and in p_T is very good.

Figure 5.14: The Z^0 mass (left) and the $Z^0_{p_T}$ (right) reconstructed from $\mu^+\mu^-$ (red) and e^+e^- (grey).

its value is currently not correctly determined in FAMOS and the resolution worse than in ORCA. However, the impact of this error to this analysis is small.

When a very high energetic photon hits the centre of a crystal, it is possible to have saturation (at about $E_{crystal} > 1.7$ TeV). First studies shows that in this case in our sample the energy can be reconstructed up to 5 % too low. A method to correct the energy using the energy deposition in the surrounding crystals has been recently presented and can be applied for this case [140]. However, in all samples used for this study the probability to have a photon in this p_T region is smaller than 1%. The effect from this potential inaccuracy can thus be safely ignored. Another interesting fact noticed during this study was that the Z^0 mass resolution from electrons is better than the resolution obtained using muons for the samples with $CKIN(3) > 400$ GeV, as can be seen in Figure 5.14. This is due to the fact, that

Reconstructed object		Resolution ORCA	Resolution FAMOS
photon	$\Delta\eta$	$1.5 \cdot 10^{-03}$	$1.7 \cdot 10^{-02}$
	$\Delta\phi$	$1.2 \cdot 10^{-03}$	$1.5 \cdot 10^{-03}$
	$\Delta p_T(rel)$	1.1 %	1.9 %
muon	$\Delta\eta$	$2.7 \cdot 10^{-04}$	$4.1 \cdot 10^{-04}$
	$\Delta\phi$	$1.5 \cdot 10^{-04}$	$4.4 \cdot 10^{-04}$
	$\Delta p_T(rel)$	1.5 %	1.5 %
electron	$\Delta\eta$	$3.5 \cdot 10^{-04}$	$3.8 \cdot 10^{-04}$
	$\Delta\phi$	$5.5 \cdot 10^{-04}$	$5.8 \cdot 10^{-04}$
	$\Delta p_T(rel)$	2.0 %	1.8 %
	ΔE_T^{miss}	17.1 GeV	19.7 GeV
	$\Delta\phi(E_T^{miss})$	$4.2 \cdot 10^{-02}$	$4.9 \cdot 10^{-02}$
Z^0	$\Delta\eta$	$2.1 \cdot 10^{-03}$	$2.3 \cdot 10^{-03}$
	$\Delta\phi$	$2.1 \cdot 10^{-03}$	$2.4 \cdot 10^{-03}$
	$\Delta p_T(rel)$	3.4 %	3.7 %

Table 5.6: Overview over resolution ΔX for the relevant objects in this analysis.

in this p_T region the measurement in the calorimeter becomes better than the muon p_T measurement from the tracker, as can be seen in Figure 5.15.

After having defined a common criterion whether a final state particle has been correctly reconstructed or not, efficiency and purity are compared as function of η and p_T.
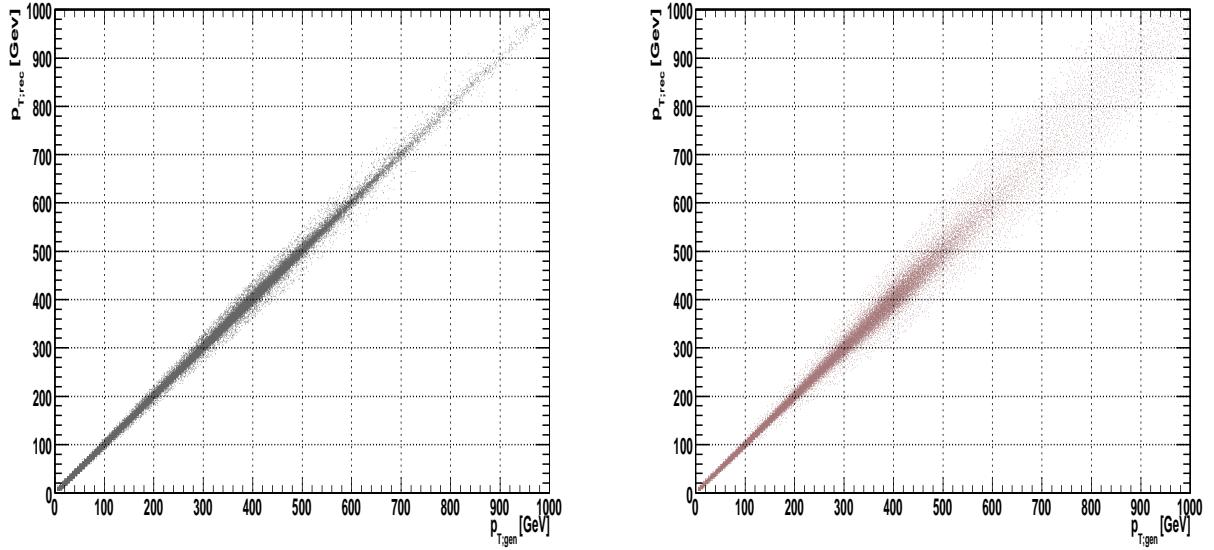
In this context, efficiency and purity are defined as:

$$\text{eff} = \frac{\# \text{ matched particles}}{\#\text{generated particles}} \tag{5.4}$$
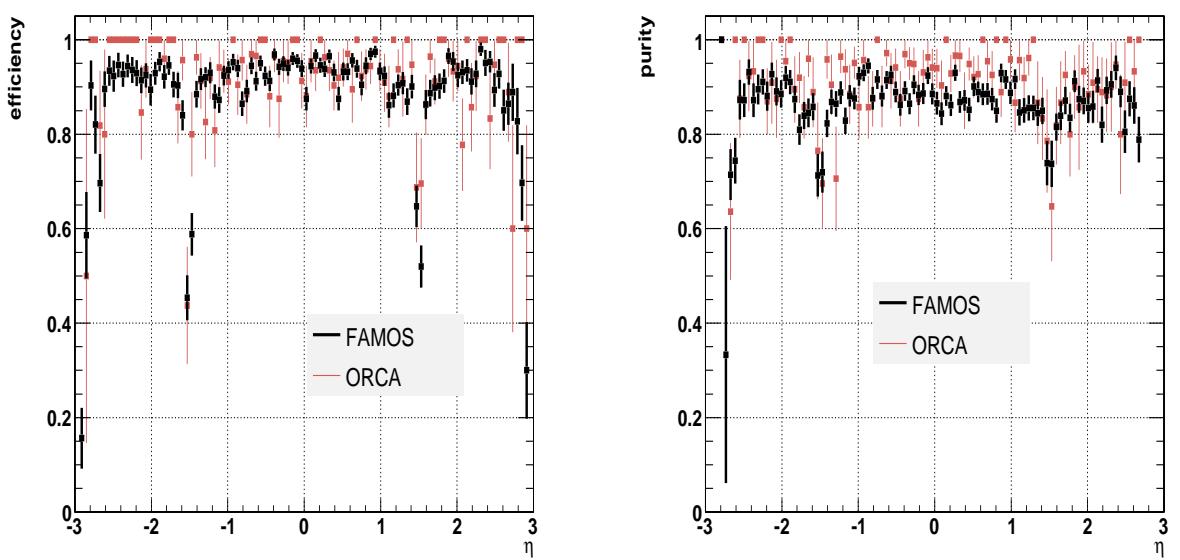
$$\text{pur} = \frac{\# \text{ matched partcles}}{\#\text{reconstructed particles}} \tag{5.5}$$

Again, a good agreement was found for all considered objects, as is shown in Figure 5.16 for photons.

This comparison has shown that the performance of FAMOS for efficiency, purity and resolution is quite good and compares well with ORCA. Therefore the use of FAMOS in order to increase the statistical precision and save computing time is justified.

Figure 5.15: p_T of reconstructed electrons (left) and muons (right) versus the corresponding generated particles.

Figure 5.16: Efficiency and purity for photons as function of η .

5.5   The $Z^0 + \gamma$ "Candle" calibration

In this section a method is described on how the full $\gamma + Z^0 \rightarrow \nu_i \bar{\nu}_i$ spectrum can be measured from $\gamma + Z^0 \rightarrow \mu^+\mu^-/e^+e^-$ events. First a conservative set of selection cuts is chosen to be able to reconstruct the "candle" from the final state particles. Then the total acceptance for events which passed the candle selection is studied as well as the estimate of the reconstruction efficiencies. This study has been mostly performed in the context of a Master's thesis, which was affiliated to this analysis [54].

5.5.1 $\gamma + Z^0 \rightarrow \mu^+\mu^-/e^+e^-$ selection

In order to reconstruct the lepton pair reliably with good precision, some kinematic and topological constraints are imposed. For the selection of $\gamma + Z^0 \longrightarrow \mu^+\mu^-$ events the following selection criteria on the reconstructed final state particles are applied:

- The single hard photon has to be found in a pseudo-rapidity range of $|\eta_\gamma| < 2.7$ in the ECAL. In the high-p_T range of interest ($p_T^\gamma > 400$ GeV) practically all photons in $\gamma + Z^0$ signals will be in that range.

- The selection criteria of the muons are chosen as follows:

  - Both muons from the Z^0 decay are required to have a minimum transverse momentum $p_T^{\mu^\pm} > 20$ GeV to be reliably found by the muon trigger (the single muon trigger uses a nominal cut of 14 GeV for the L1 trigger and 19 GeV for the HLT).

  - In order to avoid effects on the edge of the muon system, both muons are required to be within $\eta_{\mu^\pm} < 2.3$. The muon reconstruction efficiency would quickly drop at the edges of the muon system coverage and impose unwanted uncertainties on the reconstruction efficiency otherwise.

Similarly the following criteria are applied for the selection of $\gamma + Z^0 \longrightarrow e^+e^-$ events:

- The electrons are identified using a likelihood approach (standard electron likelihood module included in the ORCA reconstruction package) with a discriminator cut at 0.65.

- The electrons are required to have a minimum transverse momentum $p_T^{e^\pm} > 20$ GeV like the muons.

- For the electron identification it is important to find the electron track, so the η limit is imposed by the tracking system and electrons are only accepted with $\eta_{e^\pm} < 2.4$.
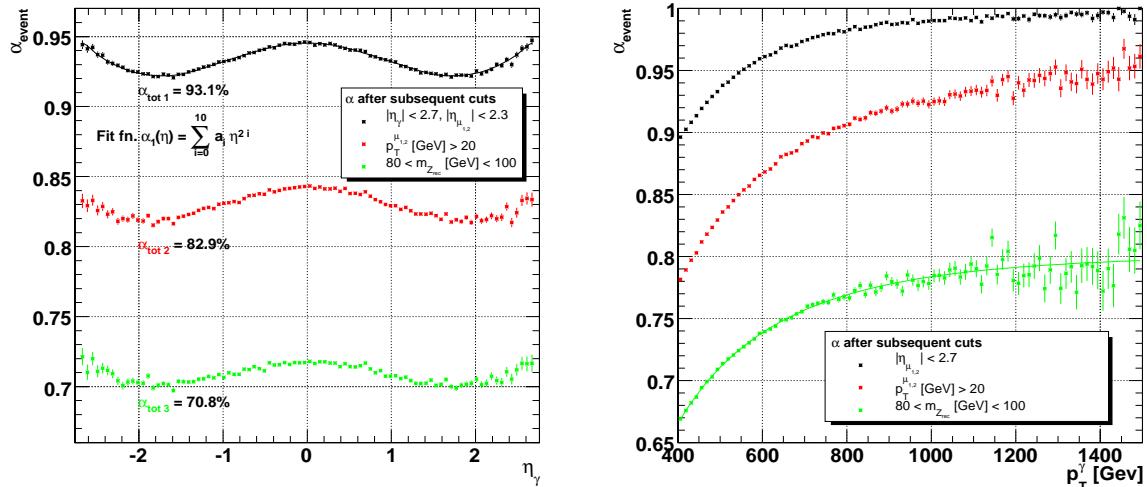
For both kind of events the common selection criteria on the photon and the reconstructed Z^0 are:

- The reconstructed Z^0 is required to be found within the mass window of 80 GeV $<$ m$_{Z^0}$ $<$ 100 GeV, as shown in Figure 5.14.

- The γ and Z^0 are required to be back-to-back in the $x - y$ plane, $\Delta\phi(\gamma, Z^0) >$ 2.5

- Both particles form the decay of the Z are required to be within 50% of their average p_T as follows: $|\frac{p_T^\gamma - p_T^{Z^0}}{p_T^\gamma + p_T^{Z^0}}| < 0.25$

5.5.2 $\gamma + Z^0 \rightarrow \mu^+\mu^- / e^+e^-$ acceptance

In the following, the reconstruction efficiency and the detector acceptance are studied separately. The reconstruction efficiency can be approximated via simple p_T dependent functions. The detector acceptance is highly dependent on the topology of the event.

To reliably normalise the $\gamma + Z^0 \rightarrow \nu_i\bar{\nu}_i$ predictions using the $\gamma + Z^0 \rightarrow \mu^+\mu^-$ data, the detector acceptance is parameterised as a function of the p_T and η of the photon. The acceptance α for high-p_T events ($p_T^\gamma > 400$ GeV) after the subsequent selection cuts is shown in Figure 5.17.

Figure 5.17: Acceptance α for high-p_T events ($p_T^\gamma > 400$ GeV) after using all candle selection criteria.

The total acceptance as a function of η is not constant for different p_T^γ regions. It is rather different in the low-p_T range ($p_T^\gamma \approx 100$ GeV) where the pseudo-rapidity distribution of the Z^0 is similar to the distribution of a single Z^0 production.

The detector acceptance is parameterised using a two-dimensional function $\alpha(p_T^\gamma, \eta_\gamma)$. The inverse of this function is used as a weighting function for accepted

events to transform the measured photon p_T distribution to the full p_T^γ spectrum. With this method, the p_T^γ spectrum of $\gamma + Z^0 \to \nu_i \bar{\nu}_i$ can be normalised to the one weighted for acceptance and efficiency from the candle sample.

The acceptance function $\alpha(p_T^\gamma, \eta_\gamma)$ is obtained by fitting even Tchebycheff polynomials of sixth order (four parameters) in different p_T^γ slices in the range between $100 \text{ GeV} < p_T^\gamma < 1200 \text{ GeV}$ and then describing the Tchebycheff coefficients in turn by fifth-order polynomials. The overall fit χ^2/ndf is close to 1.

The acceptance for the electron based calibration is done in an identical way and only differs by the slightly larger electron η cut. The average acceptance numbers resulting from the study are shown in Table 5.7.

$\gamma + Z^0 \to \mu^+\mu^-$		$\gamma + Z^0 \to e^+e^-$	
cut	acceptance	cut	acceptance
$|\eta_\gamma| < 2.7, |\eta_{\mu^\pm}| < 2.3$	93.1%	$|\eta_\gamma| < 2.7, |\eta_{e^\pm}| < 2.4$	94.6%
$p_T^{\mu^\pm} > 20 \text{ GeV}$	82.9%	$p_T^{e^\pm} > 20 \text{ GeV}$	84.1%
$80 < m_{Z^0} < 100 \text{ GeV}$	70.8%	$80 < m_{Z^0} < 100 \text{ GeV}$	71.8%

Table 5.7: Remaining $\gamma + Z^0 \to \mu^+\mu^-/e^+e^-$ after each cut for $p_T^\gamma > 400 \text{ GeV}$
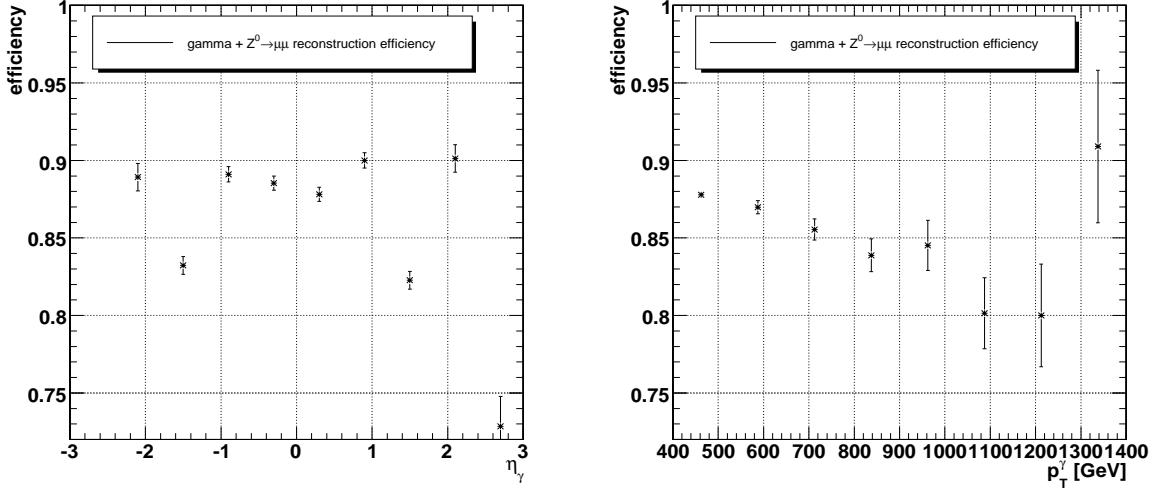
5.5.3 $\gamma + Z^0 \to \mu^+\mu^-/e^+e^-$ reconstruction efficiency

The transformation method based on the generator study using $\alpha(p_T^\gamma, \eta_\gamma)$ is now tested against the detector simulation in order to parameterise the reconstruction efficiency effects. The simulation has been mostly done with the fast simulation FAMOS at high statistics and compared with small fully simulated samples.

After the transformation, the number of events in the different p_T^γ and η_γ bins is compared to the number of expected events in these bins assuming an ideal detector with full 4π coverage (i.e. the generator information). The reconstruction efficiency thus obtained with FAMOS is shown in Figure 5.18. Again the results for the electrons are very similar and not shown explicitly.

The reconstruction efficiency is composed of several factors. One photon and two muons have to be reconstructed. Furthermore, the reconstructed Z^0 has to pass the mass window constraint. The reconstruction of its invariant mass requires an accurate measurement of the muon kinematics. The main limiting factor here is the momentum measurement, especially for muons with high p_T values since their tracks become rather straight and a precise momentum measurement is challenging. This leads to a smearing of the Z^0 mass peak and deteriorate the efficiency in the high-p_T range.

The reconstruction efficiencies are mostly flat as function of η_γ except for the ECAL gap between barrel and endcap (at about $\eta = 1.5$). As long as there is no

Figure 5.18: Reconstruction efficiency ϵ_{rec} for high-p_T events ($p_T^\gamma > 400$ GeV) passing all selection criteria as function of p_T^γ and η^γ .

interest in precise measurement of the η distributions the efficiency can be assumed to be constant in η_γ for a given p_T^γ and is slightly falling for larger p_T^γ values. A very simple approximation is done here via a linear fit through the data points of the FAMOS simulated efficiency:

$$\epsilon_{rec}(p_T^\gamma) = \epsilon_{rec}^0 + \epsilon_{rec}^1 \cdot p_T^\gamma \tag{5.6}$$

The total reconstruction efficiency ϵ_{tot} can be expressed as

$$\epsilon_{tot} = \alpha(p_T^\gamma, \eta_\gamma) \cdot \epsilon_{rec}(p_T^\gamma) \tag{5.7}$$

In Table 5.8, the detector acceptance and the reconstruction efficiencies using the fast (FAMOS) and the full detector simulation (ORCA) are listed.
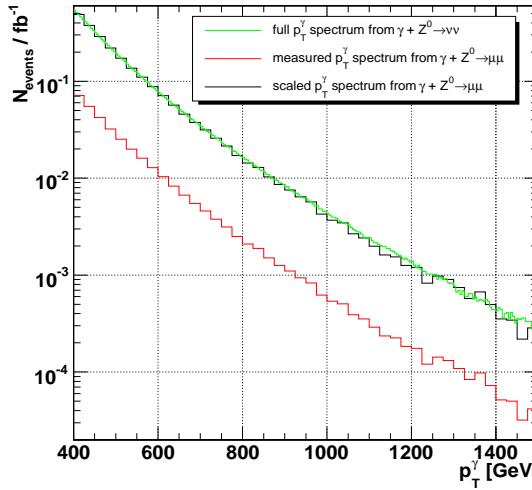
Cut GeV	Sample	Detector acceptance	Reconstruction eff. Fast	Full	Total eff. Fast	Full
$p_T^\gamma > 100$	$\gamma + Z^0 \rightarrow \mu^+\mu^-$	39.2%	94%	93%	37%	36%
	$\gamma + Z^0 \rightarrow e^+e^-$	45.3%	90%	89%	41%	40%
$p_T^\gamma > 400$	$\gamma + Z^0 \rightarrow \mu^+\mu^-$	70.8%	87%	83%	62%	59%
	$\gamma + Z^0 \rightarrow e^+e^-$	71.8%	82%	83%	59%	60%

Table 5.8: Detector acceptance, reconstruction efficiencies and the total efficiency using the fast (FAMOS) and the full detector simulation (ORCA).

As can be seen in Table 5.8, FAMOS and ORCA slightly differ in the reconstruction efficiency. The uncertainty is of the same order of magnitude as the statistical uncertainty from the total number of observable events ($< 3\%$ after $30 fb^{-1}$ of $\gamma + Z^0 \rightarrow \mu^+\mu^-/e^+e^-$).

5.5.4 Kinematics and E_T^{miss} in $\gamma + Z^0 \to \mu^+\mu^-/e^+e^-$ and $\gamma + Z^0 \to \nu\bar{\nu}$
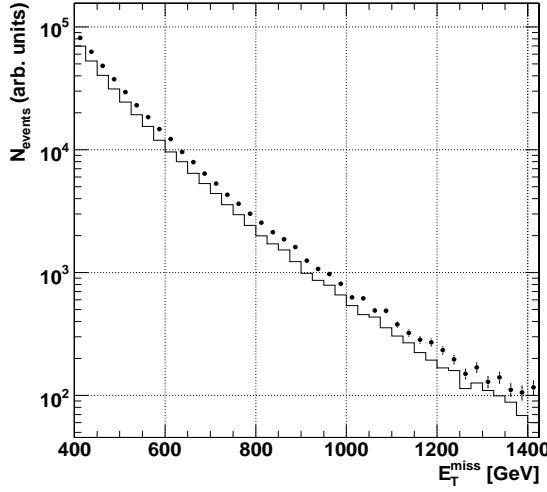
To prove that the normalisation method using the measured $\gamma + Z^0 \to \mu^+\mu^-$ events corrected for acceptance × efficiency (Equation 5.7) can be used to calibrate the $\gamma + Z^0 \to \nu_i\bar{\nu}_i^-$ events, the p_T distributions for the γ and the Z^0 (reconstructed from the muons for $Z^0 \to \mu^+\mu^-$ and E_T^{miss} in the $Z^0 \to \nu_i\bar{\nu}_i$ case) are compared.

Figure 5.19:   Number   of   expected p_T^γ events   per   25 GeV   bin   at   1fb^{-1}
from   measured $\gamma + Z^0 \to \mu^+\mu^-$ events   before   and   after   trans-
formation   compared   with   the   generator   distribution   for
$\gamma + Z^0 \to \nu_i\bar{\nu}_i$. The transformed muon distribution models well the $\nu_i\bar{\nu}_i$ spectrum.

Figure 5.19 shows the measured and the derived (i.e. corrected for acceptance × efficiency and scaled with the branching ratio) p_T^γ spectrum from $\gamma + Z^0 \to \mu^+\mu^-$ in comparison with the true generator spectrum for $\gamma + Z^0 \to \nu_i\bar{\nu}_i$ events. Since the p_T spectrum of the Z^0 at generator level corresponds to the photon p_T^γ spectrum, the weighted $\gamma + Z^0 \to \mu^+\mu^-$ spectrum delivers a precise approximation of both true spectra.

The particle balancing the transverse momentum of the photon in $\gamma + Z^0$ events is the Z^0. While the Z^0 can be reconstructed from the leptons ($\mu^+\mu^-$ and e^+e^- respectively) it shows up as missing transverse energy (E_T^{miss}) when the Z^0 decays into neutrinos. The $E_T^{Z^0}$ spectrum from the derived $\gamma + Z^0 \to \mu^+\mu^-$ events compared with the reconstructed E_T^{miss} in the $\gamma + Z^0 \to \nu_i\bar{\nu}_i$ case can be seen in Figure 5.20. The distributions are not expected to be identical: the reasons for the difference is that the E_T^{miss} a complicated object and its reconstruction in CMS (and in general) is not very accurate compared to the precise reconstruction of the Z^0 from muons or electrons. Therefore the derived spectrum gives a better description of the true $Z^0 \to \nu_i\bar{\nu}_i$ distribution.

Figure 5.20: Reconstructed E_T^{miss} from $\gamma+Z^0 \to \nu_i\bar{\nu}_i$ (dots) in comparison with transformed $\gamma+Z^0 \to \mu^+\mu^-$ events (line) ($E_T^{Z^0}$ used as E_T^{miss}).

The average multiplicative factors going into the derivation are shown in Table 5.9.

Cut	$\gamma+Z^0 \to \mu^+\mu^-$			$\gamma+Z^0 \to e^+e^-$		
	$1/\epsilon_{tot}$	$\frac{br_{Z^0\to\nu_i\bar{\nu}_i}}{br_{Z^0\to\mu^+\mu^-}}$	**total**	$1/\epsilon_{tot}$	$\frac{br_{Z^0\to\nu_i\bar{\nu}_i}}{br_{Z^0\to e^+e^-}}$	**total**
$p_T^\gamma > 100$ GeV	2.71	5.96	16.2	2.45	5.96	14.6
$p_T^\gamma > 400$ GeV	1.62		9.68	1.70		10.2

Table 5.9: Transformation factors for $\gamma+Z^0 \to \nu_i\bar{\nu}_i$ calibration.

5.5.5 Statistical and systematical limitations at high p_T

The total number of expected events from $\gamma+Z^0 \to \mu^+\mu^-$ and $\gamma+Z^0 \to e^+e^-$ in the high- and low-p_T range ($p_T > 400$ GeV and $p_T > 100$ GeV respectively) as well as the number of $\gamma+Z^0 \to \nu_i\bar{\nu}_i$ events that are used for the calibration are shown in table Table 5.10.

Due to the very small cross-section in the high-p_T range above 400 GeV the whole study has been extended down to the much lower $p_T > 100$ GeV cut to get more statistics. Doing this, however, raises the problem of how the distribution obtained can be extrapolated into the high-p_T range. If the Monte Carlo prediction for the shape of the p_T^γ distribution can be trusted it can be compared to the measured shape of the spectrum. Here no K-factors which might increase the expected statistics of the candle sample are taken into account. The K factor amounts for the ratio of a highly accurate cross-section calculation to a less accurate one, typically a leading-

order (LO) calculation. It is defined as

$$K = \frac{\sigma_{\mathrm{NLO}}}{\sigma_{\mathrm{LO}}} \qquad (5.8)$$

if one considers the Next-to-leading order calculations(NLO) with respect to leading order calculations (LO).

Events		$p_T^\gamma > 100$ GeV			$p_T^\gamma > 400$ GeV		
			observ.	stat.		observ.	stat.
sample	$\mathcal{L}_{int}$	all	events	error	all	events	error
$\gamma + Z^0 \rightarrow \mu^+ \mu^-$	$10 fb^{-1}$	485	177	7.5%	3.8	2.2	67%
	$30 fb^{-1}$	1460	530	4.3%	11.4	6.7	38%
$\gamma + Z^0 \rightarrow e^+ e^-$	$10 fb^{-1}$	485	196	7.1%	3.8	2.6	61%
	$30 fb^{-1}$	1460	590	4.1%	11.4	8.0	35%
combined	$10 fb^{-1}$	970	390	5.1%	7.6	5.3	45%
	$30 fb^{-1}$	2910	**1170**	**2.9%**	23	16	26%
$\gamma + Z^0 \rightarrow \nu_i \bar{\nu}_i$	$10 fb^{-1}$				23	21	22%
	$30 fb^{-1}$				69	**62**	**13%**

Table 5.10: Number of events (efficiency estimations from OSCAR/ORCA) and resulting statistical uncertainty.
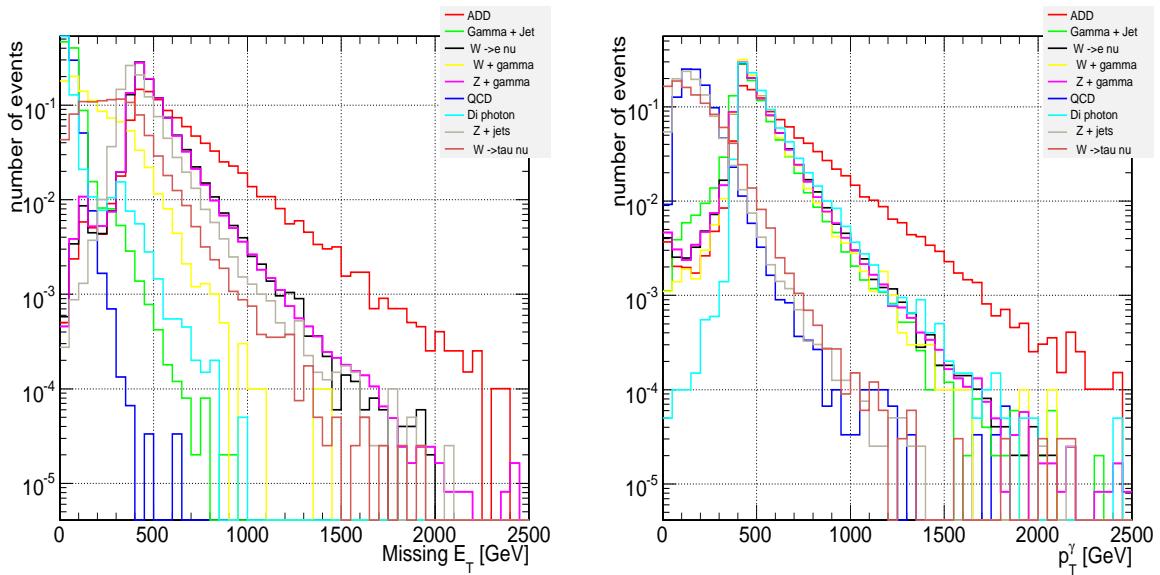
The acceptance correction function has been obtained using the leading order event generator PYTHIA. This contributes an unknown systematics uncertainty from the Monte Carlo calculations that cannot be corrected away by the calibration. The acceptance correction relies on the correct prediction of the angle distribution between the Z^0 and the photon at different energies. Since both particles are not charged the error is estimated to be small but the availability of next to leading order (NLO) calculations would improve the situation.

5.6   Trigger path

The topology of signal events is simple. The main trigger path will be the single photon trigger, both at the fast Level 1 trigger(L1) and the High Level Trigger(HLT). Presently the single photon trigger has a HLT level threshold of 80 GeV, which is far below the selection cut for events with isolated photons above 400 GeV. Hence the expected trigger efficiency is close to 100%. The efficiency can be monitored from data with a $\not{E}_T$ trigger, which will have a threshold in the range of 200-300 GeV, well below the acceptance of the bulk of the signal data. The trigger in CMS are described in [6].
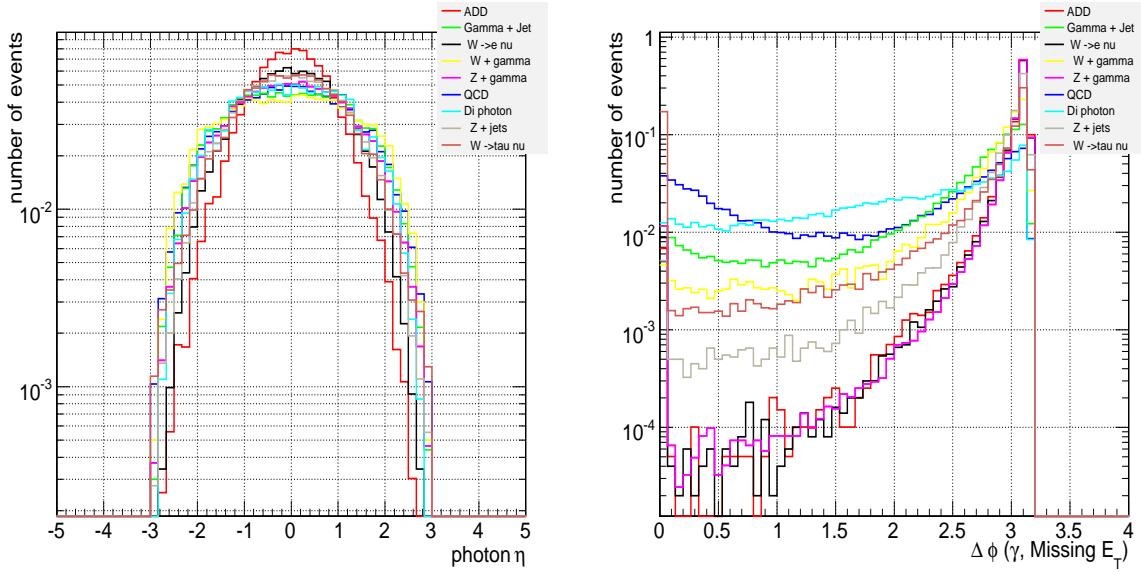
5.7 Analysis path and cut efficiency on signal and backgrounds

Besides of the kinematic cut on the partonic centre of mass p_T $(\widehat{p_T})$ at generator level, only photons with a transverse momentum larger than $15\,\text{GeV}$ have been considered in this analysis, since only very high-energetic photons are relevant for this study. With a simple set of cuts already a notable suppression of the backgrounds is possible. Depending on the model parameters a more or less significant excess of $\gamma + \not{E}_T$ events can be observed. The following analysis cuts have been chosen:

Figure 5.21: Normalised distributions for signal and background for $\not{E}_T$ (left) and p_T^γ (right).

1. At least $\not{E}_T > 400\,\text{GeV}$ is requested. This cut significantly reduces the QCD, the γ+jets and di-photon background where no high $\not{E}_T$ is expected. The normalised $\not{E}_T$ distributions for signal (as an example signal a scenario with $\text{M}_\text{D} = 5\,\text{TeV}$, $n = 2$ is chosen for the following plots) and background can be seen in Figure 5.21.

2. The photon p_T has to be above $400\,\text{GeV}$, too. This reduces the background with softer photons as can be seen in Figure 5.21.

3. The final state photon and graviton are back-to-back -therefore a cut on the difference in ϕ can be applied to reduce background which do not have this characteristic, see Figure 5.22. We demand a $\Delta\phi(\not{E}_T, \gamma) > 2.5$.

4. Since the signal photons are produced in the central detector region (Figure 5.22), a $\|\eta\| \leq 2.4$ is required.
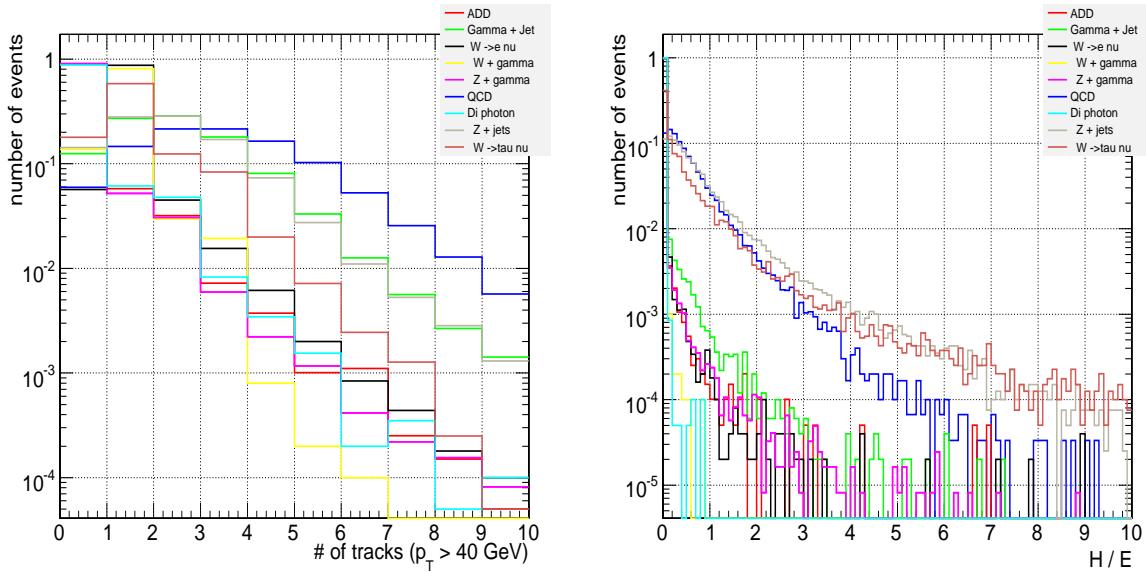
Figure 5.22:   Normalised distributions for signal and background showing the pseudorapidity of the photon, η^γ, (left) and the difference $\Delta\phi$ between the photon and $\not{E}_T$ (right).

5. A track veto for high p_T tracks > 40 GeV is applied. This is a powerful criterion to reduce all background containing high-energetic charged particles (e, μ, jets) (see Figure 5.23.)

6. During the analysis a contamination with fake photons originating from jets has been detected, which results in a non negligible background contribution due to the high cross-section. Therefore, an Isolated Photon Likelihood L has been applied as well.

To reduce the backgrounds containing jets an H/E cut or a cut on the number of jets have been also studied. H/E is the ratio of the energy deposited in the hadronic calorimeter divided by the energy deposited in the electromagnetic calorimeter. This criterion is well suited to distinguish processes containing photons from jets, which have naturally higher H/E values, as can be seen in Figure 5.23. This cut significantly reduces the QCD and γ+jets background, which are already highly suppressed by the previous cuts. It does also suppress the Z^0 + jets background, but due to the high cross-section the number of remaining events is still very large, i.e. about the same number of events as the main background $Z^0 \rightarrow \nu\bar{\nu}$. Therefore an *Isolated Photon Likelihood* has been introduced to reject jets faking photons or non-isolated photons coming from jets. It was designed following the example of the Electron Likelihood in ORCA and calculates the Likelihood from a set of reference histograms for signal and background. It uses the following input variables:

- $\frac{EMax}{E3x3}$, i.e. the ratio of the energy deposition in the highest-energetic ECAL

Figure 5.23: Normalised distributions for signal and background showing the number of tracks with a $p_T > 40$ GeV (left) and H/E for the most energetic supercluster (right.)

crystal relative to the 3x3 matrix as shower shape variable to suppress pions.

- $\frac{E3x3}{E5x5}$ to also take the energy deposition in the 3x3 matrix with respect to the 5x5 matrix into account.

- The total momentum of all tracks around the photon in a $\Delta R < 0.3$ cone with $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

- The relative amount of energy in the hadronic calorimeter(HCAL) in all clusters around the photon in a $\Delta R < 0.3$ cone compared to the energy deposited in the ECAL.

- The distance to the nearest track.

With this approach the misidentification of jets as photons can be completely suppressed. In a small fraction of Z^0+jets events one of the quarks can radiate an isolated high-energetic photon while the jet is very soft and not reconstructed, which makes the event look like a signal candidate and irreducible. This topology is very unlikely, but due to the high total cross-section of the Z^0+ jets prodcution it still delivers a non negligible contribution. The candle calibration method from data presented in the last section will take this type of events intrinsically into account. For all data samples the signal acceptance and background rejection have been evaluated. Signal samples corresponding to ADD scenarios with 2, 3, 4, 5 and 6 extra dimensions have been investigated - as for the second model parameter, the fundamental scale M_D, it turns out that M_D is only a scale factor of the cross-section
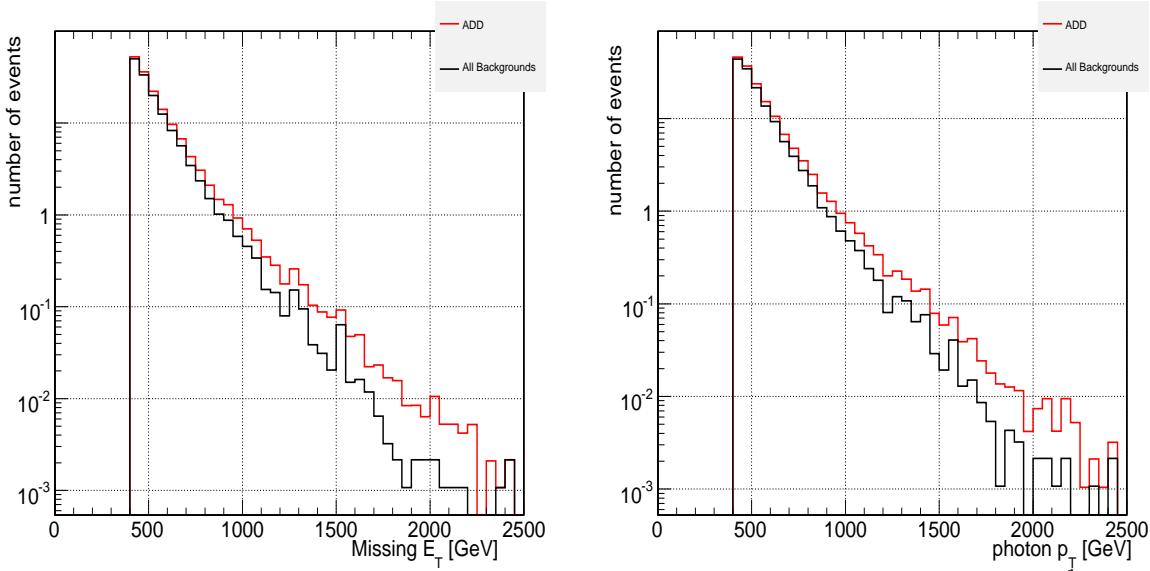
and does not distort the distributions - therefore, for different M_D's only the total number of expected events has been scaled, since the selection efficiency remains constant. The calculation of the number of expected ADD events is challenging from the theoretical point of view: a fraction of the events has a partonic centre of mass energy above the effective Planck scale, which leads to transplanckian graviton production. The ADD model, however, is valid only below M_D which is the scale where gravity becomes strong and only a (not available) theory of quantum gravity would be able to make predictions in this region (see 4.4.5). Therefore, the ADD cross-sections are rescaled by an acceptance factor α, which only chooses events with a graviton mass below the effective Planck scale, $M_D > m_G$ (hard truncation). The (rescaled) cross-sections of the ADD signal and its major backgrounds, the cut performance and the number of expected events for 30 fb^{-1} and 60 fb^{-1} are summarised in Table 5.11.

Sample	$\not{E}_T >$ 400 GeV	$p_T^\gamma >$ 400 GeV	$\|\eta_\gamma\|$ < 2.4	$\Delta\Phi$ > 2.5	$\not{\exists}$ **track**$_{p_T}$ > 40 GeV	$L >$ 0.2	**events** 30 fb^{-1}
ADD	88.60%	85.52%	85.52%	84.67%	77.40%	75.10%	8.1
$\gamma + Z^0$	81.29%	75.66%	74.61%	74.11%	68.44%	67.42%	43.7
$\gamma + W^\pm$	8.59%	8.42%	8.39%	8.35%	3.35%	3.32%	0.8
QCD	0.01%	0.01%	0.01%	0.01%	0.00%	0.00%	< 3
di-γ born	1.19%	1.16%	1.16%	1.12%	1.00%	0.98 %	1.5
di-γ box	0.75%	0.61%	0.61%	0.44%	0.34%	0.34%	0.01
$W^\pm \to e\nu$	82.27%	76.05%	75.75%	75.11%	3.96%	3.50%	19.1
$W^\pm \to e\mu$	88.34%	0.20%	0.19%	0.09%	0.00%	0.00%	< 3
$W^\pm \to e\tau$	21.15%	4.21%	4.20%	4.11%	0.92%	0.40%	2.2
$\gamma +$jets	0.31%	0.05%	0.05%	0.03%	0.01%	0.00%	< 3
$Z^0 +$jets	52.86%	2.78%	2.76%	2.59%	0.29%	0.04%	8.2

Table 5.11: Signal ($M_D = 5$ TeV, $n = 2$) and background efficiency for the applied cuts and number of expected events for an integrated luminosity of 30 fb^{-1}.

.

A detailed study of the expected signal events for a set of sample points in the M_D, n parameter space has been performed using Pythia. In Table 5.12 the total cross-sections of the ADD Graviton + Photon production are listed. As described above, the cross-section are truncated and events with $M_D < m_G$ are rejected, since they have been produced in the trans-Planckian region. The acceptance naturally gets smaller at lower values of M_D. Since the Graviton gets heavier with increasing number of extra dimensions the acceptance also gets lower with increasing n - this can be seen in Table 5.13. The influence of this hard truncation method is shown in table Table 5.14, where the effective cross sections are listed. In the next table
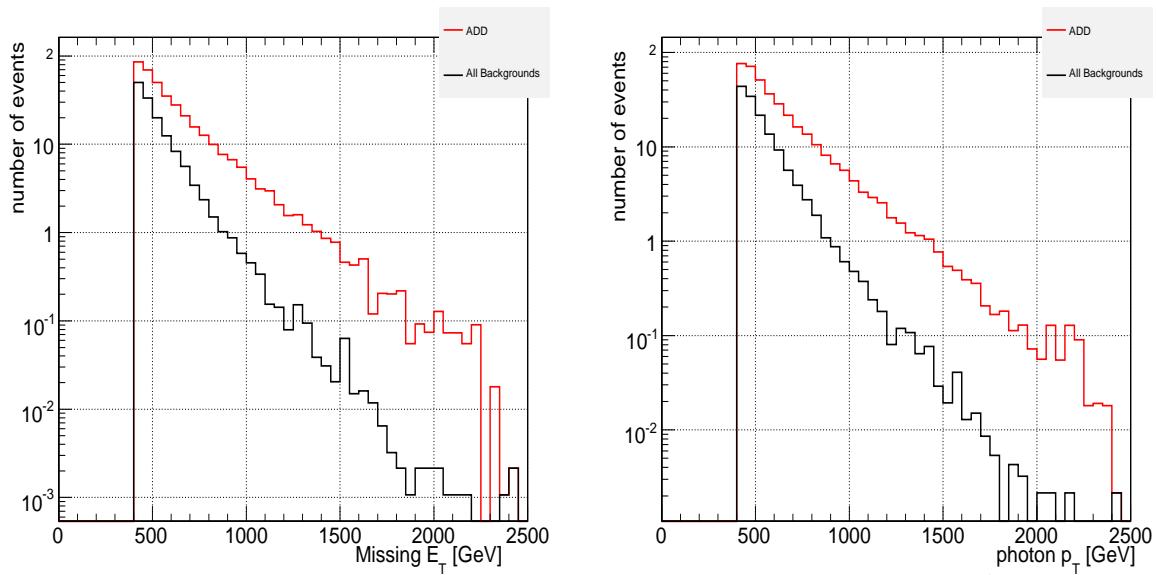
Table 5.15 the cut efficiency ϵ - i.e. the percentage of signal events surviving all applied cuts - is shown. With this strategy one can calculate the number of expected events as $N_{exp} = \sigma_{\mathrm{tot}} * \alpha(\mathrm{m_G}) * \epsilon$.

Figure 5.24: Signal and all backgrounds for $\not{E}_T$ after all cuts normalised to 60 fb^{-1} for a fundamental scale $\mathrm{M_D} = 5$ TeV, $n = 2$ ($\not{E}_T$ left, reconstructed photon p_{T} right).

The signal would show up as an excess over the expected number of Standard Model background events - this is exemplified in Figure 5.24 and Figure 5.25 , where the photon spectrum and the $\not{E}_T$ spectrum are shown in the case of a discovery of a $\mathrm{M_D} = 5$ TeV, $n = 2$ and $\mathrm{M_D} = 2.5$ TeV, $n = 2$ scenario. In Table 5.17 the significance Sig $= 2(\sqrt{\mathrm{S+B}} - \sqrt{\mathrm{B}})$ is calculated for each ADD scenario. It can be seen that up to $\mathrm{M_D} = 3$ a $5\,\sigma$ discovery for all n is possible. It should be noted that due to the hard truncation this is a conservative approach and should be considered as lower bound for the expected significances. A less conservative approach is to reduce the cross-section by a damping factor. This has been applied for example by ATLAS [116] using the damping factor $M_D^4/\hat{s}^2$ when $\hat{s}^2 > M_D^2$ (soft truncation).

Based on the calculated significances in Table 5.17, the integrated luminosity necessary for a $5\,\sigma$ discovery can be calculated and is shown in Table 5.18. If an ADD scenario with a low $\mathrm{M_D} < 3$ TeV is realized in nature, a discovery would be possible even in the first years of the LHC data taking. Disentangling the number of extra dimensions however is going to be challenging. The reach of CMS to find extra dimensions in the graviton and photon channel for 30 fb^{-1} and 60 fb^{-1} is shown in Figure 5.26 and Figure 5.27. For comparison: a study of the same channel in ATLAS [116] claims significances of 6.8 or 2.8 (conservative estimate) for $M_D = 4$ TeV, $n = 2$ after 100 fb^{-1}, in this study we obtain a value of 3.8 (without considering the systematics). However, the ATLAS study considers only the main background
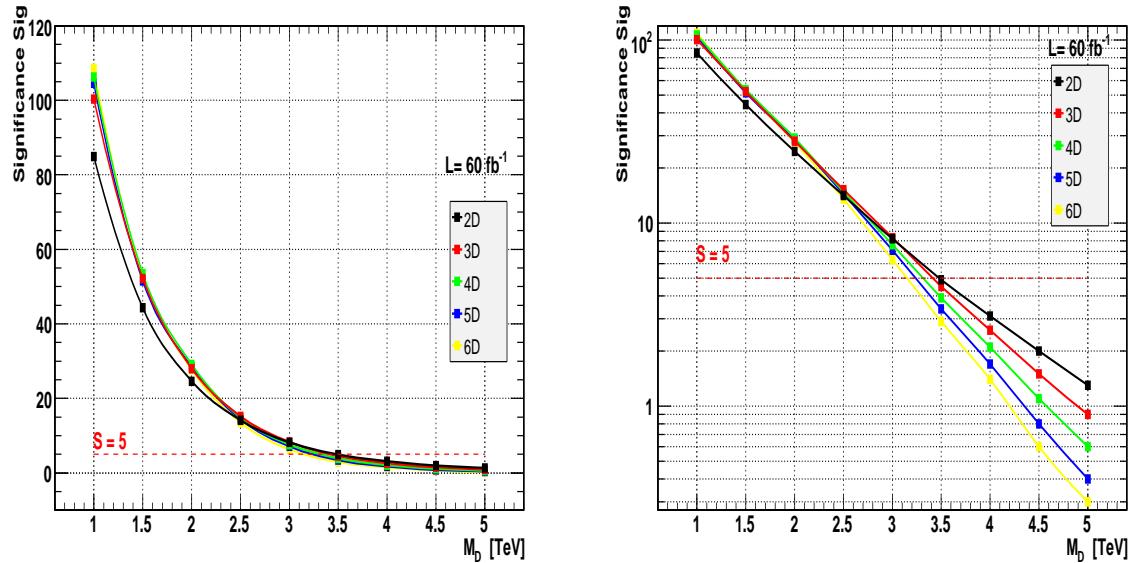
Figure 5.25: Signal and all backgrounds after all cuts normalised to 30 fb^{-1} for a fundamental scale $M_D = 2.5\,\mathrm{TeV}, n = 2$ ($\not{E}_T$ left, reconstructed photon p_T right).
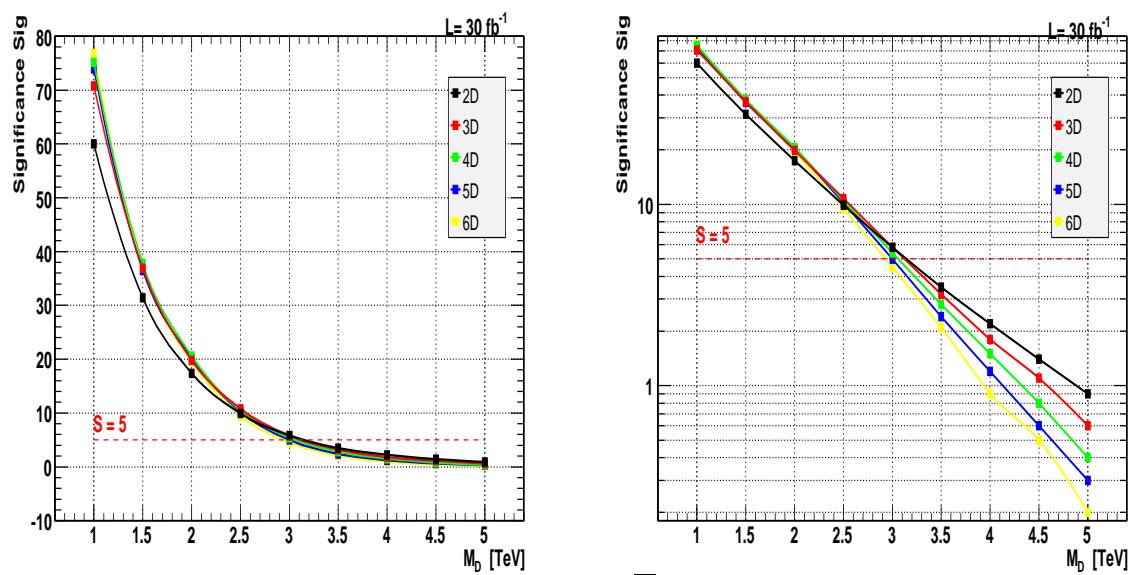
$Z^0 \to \nu\bar{\nu}$.

M_D / n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
$M_D = 1.0\,\mathrm{TeV}$	0.22 pb	0.75 pb	2.69 pb	10.07 pb	39.18 pb
$M_D = 1.5\,\mathrm{TeV}$	43.81 fb	99.28 fb	0.23 pb	0.59 pb	1.52 pb
$M_D = 2.0\,\mathrm{TeV}$	13.86 fb	23.56 fb	42.10 fb	78.64 fb	153.0 fb
$M_D = 2.5\,\mathrm{TeV}$	5.67 fb	7.72 fb	11.03 fb	16.49 fb	25.67 fb
$M_D = 3.0\,\mathrm{TeV}$	2.73 fb	3.10 fb	3.69 fb	4.60 fb	5.97 fb
$M_D = 3.5\,\mathrm{TeV}$	1.47 fb	1.43 fb	1.46 fb	1.56 fb	1.74 fb
$M_D = 4.0\,\mathrm{TeV}$	0.86 fb	0.73 fb	0.65 fb	0.61 fb	0.59 fb
$M_D = 4.5\,\mathrm{TeV}$	0.54 fb	0.40 fb	0.32 fb	0.27 fb	0.23 fb
$M_D = 5.0\,\mathrm{TeV}$	0.35 fb	0.24 fb	0.17 fb	0.12 fb	0.10 fb

Table 5.12: Total ADD cross section σ_{tot} for different M_D, n parameter values.

Figure 5.26: Significance Sig $= 2(\sqrt{S + B} - \sqrt{B})$ after an integrated luminosity of 60 fb^{-1}.

Figure 5.27: Significance Sig $= 2(\sqrt{S + B} - \sqrt{B})$ after an integrated luminosity of 30 fb^{-1}.

$M_D\,/n$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$M_D = 1.0$ TeV	26.46%	10.21%	3.23%	0.80%	0.23%
$M_D = 1.5$ TeV	49.34%	27.13%	12.15%	4.76%	1.95%
$M_D = 2.0$ TeV	68.48%	46.88%	27.62%	14.73%	7.24%
$M_D = 2.5$ TeV	81.50%	64.28%	44.09%	28.91%	17.16%
$M_D = 3.0$ TeV	89.74%	77.84%	60.68%	44.94%	30.61%
$M_D = 3.5$ TeV	94.53%	86.69%	73.46%	59.96%	45.26%
$M_D = 4.0$ TeV	97.22%	92.69%	83.48%	73.00%	60.55%
$M_D = 4.5$ TeV	98.74%	96.11%	90.62%	83.24%	73.88%
$M_D = 5.0$ TeV	99.40%	97.91%	94.85%	90.51%	83.61%

Table 5.13: Acceptance $\alpha(M_G)$ for signal events required to have $M_G < M_D$ in order to select only events from the region where the effective ADD theory is valid.

$M_D\,/n$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$M_D = 1.0$ TeV	58.0 fb	76.5 fb	86.8 fb	80.5 fb	90.1 fb
$M_D = 1.5$ TeV	21.6 fb	26.96 fb	27.8 fb	28.0 fb	29.8 pb
$M_D = 2.0$ TeV	9.48 fb	11.0 fb	11.6 fb	11.1 fb	11.1 fb
$M_D = 2.5$ TeV	4.6 fb	4.97 fb	4.85 fb	4.77 fb	4.31 fb
$M_D = 3.0$ TeV	2.43 fb	2.38 fb	2.21 fb	2.07 fb	1.82 fb
$M_D = 3.5$ TeV	1.38 fb	1.23 fb	1.07 fb	0.93 fb	0.78 fb
$M_D = 4.0$ TeV	0.83 fb	0.67 fb	0.54 fb	0.44 fb	0.35 fb
$M_D = 4.5$ TeV	0.53 fb	0.39 fb	0.29 fb	0.22 fb	0.17 fb
$M_D = 5.0$ TeV	0.35 fb	0.24 fb	0.16 fb	0.11 fb	0.09 fb

Table 5.14: Effective ADD cross section after truncation for different M_D,n parameter values ($\sigma_{\mathrm{eff}} = \sigma_{\mathrm{tot}} * \alpha$).

M_D/n	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
M_D= 1 TeV	77.6 %	77.9 %	78.0 %	78.6 %	69.6 %
M_D=1.5 TeV	76.0 %	78.5 %	77.0 %	74.2 %	70.3 %
M_D=2 TeV	75.6 %	77.8 %	77.7 %	75.9 %	75.4 %
M_D=2.5 TeV	75.4 %	77.8 %	76.7 %	75.2 %	75.3 %
M_D=3.0 TeV	75.2 %	77.2 %	76.1 %	74.9 %	74.6 %
M_D=3.5 TeV	72.5 %	76.9 %	76.1 %	75.3 %	74.6 %
M_D=4. TeV	75.2 %	76.7 %	75.8 %	75.1 %	74.1 %
M_D=4.5 TeV	75.2 %	76.8 %	75.5 %	75.3 %	74.2 %
M_D=5. TeV	75.1 %	76.8 %	75.6 %	75.2 %	73.8 %

Table 5.15: Accepted ADD signal events after all cuts for different sampling points in the M_D, n space.

M_D / n	n =2	n=3	n=4	n=5	n= 6
M_D = 1.0 TeV	2726 /1363	3594/1797	4034/2017	3799/1899	3784/1892
M_D = 1.5 TeV	984/492	1267/633	1322/661	1232/616	1257/628
M_D = 2.0 TeV	430/215	514/257	541/270	526/263	501/250
M_D = 2.5 TeV	210/104	231/115	223/111	215/107	200/99
M_D = 3.0 TeV	110/55	111/56	102/51	92/46	82/41
M_D = 3.5 TeV	60/30	57/29	49/24	42/21	36/17
M_D = 4.0 TeV	37/19	32/15	25/12	20/10	16/8
M_D = 4.5 TeV	24/12	18/9	13/6	10/5	8/4
M_D = 5.0 TeV	16/8	11/5	7/3	5/3	4/2

Table 5.16: Number of expected events after an integrated luminosity of 60 fb^{-1} and 30 fb^{-1}.

M_D /n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
$M_D = 1.0$ TeV	82.9/58.6	97.9/69.3	104.9/74.2	101.3/71.6	101.1/71.4
$M_D = 1.5$ TeV	42.9/30.4	50.9/35.9	52.3/37.0	49.9/35.3	50.6/35.8
$M_D = 2.0$ TeV	23.7/16.7	27.1/19.2	28.1/19.9	27.6/19.5	26.6/18.8
$M_D = 2.5$ TeV	13.4/9.5	14.6/10.4	14.2/10.0	13.7/9.7	12.9/9.1
$M_D = 3.0$ TeV	7.8/5.5	7.9/5.6	7.3/5.2	6.7/4.7	5.9/4.2
$M_D = 3.5$ TeV	4.5/3.2	4.3 /3.0	3.7/2.6	3.3/2.3	2.7/1.9
$M_D = 4.0$ TeV	2.9/2.1	2.4/1.7	1.9/1.4	1.6/1.1	1.3/0.9
$M_D = 4.5$ TeV	1.9/1.3	1.5/1.0	1.1/0.7	0.8/0.6	0.6/0.4
$M_D = 5.0$ TeV	1.3/0.9	0.9/0.6	0.6/0.4	0.4/0.3	0.3/0.2

Table 5.17: Significance Sig $= 2(\sqrt{S+B} - \sqrt{B})$ for an integrated luminosity of 60 fb^{-1} and 30 fb^{-1}.

M_D /n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
$M_D = 1.0$ TeV	0.21 fb^{-1}	0.15 fb^{-1}	0.13 fb^{-1}	0.14 fb^{-1}	0.14 fb^{-1}
$M_D = 1.5$ TeV	0.81 fb^{-1}	0.57 fb^{-1}	0.55 fb^{-1}	0.60 fb^{-1}	0.58 fb^{-1}
$M_D = 2.0$ TeV	2.6 fb^{-1}	2.0 fb^{-1}	1.8 fb^{-1}	1.9 fb^{-1}	2.1 fb^{-1}
$M_D = 2.5$ TeV	8.2 fb^{-1}	7.0 fb^{-1}	7.4 fb^{-1}	7.9 fb^{-1}	8.8 fb^{-1}
$M_D = 3.0$ TeV	24.4 fb^{-1}	24.0 fb^{-1}	28.1 fb^{-1}	33.3 fb^{-1}	41.9 fb^{-1}
$M_D = 3.5$ TeV	72.0 fb^{-1}	80.2 fb^{-1}	107.0 fb^{-1}	141.2 fb^{-1}	199 fb^{-1}
$M_D = 4.0$ TeV	173.0 fb^{-1}	249.0 fb^{-1}	387.8 fb^{-1}	581.3 fb^{-1}	904 fb^{-1}
$M_D = 4.5$ TeV	413.9 fb^{-1}	720.1 fb^{-1}	1310 fb^{-1}	2242 fb^{-1}	3884 fb^{-1}
$M_D = 5.0$ TeV	903.3 fb^{-1}	1846.2 fb^{-1}	4147 fb^{-1}	8183 fb^{-1}	16343 fb^{-1}

Table 5.18: Integrated luminosity necessary for a 5σ discovery.

5.8   Systematic uncertainties

The estimated significances can be affected by systematic uncertainties of the measurement. If we assume that the measurement of the photon p_T^γ in the electromagnetic calorimeter has an uncertainty of 2%, the cut efficiencies will be modified. In this case the background increases by 3.1 %, corresponding to 2.3 events. (The numbers of events given in this section as example always corresponds to 30 fb^{-1}.) We also investigated the effect on the significance by a miss-measurement of the $\not{E}_T$ assuming an uncertainty of 5 %. Under this assumption the background gets larger by 4.0 % or 3 events. Another source of systematic uncertainty originates from the parton distribution function (PDF): The parton distribution functions of interacting particles describe the probability density for partons undergoing hard scattering at the hard process scale and taking a certain fraction of the total particle momentum. In this study, all cross sections and samples were obtained using CTEQ6L. In order to estimate the cross section uncertainties originating from PDF uncertainties in this analysis the master equations were used:

$$\Delta X_1 = \frac{1}{2}\sqrt{\sum_{i=1}^{N}(X_i^+ - X_i^-)^2}\,;\ \ \Delta X_2 = \frac{1}{2}\sqrt{\sum_{i=1}^{2N} R_i^2} \tag{5.9}$$

$$\Delta X_{\max}^+ = \sqrt{\sum_{i=1}^{N}[\max(X_i^+ - X_0, X_i^- - X_0, 0)]^2}\,; \tag{5.10}$$

$$\Delta X_{max}^- = \sqrt{\sum_{i=1}^{N}[\max(X_0 - X_i^+, X_0 - X_i^-, 0)]^2} \tag{5.11}$$
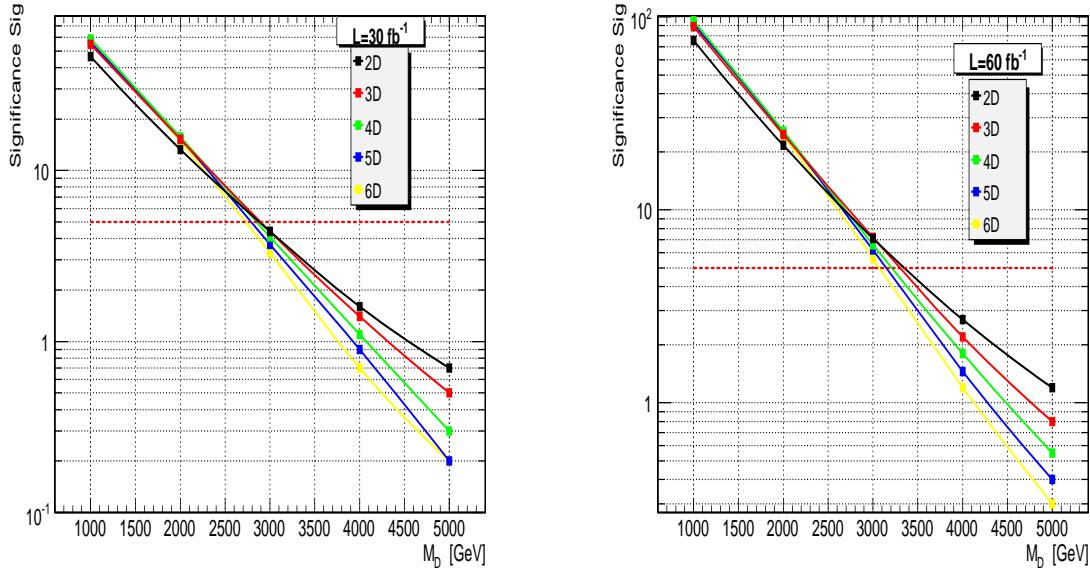
This leads to the following values :

$$W \to e\nu : \Delta X_1 = 7.81\%, \Delta X_2 = 8.64\%;\ \Delta X^+ = 8.47\%, \Delta X^- = 8.34\% \tag{5.12}$$

$$\gamma + Z \to \nu\bar{\nu} : \Delta X_1 = 7.92\%, \Delta X_2 = 8.81\%;\ \Delta X^+ = 8.13\%, \Delta X^- = 8.99\%. \tag{5.13}$$

If we assume the maximum uncertainty for these two main background components, the total background is increased by 7.5 % (5.6 events).

In conclusion, we have a total systematic error on the background of 9 %. The effect of the systematic error is shown in Figure 5.28 and Table 5.19, where the significances and the required luminosity for a 5 σ discovery are recalculated including systematics. On can see in Table 5.19 that with this background uncertainty and under the assumption of a Gaussian shape of the error a discovery with 5 σ is not possible anymore above around 3.0.

Figure 5.28: Significance Sig $= 2(\sqrt{S + B} - \sqrt{B})$ for an integrated luminosity of 30 and 60 fb^{-1} including systematic uncertainties.

M_D / n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
$M_D = 1.0$ TeV	0.21 fb^{-1}	0.16 fb^{-1}	0.14 fb^{-1}	0.15 fb^{-1}	0.15 fb^{-1}
$M_D = 1.5$ TeV	0.83 fb^{-1}	0.59 fb^{-1}	0.56 fb^{-1}	0.61 fb^{-1}	0.59 fb^{-1}
$M_D = 2.0$ TeV	2.8 fb^{-1}	2.1 fb^{-1}	1.9 fb^{-1}	2.1 fb^{-1}	2.3 fb^{-1}
$M_D = 2.5$ TeV	9.9 fb^{-1}	8.2 fb^{-1}	8.7 fb^{-1}	9.4 fb^{-1}	10.9 fb^{-1}
$M_D = 3.0$ TeV	47.8 fb^{-1}	46.4 fb^{-1}	64.4 fb^{-1}	100.8 fb^{-1}	261.2 fb^{-1}
$M_D = 3.5$ TeV	**5 σ discovery not possible anymore**				

Table 5.19: Integrated luminosity necessary for a $5\,\sigma$ discovery including systematics.

Conclusion

In the preparation time before the LHC start-up in 2008, a multitude of simulated Monte Carlo events are needed in order to understand the detectors and perform physics analyses at a centre-of-mass energy of $\sqrt{s} = 14$ TeV. In this thesis several important contributions to the simulation software of the Compact Muon Solenoid experiment (CMS) were made. When simulating proton-proton collisions at LHC energies, a significant amount of time is spent for electromagnetic showers in the calorimeter. The solution presented in this thesis, namely the usage of shower parameterisation, can speed-up the simulation considerably without sacrificing too much precision. A package based on this concept, GFlash, was implemented in the general simulation framework of Geant4. The new GFlash implementation has proven to work with an impressive speed (Table 3.2) and acceptable accuracy. It can be employed in a homogenous or sampling calorimeter of any high energy physics experiment, or in another simulation application based on Geant4. In order to improve the simulation accuracy, automated tuning tools have been setup, which can be used as well in an experiment independent way to tune the parameterisation. GFlash has been successfully integrated and tested in the CMS detector simulation. The transverse and longitudinal shower profiles in the CMS calorimeter are well modeled to within 1-3%. The shower parameterisation allows for a significant gain in time performance in the CMS simulation with speed increases in the range of 2-10 for single electrons or photons. The speed-up for the simulation of a proton-proton collision including the full detector geometry depends on the event type, especially on the energy and the angles of the particles hitting the detector. For example, the process $pp \rightarrow \gamma + G$ (graviton) with a single high energetic photon with $p_T^\gamma > 400$ GeV, is simulated around 3.3 times faster (Table 3.9).

The graviton emission process $pp \rightarrow \gamma + G$ mentioned above gives rise to a final state with a photon and missing transverse energy, a possible signature of large extra spatial dimensions. This signature studied in this thesis is forseen by a model that assumes the existence of n large extra dimensions (up to a fraction of a millimeter) and predicts effects that might be detectable at the LHC, if the model parameter M_D, the effective Planck scale, is in the TeV range. Simulation studies were performed with signal samples for various model parameters M_D and n and a multitude of possible background samples. Since a large number of simulated events

was needed, the standalone CMS fast simulation has been used for this study. The reconstruction performance and efficiency obtained with the fast simulation has been verified to compare with the detailed, Geant4 based, simulation. A normalisation method is proposed to measure the main background, the di-boson production of a photon and a Z^0 ($Z^0 \to \nu\bar{\nu}$), with high precision using reference rates and spectra from $\gamma + Z^0(\to \mu\mu)$ and $\gamma + Z^0(\to ee)$. This "Candle" calibration allows to control the background in the region of interest to about 5% after an intergrated luminosity of 10 fb^{-1} (Table 5.10). A 5 σ discovery can be made with less than 1 fb^{-1} of data for scenarios with M_D in the range of 1.0-1.5 TeV, and less than 10 fb^{-1} for values of M_D in the range of 2.0-2.5 TeV, largely independent of the number of extra dimensions (Figure 5.28). The discovery reach via this channel with 60 fb^{-1} is about 3.0-3.5 TeV. These estimates are conservative taking into account only the events for which the graviton mass is smaller than M_D and should be considered as a lower bound.

List of Figures

List of Tables

Appendix A

XCMSI - a CMS software installation tool

The product XCMSI is a tool to ease the installation of the entire CMS software on computing resources ranging from physicists' notebooks up to grid-enabled clusters. To ensure the installation of a software which is working and producing reliable results in such an inhomogeneous environment with different hardware and operating systems is not a trivial task, especially in the case of such a complex software as for the new LHC detectors. A typical CMS installation comprise in total around 75 compressed packages, which themselves take around 1 GB, the complete unpacked installation requires 4 GB of disk space. The solution presented here is based on perl for an automated production of RPM packages and xcmsi, a tool written in perl and perl/Tk, to facilitate installing, updating and verifying the RPM packaged software. The project web page with further information can be found at [57].

A.1   Features and requirements of XCMSI

The main parts of XCMSI are `xcmsi.pl`, a perl script using perl-Tk to provide a graphical interface for a user friendly configuration of the environment, and `cmsi.pl`, a perl script to do the actual installation. In addition, `cmsv.pl` serves to validate the installed software. The latter two can also be called in batch mode provided they are given access to a proper configuration file.

By design, the following requirements are met by this tool:

- Arbitrary installation directory `$CMS_PATH` to relocate packages

- No `root` privileges required

- Network access for automated downloads is helpful but not mandatory

- Command-line (batch mode) installation possible

- Separate validation suite included

- A graphical interface (GUI) for a concise configuration also by less experienced users

- Save-able and (re)loadable configuration files

- Configuration files directly usable to set the CMS environment for users

- Multi-platform support

- Support for multiple non-interfering installations

The tool is designed for a standard Linux PC with the following packages installed:

- perl, version 5.6.0 or higher

- perl-Tk

- rpm, version 3.0.6 or higher

Perl is mandatory but should be available on almost all Linux systems anyway. Perl-Tk is only necessary to use the GUI.
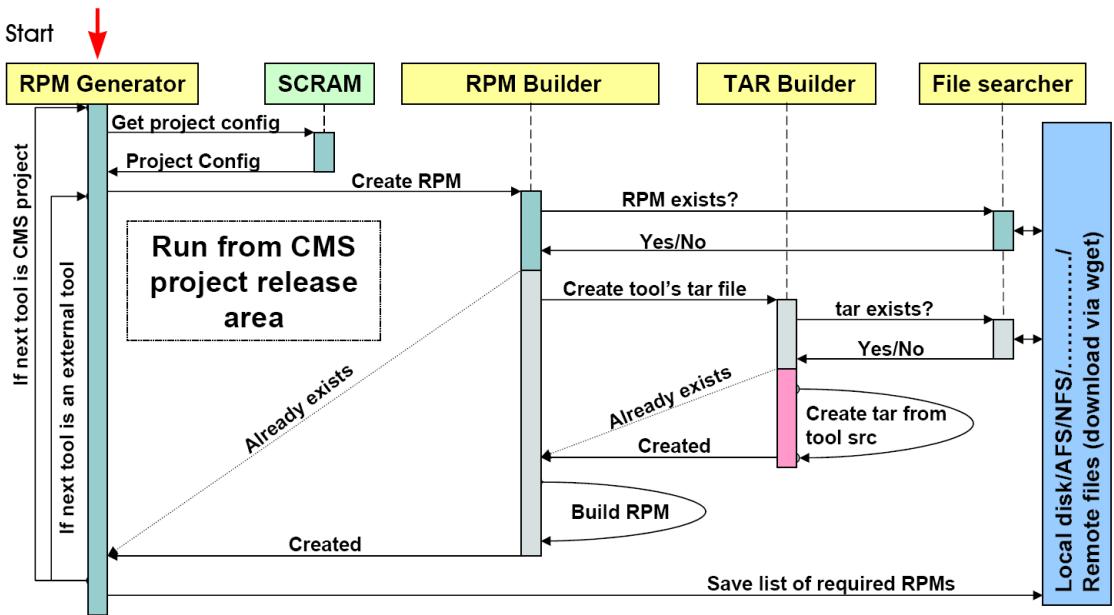
A.1.1   Generation of rpm-packages

The basic development and final testing of CMS software is performed on the central repositories at CERN. Therefore, all RPMs are based on this central installation of all experiment and external software. In total, this comprises the experiment specific programs of the CMS collaboration, the installed versions of packages of the LCG project [58] and the CMS specific installations of other external programs. In addition, an RPM of the SCRAM tool and the necessary compilers, currently the GNU compiler collection GCC 3.2.3, are provided. The central script to gather the additional intelligence from SCRAM and to prepare the dependency information is `ProjectDist.pl`, a schematic view of the procedure is presented in figure. A.1. Further scripts are employed to perform consistency checks, produce tar archives, take proper care of symbolic links and to finally generate the RPMs.

A.1.2   Configuration with xcmsi.pl

The main window of xcmsi (see figure. A.2) provides six rows and columns of push buttons representing approximately the corresponding sequence of steps to follow for a successful installation.

In the `Select Tags` section, the user can choose between download tags representing different CMS software projects — all necessary software for the chosen

Start

Figure A.1: Schematic view of the iterative RPM generation procedure initiated by
`ProjectDist.pl`. It differentiates between SCRAM managed projects,
where additional dependency information is extracted in order to run the
proper SCRAM set-up in the installation phase, and other tools where no
additional treatment is necessary.

project, even additional projects if required, is downloaded and installed. The currently available download tags are obtained from the CMS central webpage and can be chosen by clicking. The user can chose one or several tags.
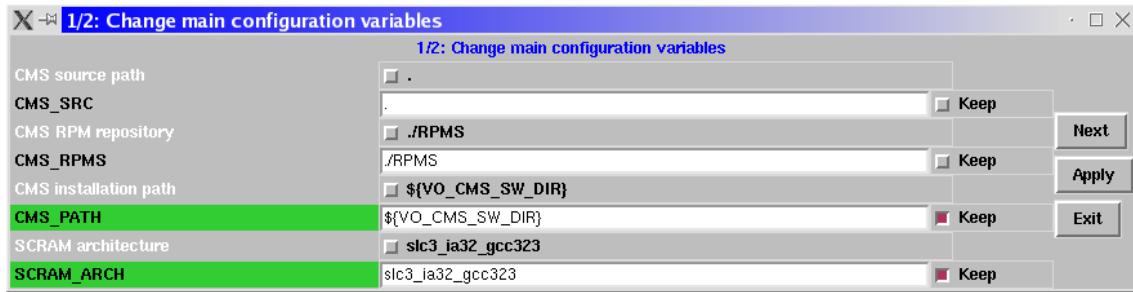
The Select Archives section shows the already installed RPMs and offers the possibility to choose new RPMs for the update mode. The last three buttons serve to load a previously saved configuration file and to edit and save the current settings. There are about 30 configurable settings but usually only three of them have to be adapted to the users' needs and only these three are shown in the first configuration window, which can be seen in figure. A.3 :

1. CMS_SRC, the location where XCMSI is to be found

2. CMS_RPMS, the directory for depositing all the RPMs

3. CMS_PATH, the software installation path

For new users, normally nothing has to be changed, if they have set ${VO_CMS_SW_DIR}, (a environment needed by LCG) correctly: $CMS_PATH is then set to ${VO_CMS_SW_DIR}/cms and all other variables depending on $CMS_PATH are propagated. If one changes $CMS_PATH again all variables depending on $CMS_PATH are propagated with the new value. These additional

Figure A.2: The main window of `xcmsi.pl`. The arrangement of push buttons from top to bottom and left to right represents approximately the sequence of steps to follow for normal usage.
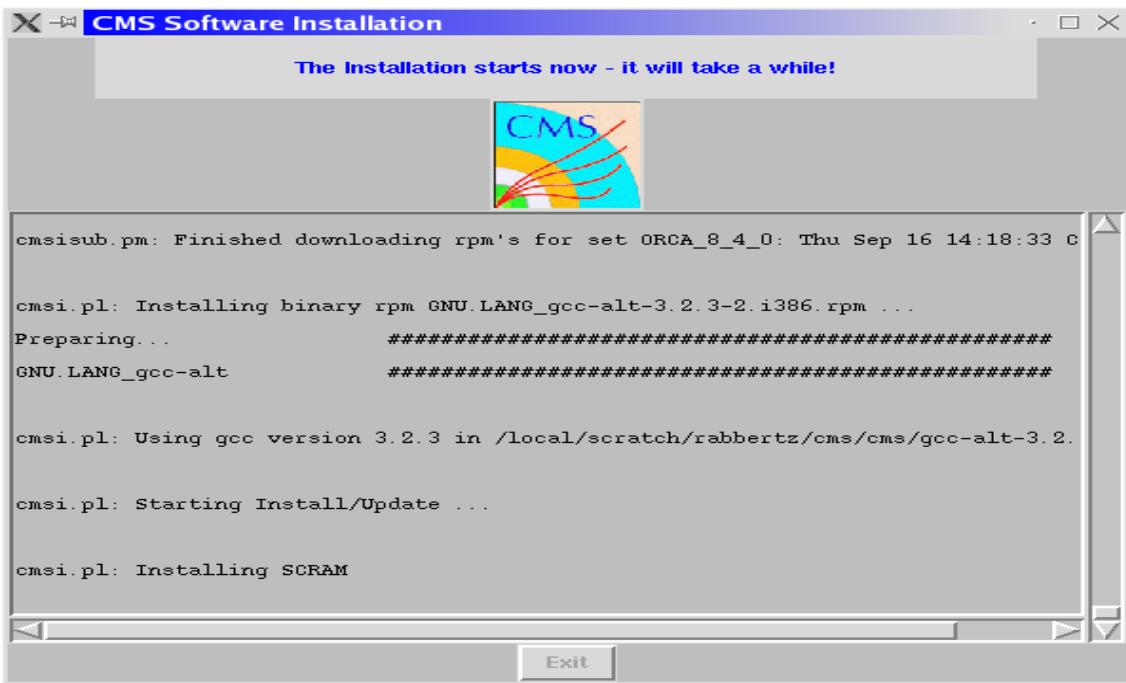
Figure A.3: The first ChangeConfig window of xcmsi.

settings are hidden behind an "Expert Options "button and should only be changed if necessary. This concept makes an installation for a new user very simple: he / she has only to set one top directory, everything else is taken care of by xcmsi !

Pushing the `Options` button additional features can be activated which are required for certain install versions. The `Select Tags` window actually enters the software projects to install in the "Download additional RPMs "entry field. If other Linux distributions than RedHat or Scientific Linux are tried, say SuSE or even Debian, it might be necessary to pass additional arguments to the RPM commands for the installation.

Finally, one can start, verify or update the installation and quit the program. Since the graphical user interface is employed only for the configuration itself, the settings have still to be passed to the perl scripts `cmsi.pl` and `cmsv.pl` which do the real work. Thereby it is ensured that the whole procedure can be performed from the command line (or in batch mode) without duplicating functionalities. In addition, the configuration files are designed in such a way that they can directly be used to set up a users environment for developing its code with the installed software. The user also has the possibility to uninstall packages by clicking on the corresponding tag under "Uninstall Packages ". With the "Recalculate Dependencies"button the consistency of the system should be insured. A sccrenshot of an

ongoing installation can be seen in figure A.4. As mentioned before, the installa-

Figure A.4: Installation window (The loging file is automatically created and saved).

tion can also be performed without the GUI. Currently, a "default "command of a
text mode installation could look like:

```
./cmsi.pl -f cmsset_default.csh -g config
% -d "ORCA_8_4_0 OSCAR_3_4_0"
```

In total, `cmsi.pl` accepts the following options:

```
cmsi.pl Version V0.8
Usage: cmsi.pl [switches/options]
              [list of rpm's]
  -d software set Download/install or erase
                  pre-defined software set(s),
                  multiple sets have to be
                  quoted ""!
  -e              Erase a set of rpms
  -f configfile   Use configfile instead of
                  cmsset_default.csh
  -g gcc-path     Adapt SCRAM to use gcc-path
                  or with special keyword
  -g config       adapt SCRAM to use gcc-path
                  from configfile, the new
```

```
                  default!
 -h               Print this text
 -i               Overrule installation abort
                  in case of existing rpmdb
 -m mirror        Use mirror instead of
                  CERN download server
 -r rpm options   Additional options to rpm
                  install command, multiple
                  options have to be
                  quoted ""!
 -u               Run in update mode
```

After a successful installation a validation procedure is offered, which checks at the moment the main ingredients of the software : the reconstruction toolkit ORCA and the simulation software OSCAR. The verification process just needs to know which configuration file to use. More details can be found in the online documentation.[57]

A.1.3   Installation via LCG
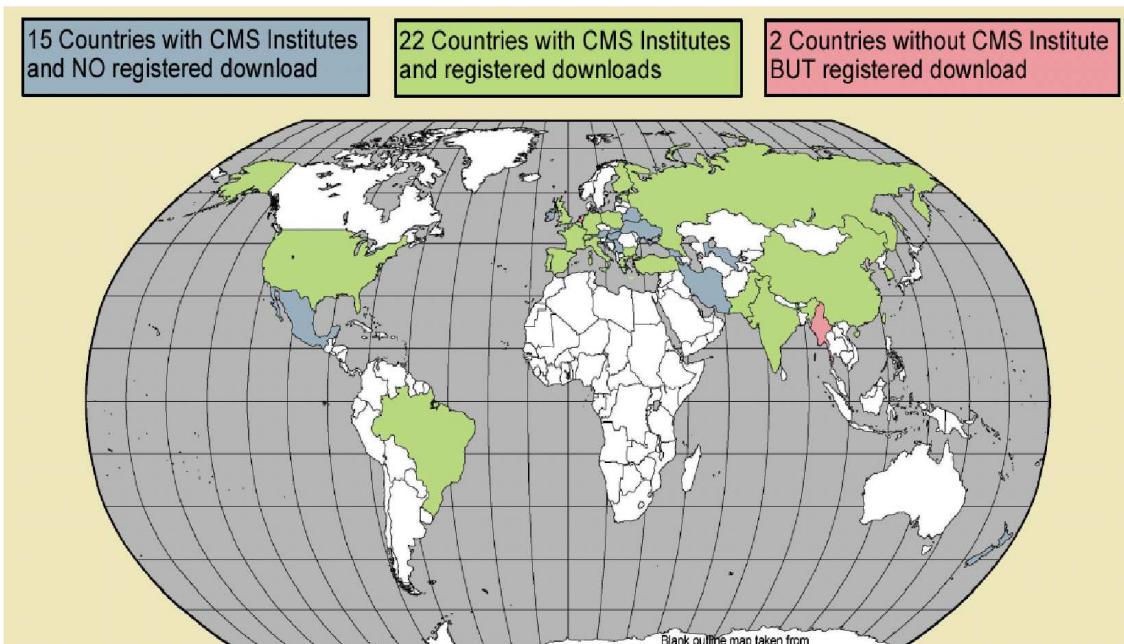
Every **V**irtual **O**rganization (VO) in the LCG scope has an **E**xperiment **S**oftware **M**anager (ESM) who is mapped onto a privileged grid account with write permissions on a specific software area. Currently, LCG sites have two distinct policies for software installation. The first one allows the ESM to install software on a shared area common to all **W**orker **N**odes where it is mounted via a network file system (NFS, GPFS). This shared area may reside on the site's **C**omputing **E**lement (CE). The second one without a shared area, foresees an installation procedure where the experiment software is first installed on one WN and then propagated to all the others. This second mechanism is still under test. From the point of view of disk space consumption and fault tolerance it is less favourable.

The grid installation job runs on a WN creating, on the shared area, the directory structure containing CMS software. The core tool which performs installations on LCG sites is `cmsi.pl`. Around it, a perl script (`CmsSwGridInstall.pl`) has been developed which provides some features to automatise the creation and submission of installation jobs.

Basically, the script queries the ldap server (GRIS) of every LCG Computing Element to retrieve published tags and decides the action to perform, either a brand new installation or an update. Once all information has been collected, it creates shell scripts, which are wrappers of `cmsi.pl`, and the necessary job description files (jdl files). All jdl's are then submitted and their status monitored to retrieve output files as soon as they finish.

Once the job finishes, the Information Service is made aware of the newly installed distribution publishing a new tag describing it. Every release is actually marked with a specific tag which describes the software contents. This system allows users to drive production or analysis jobs only on sites publishing the needed tag.

A.1.4   Current deployment in CMS

Figure A.5: World map showing the current deployment status of xcmsi.

The described procedure has routinely been used to distribute and install new releases of CMS software on more than 20 LCG2 sites in France, Germany, Great Britain, Italy, Pakistan, Russia, Spain, Switzerland (CERN) and Taiwan. Since by design the installations can be performed on any PC or notebook complete statistics on the usage are not available. As example, in two weeks the corresponding http download server has counted about 1000 hits per day, the counter for SCRAM downloads is at 185 meaning about 12 base installations per day. This, however, does not take into account double or incomplete downloads. Started in 2003 as private project, the xcmsi installation tool has become the official (see PTDR volume 1 [6]) and widely used installation tool within the CMS collaboration. Possible or foreseen improvements concern the grid installation and the validation procedures, automated updates and the extension to platforms that are not binary compatible to the CMS operating system of reference via source RPMs.

Appendix B

Summary of formulae

B.0.5 Fluctuated longitudinal profiles–original parameters

$$
\begin{aligned}
\langle \ln T \rangle &= \ln(\ln y - 0.812) \\
\sigma(\ln T) &= (-1.4 + 1.26 \ln y)^{-1} \\
\langle \ln \alpha \rangle &= \ln\left(0.81 + (0.458 + 2.26/Z) \ln y\right) \\
\sigma(\ln \alpha) &= (-0.58 + 0.86 \ln y)^{-1} \\
\rho(\ln T, \ln \alpha) &= 0.705 - 0.023 \ln y
\end{aligned}
$$

B.0.6 Fluctuated longitudinal profiles–tuned parameters

$$
\begin{aligned}
\langle \ln T \rangle &= \ln(1.10 \ln y - 1.508) \\
\sigma(\ln T) &= (-1.73 + 1.24 \ln y)^{-1} \\
\langle \ln \alpha \rangle &= \ln\left(-1.12 + (0.37 + 25.8/Z) \ln y\right) \\
\sigma(\ln \alpha) &= (1.9 + 0.6 \ln y)^{-1} \\
\rho(\ln T, \ln \alpha) &= 0.705 - 0.023 \ln y
\end{aligned}
$$

B.0.7   Average radial profiles

$$
\begin{aligned}
R_{C,}(\tau) &= z_1 + z_2\tau \\
R_{T,}(\tau) &= k_1\{\exp(k_3(\tau - k_2)) + \exp(k_4(\tau - k_2))\} \\
p(\tau) &= p_1 \exp\left\{\frac{p_2 - \tau}{p_3} - \exp\left(\frac{p_2 - \tau}{p_3}\right)\right\}
\end{aligned}
$$

with

$$
\begin{aligned}
z_1 &= 0.0251 + 0.00319\ln E \\
z_2 &= 0.1162 + -0.000381Z \\
k_1 &= 0.659 + -0.00309Z \\
k_2 &= 0.645 \\
k_3 &= -2.59 \\
k_4 &= 0.3585 + 0.0421\ln E \\
p_1 &= 2.632 + -0.00094Z \\
p_2 &= 0.401 + 0.00187Z \\
p_3 &= 1.313 + -0.0686\ln E
\end{aligned}
$$

Average radial profiles–corrected weight p

Correction factor $k(\tau) = (0.84 + (-0.033\,\tau))$;

Fluctuated radial profiles

$$
\begin{aligned}
\tau_i &= \frac{t}{\langle t\rangle_i}\frac{\exp(\langle\ln\alpha\rangle)}{\exp(\langle\ln\alpha\rangle) - 1} \\
N_{Spot} &= 93\ln(Z)E^{0.876} \\
T_{Spot} &= T(0.698 + 0.00212Z) \\
\alpha_{Spot} &= \alpha(0.639 + 0.00334Z)
\end{aligned}
$$

Bibliography

[1] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. `doi:10.1103/PhysRevLett.13.508`. 5, 89

[2] P. W. Higgs, "Broken symmetries, massless particlees and gauge fields," *Phys. Lett.* **12** (Sep, 1964) 132–133. `doi:10.1016/0031-9163(64)91136-9`. 5, 89

[3] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. `doi:10.1103/PhysRevLett.13.321`. 5, 89

[4] J. Weng, "GFLASH - parameterised electromagnetic shower in CMS," in *Proceedings of Computing in High Energy (CHEP 2006).* , Mumbai, India, 13-20 Feb, 2006. 6, 84

[5] J. Weng, E. Barberio, and A. Waugh, "Fast shower parameterisation for electromagnetic cascades," in *Proceedings of 9th ICATPP Conference on Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications.* , Villa Olmo, Como, Italy, 17-21 Oct, 2005. 6, 49, 84

[6] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 1," *CERN/LHCC* **2006-001** (2006). CMS TDR 8.1. 7, 9, 34, 84, 121, 132, 157

[7] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The hierarchy problem and new dimensions at a millimeter," *Phys. Lett.* **B429** (1998) 263–272, `arXiv:hep-ph/9803315`. 7, 29, 99, 109

[8] J. Weng, C. Saout, G. Quast, A. De Roeck, and M. Spiropulu, "Search for ADD Direct Graviton Emission in Photon plus Missing Transverse Energy Final State at CMS," **CMS NOTE-2006/129** (2006). 7

[9] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 2," *CERN/LHCC* **2006-021** (2006). CMS TDR 8.2. 7, 103

[10] CMS Collaboration, "The Tracker Project Technical Design Report," *CERN/LHCC* **98-006** (1998). CMS TDR 5, Addendum CERN/LHCC 2000-016. 18

[11] CMS Collaboration, "The Electromagnetic Calorimeter Technical Design Report," *CERN/LHCC* **97-033** (1997). CMS TDR 4, Addendum CERN/LHCC 2002-027. 19, 51

[12] CMS Collaboration, "The Hadron Calorimeter Technical Design Report," *CERN/LHCC* **97-031** (1997). CMS TDR 2. 19

[13] CMS Collaboration, "The Magnet Project Technical Design Report," *CERN/LHCC* **97-010** (1997). CMS TDR 1. 20

[14] CMS Collaboration, "The Muon Project Technical Design Report," *CERN/LHCC* **97-32** (1997). CMS TDR 3. 20

[15] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems," *CERN/LHCC* **2000-38** (2000). CMS TDR 6.1. 21

[16] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger," *CERN/LHCC* **2002-26** (2002). CMS TDR 6.2. 21

[17] V. Karimaki et al., "CMKIN v3 User's Guide," *CMS IN* **2004-016** (2004). 24, 111

[18] T. Sjostrand, L. Lonnblad, and S. Mrenna, "PYTHIA 6.2: Physics and manual," `arXiv:hep-ph/0108264`. 24, 26, 110

[19] "HBOOK Reference Manual." Located at `http://wwwasdoc.web.cern.ch/wwwasdoc/hbook/HBOOKMAIN.html`. 24

[20] A. Pukhov et al., "CompHEP: A package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33," `arXiv:hep-ph/9908288`. 24, 28, 116

[21] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph," *JHEP* **02** (2003) 027. `doi:10.1088/1126-6708/2003/02/027`. 24, 28, 116

[22] "COBRA: Coherent Object-oriented Base for simulation, Reconstruction and Analysis." Site located at `http://cmsdoc.cern.ch/cobra`. 24

[23] "OSCAR: CMS Simulation Package Home Page." Site located at `http://cmsdoc.cern.ch/oscar`. 24, 36, 81, 117

[24] **GEANT4** Collaboration, S. Agostinelli et al., "GEANT4: A simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250–303. 24, 35

[25] "Geant4." Located at `http://wwwinfo.cern.ch/asd/geant4/geant4.html`. 24, 35, 52

[26] "ORCA: CMS Reconstruction Package." Site located at `http://cmsdoc.cern.ch/orca`. 24, 117

[27] **CMS** Collaboration, D. Acosta et al., "CMS Physics TDR Volume 1, Section 2.6: Fast simulation," *CERN/LHCC* **2006-001** (2006) 55. 24, 35, 117

[28] "IGUANA Web Site." Located at `http://iguana.web.cern.ch/iguana/`. 25

[29] "SCRAM." Located at `http://cmsdoc.cern.ch/Releases/SCRAM/doc/scramhomepage.html`. 25

[30] "CVS." Located at `http://www.nongnu.org/cvs/`. 25

[31] T. Gleisberg et al., "SHERPA 1.alpha, a proof-of-concept version," *JHEP* **02** (2004) 056, `arXiv:hep-ph/0311263`. 25

[32] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 2, Appendix C," *CERN/LHCC* **2006-021** (2006) 537. 25

[33] G. Corcella et al., "HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)," *JHEP* **01** (2001) 010, `arXiv:hep-ph/0011363`. `doi:10.1088/1126-6708/2001/01/010`. 26

[34] F. E. Paige, S. D. Protopopescu, H. Baer, and X. Tata, "ISAJET 7.69: A Monte Carlo Event Generator for pp, $\bar{p}p$, and e^+e^- Reactions," `arXiv:hep-ph/0312045`. 26

[35] F. Caravaglios and M. Moretti, "An Algorithm to Compute Born Scattering Amplitudes without Feynman Graphs," *Physics Letters B* **358** (1995) 332. 27

[36] G. P. Lepage, "VEGAS: AN ADAPTIVE MULTIDIMENSIONAL INTEGRATION PROGRAM,". CLNS-80/447. 27, 28

[37] F. Krauss, R. Kuhn, and G. Soff, "AMEGIC++ 1.0: A matrix element generator in C++," *JHEP* **02** (2002) 044, `arXiv:hep-ph/0109036`. 28, 29

[38] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, "QCD matrix elements + parton showers," *JHEP* **11** (2001) 063, `arXiv:hep-ph/0109231`. 28

[39] F. Krauss, "Matrix elements and parton showers in hadronic interactions," *JHEP* **08** (2002) 015, `arXiv:hep-ph/0205283`. 28

[40] H. E. Haber and G. L. Kane, "The search for supersymmetry: Probing physics beyond the Standard Model," *Physics Reports* **117** (1985) 75–263, `arXiv:hep-ph/0109036`.
http://www.sciencedirect.com/science/article/B6TVP-46SWYSG-3V/2/50bf78d167cac42f6109d17f2d4ae8ef.
29

[41] R. Kleiss and W. J. Stirling, "Spinor techniques for calculating pp $W^{+/-}/Z^0$ + jets," *Nuclear Physics B* **262** (December, 1985) 235–262.
`doi:10.1016/0550-3213(85)90285-8`. 29

[42] A. Ballestrero, E. Maina, and S. Moretti, "Heavy quarks and leptons at e+ e- colliders," *Nucl. Phys.* **B415** (1994) 265–292, `arXiv:hep-ph/9212246`. 29

[43] "SHERPA." Located at `http://sherpa.de`. 29, 111

[44] "HepMC a C++ Event Record for Monte Carlo Generators." Located at `http://mdobbs.web.cern.ch/mdobbs/HepMC/`. 29

[45] CMS, "CMS Software Page.". 30, 84

[46] M. Erdmann, D. Hirschbühl, C. Jung, S. Kappler, Y. Kemp, M. Kirsch, D. Miksat, C. Piasecki, G. Quast, K. Rabbertz, P. Schemitz, A. Schmidt, T. Walter, and C. Weiser, "Physics Analysis Expert PAX: First Applications," in *Computing in High Energy and Nuclear Physics (CHEP03)*. La Jolla (USA), 2003. `arXiv:physics/0306085`. See also `http://pax.web.cern.ch/pax/`. 31, 111, 117

[47] "The ALEPH experiment." Located at `http://aleph.web.cern.ch/aleph`. 31, 105

[48] "The H1 experiment." Located at `http://www-h1.desy.de/`. 31, 35

[49] L. Lonnblad, "CLHEP: A project for designing a C++ class library for high-energy physics," *Comput. Phys. Commun.* **84** (1994) 307–316. 32

[50] "ROOT : an Object Oriented Data Analysis Framework." Information available at `http://root.cern.ch/`. 32

[51] M. Erdmann, U. Felzmann, D. Hirschbühl, C. Jung, S. Kappler, M. Kirsch, G. Quast, K. Rabbertz, J. Rehn, S. Schalla, P. Schemitz, A. Schmidt, T. Walter, and C. Weiser, "New Applications of PAX in Physics Analyses at Hadron Colliders," in *Computing in High Energy and Nuclear Physics (CHEP04)*. Interlaken (Switzerland), 2004. 33

[52] M. Erdmann, U. Felzmann, A. Flossdorf, S. Kappler, M. Kirsch, G. Mueller, G. Quast, C. Saout, A. Schmidt, and J. Weng, "Concepts, Developments and Advanced Applications of the PAX Toolkit," in *Computing in High Energy and Nuclear Physics (CHEP06)*. Bombay (India), 2006. `arXiv:physics/0605063`. 33

[53] M. Erdmann, U. Felzmann, D. Hirschbühl, S. Kappler, M. Kirsch, G. Quast, A. Schmidt, and J. Weng, "The PAX Toolkit and Its Applications at Tevatron and LHC," *IEEE Transactions on Nuclear Science* **53** (April, 2006) 506–512. 33

[54] C. Saout, "$Z\gamma$ Production in $\gamma\nu\nu$ and γl^+l^- Final States in pp Collisions at 14 TeV Centre-of-Mass Energy," Master's thesis, Institut für Experimentelle Kernphysik, Universität Karlsruhe (TH), 2006. 33, 126

[55] M. Galanti and A. Giammanco, "ExRootAnalysis,". 33, 119

[56] "wxWidgets- a Cross-Platform GUI library." Located at `http://www.wxwidgets.org/`. 33

[57] "XCMSInstall Download Page for CMS OO Projects." Located at `http://cern.ch/cms-xcmsi`. 34, 151, 156

[58] LCG Collaboration, "LHC computing Grid : Technical Design Report," *CERN/LHCC* **2005-024** (2005). LCG TDR 001, See also `http://lcg.web.cern.ch/LCG/`. 34, 152

[59] K. Rabbertz, J. Weng, and et. al., "CMS Software Packaging and Distribution Tools," in *Computing in High Energy and Nuclear Physics (CHEP06)*. Mumbai, India, 2006. 34

[60] A. Nowack, K. Rabbertz, J. Weng, and et. al., "CMS Software Distribution on the LCG and OSG Grids," in *Computing in High Energy and Nuclear Physics (CHEP06)*. Mumbai, India, 2006. 34

[61] K.Rabbertz, J.Weng, A.Nowack, A.Sciaba, M. S.Wynhoff, and Sh.Muzaffar, "CMS Software Installation," in *Computing in High Energy and Nuclear Physics (CHEP03)*. Interlaken, Switzerland, 2004. 34

[62] "The Collider Detector at Fermilab." Located at `http://www-cdf.fnal.gov`. 35, 105

[63] R. e. a. Brun, "GEANT3 User's Guide," *CERN–DD/EE 84–1* (1986). 36

[64] Particle Data Group, "The Review of Particle Physics,". 37, 38, 41

[65] R. M. Sternheimer, "Density Effect for the Ionisation Loss in Various Materials," *Phys. Rev.* **88** (1952). 38

[66] "Particle Detectors". Cambridge University Press, 1996. 38, 39, 40, 42

[67] G. Grindhammer and S. Peters, "The parametrized simulation of electromagnetic showers in homogeneous and sampling calorimeters," *Int. Conf. of Monte-Carlo Simulation in High Energy and Nuclear Physics, Tallahassee, Florida, USA* (1993) arXiv:hep-ex/0001020. 39, 47, 49, 64, 75

[68] G. Molière, "Winkel- und räumliche Verteilung der Elektronenkomponente der großen Luftschauer". Springer, Berlin (1953). 44

[69] H. Bethe and W. Heitler *Proc. Roy. Soc.* **A146** (1934) 83. 44

[70] G. Grindhammer, M. Rudowicz, and S. Peters, "The Fast Simulation of Electromagnetic and Hadronic Showers," *Nucl. Instrum. and Methods* **A290** (1990) 469. doi:10.1016/0168-9002(90)90566-O. 45

[71] B. Rossi, "High-Energy Particles," *Prentice Hall, New York* (1952). 46

[72] G. e. a. Akopdjanov *Nucl. Instrum. Meth* **140, 441** (1977). 47

[73] G. A. et al., "164, 67," *Nucl. Instrum. Meth.* (1979). 47

[74] G. Ferri and et al. *Nucl. Instrum. Meth.* **A273, 123** (1988). 47

[75] J. del Peso and E. Ros *Nucl. Instrum. Meth.* **A276, 456** (1989). 47, 49

[76] S. Peters, "Die parametrisierte Simulation elektromagnetischer Schauer". PhD thesis, Universität Hamburg, 1992. 49

[77] D. Barney, "Pedagogical Introduction to the CMS Electromagnetic Calorimeter," *Conference Report* (1997). 51

[78] **LEP Working Group for Higgs boson searches** Collaboration, R. Barate et al., "Search for the standard model Higgs boson at LEP," *Phys. Lett.* **B565** (2003) 61–75, arXiv:hep-ex/0306033. 89

[79] "The LEP Electroweak Working Group." Located at http://lepewwg.web.cern.ch/LEPEWWG/, Status of July 2006. 89

[80] F. Halzen and A. D. Martin, "QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS,". New York, Usa: Wiley (1984) 396p. 89

[81] S. Fukuda, Y. Fukuda, M. Ishitsuka, Y. Itow, T. Kajita, J. Kameda,
K. Kaneyuki, K. Kobayashi, Y. Koshio, M. Miura, S. Moriyama,
M. Nakahata, S. Nakayama, Y. Obayashi, A. Okada, K. Okumura,
N. Sakurai, M. Shiozawa, Y. Suzuki, H. Takeuchi, Y. Takeuchi, T. Toshito,
Y. Totsuka, S. Yamada, M. Earl, A. Habig, and E. Kearns, "Tau Neutrinos
Favored over Sterile Neutrinos in Atmospheric Muon Neutrino Oscillations,"
Phys. Rev. Lett. **85** (Nov, 2000) 3999–4003.
doi:10.1103/PhysRevLett.85.3999. 90

[82] R. N. Mohapatra and G. Senjanović, "Neutrino Mass and Spontaneous
Parity Nonconservation," *Phys. Rev. Lett.* **44** (Apr, 1980) 912–915.
doi:10.1103/PhysRevLett.44.912. 90

[83] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon,
S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright,
C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard,
H. V. Peirs, L. Verde, and J. L. Weiland, "First Year Wilkinson Microwave
Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic
Results," *The Astrophysical Journal* **148** (2003) 1. 90

[84] U. Amaldi, W. de Boer, P. H. Frampton, H. Furstenau, and J. T. Liu,
"Consistency checks of grand unified theories," *Phys. Lett.* **B281** (1992)
374–383. 91

[85] E. Witten, "Mass Hierarchies in Supersymmetric Theories," *Phys. Lett.*
B105 (1981) 267. 92

[86] "The Official String Theory Web Site." Located at
http://superstringtheory.com/index.html. 93

[87] G. Nordstrom *Phys. Zeitschr.1* **504** (1914). 93

[88] T. Kaluza *Phys. Math. K.1* **966** (1921). 93

[89] O. Klein *Phys. Zeitschr.37* **895** (1926). 93

[90] L. Randall and R. Sundrum, "An alternative to compactification," *Phys.
Rev. Lett.* **83** (1999) 4690–4693. doi:10.1103/PhysRevLett.83.4690. 95, 98

[91] N. Arkani-Hamed, S. Dimopoulos, G. R. Dvali, and N. Kaloper, "Infinitely
large new dimensions," *Phys. Rev. Lett.* **84** (2000) 586–589,
arXiv:hep-th/9907209. 95

[92] T. G. Rizzo, "Pedagogical introduction to extra dimensions," *ECONF*
C040802 (2004) L013, arXiv:hep-ph/0409309. 96

[93] J. D. Lykken, "Phenomenology beyond the standard model," *Czech. J. Phys.* **55** (2005) B577–B598, `arXiv:hep-ph/0503148`. 96

[94] I. Antoniadis, "A Possible new dimension at a few TeV," *Phys. Lett.* **B246** (1990) 377–384. `doi:10.1016/0370-2693(90)90617-F`. 96

[95] N. Arkani-Hamed and M. Schmaltz, "Hierarchies without symmetries from extra dimensions," *Phys. Rev. D* **61** (Jan, 2000) 033005. `doi:10.1103/PhysRevD.61.033005`. 97

[96] T. Appelquist, H.-C. Cheng, and B. A. Dobrescu, "Bounds on universal extra dimensions," *Phys. Rev.* **D64** (2001) 035002. `doi:10.1103/PhysRevD.64.035002`. 97

[97] H.-C. Cheng, J. L. Feng, and K. T. Matchev, "Kaluza-Klein dark matter," *Phys. Rev. Lett.* **89** (2002) 211301, `arXiv:hep-ph/0207125`. `doi:10.1103/PhysRevLett.89.211301`. 97

[98] G. Servant and T. M. P. Tait, "Is the lightest Kaluza-Klein particle a viable dark matter candidate?," *Nucl. Phys.* **B650** (2003) 391–419, `arXiv:hep-ph/0206071`. 97

[99] A. J. Barr, "Using lepton charge asymmetry to investigate the spin of supersymmetric particles at the LHC," *Phys. Lett.* **B596** (2004) 205–212, `arXiv:hep-ph/0405052`. 97

[100] T. G. Rizzo, "Probes of universal extra dimensions at colliders," *Phys. Rev.* **D64** (2001) 095010, `arXiv:hep-ph/0106336`. 97

[101] K. Agashe, N. G. Deshpande, and G. H. Wu, "Universal extra dimensions and b → s gamma," *Phys. Lett.* **B514** (2001) 309–314, `arXiv:hep-ph/0105084`. 97

[102] T. Appelquist and B. A. Dobrescu, "Universal extra dimensions and the muon magnetic moment," *Phys. Lett.* **B516** (2001) 85–91, `arXiv:hep-ph/0106140`. 97

[103] F. J. Petriello, "Kaluza-Klein effects on Higgs physics in universal extra dimensions," *JHEP* **05** (2002) 003, `arXiv:hep-ph/0204067`. 97

[104] T. Appelquist and H.-U. Yee, "Universal extra dimensions and the Higgs boson mass," *Physical Review D (Particles and Fields)* **67** (2003), no. 5, 055002. 98

[105] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373. `doi:10.1103/PhysRevLett.83.3370`. 98

[106] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall-Sundrum gauge hierarchy model," *Phys. Rev. Lett.* **84** (2000) 2080. `doi:10.1103/PhysRevLett.84.2080.` 99

[107] **D0** Collaboration, "Search for Randall-Sundrum Gravitons in Dilepton and Diphoton Final States," *Physical Review Letters* **95** (2005), no. 9, 091801. 99

[108] **CDF** Collaboration, G. Landsberg, "Collider searches for extra dimensions," *ECONF* **C040802** (2004) MOT006, `arXiv:hep-ex/0412028.` 99, 107, 149

[109] G. F. Giudice, R. Rattazzi, and J. D. Wells, "Quantum gravity and extra dimensions at high-energy colliders," *Nucl. Phys.* **B544** (1999) 3–38, `arXiv:hep-ph/9811291.` 101, 102, 103, 107, 109

[110] T. Han, J. D. Lykken, and R.-J. Zhang, "On Kaluza-Klein states from large extra dimensions," *Phys. Rev.* **D59** (1999) 105006, `arXiv:hep-ph/9811350.` `doi:10.1103/PhysRevD.59.105006.` 101

[111] L. A. Anchordoqui, H. Goldberg, and A. D. Shapere, "Phenomenology of Randall-Sundrum black holes," *Phys. Rev. D* **66** (Jul, 2002) 024033. `doi:10.1103/PhysRevD.66.024033.` 101

[112] S. B. Giddings and S. Thomas, "High energy colliders as black hole factories: The end of short distance physics," *Phys. Rev. D* **65** (Feb, 2002) 056010. `doi:10.1103/PhysRevD.65.056010.` 101

[113] E. A. Mirabelli, M. Perelstein, and M. E. Peskin, "Collider signatures of new large space dimensions," *Phys. Rev. Lett.* **82** (1999) 2236–2239, `arXiv:hep-ph/9811337.` 101

[114] J. L. Hewett, "Indirect collider signals for extra dimensions," *Phys. Rev. Lett.* **82** (1999) 4765–4768. `doi:10.1103/PhysRevLett.82.4765.` 102, 107

[115] C. P. W. G. E. Accomando, A. Aranda, E. Ateser, C. Balazs, D. Bardin, T. Barklow, M. Battaglia, W. Beenakker, S. Berge, G. Blair, E. Boos, F. Boudjema, H. Braun, P. Burikham, H. Burkhardt, M. Cacciari, O. Cakir, A. Ciftci, R. Ciftci, B. Cox, C. Da Via, A. Datta, S. De Curtis, A. De Roeck, M. Diehl, A. Djouadi, D. Dominici, J. Ellis, A. Ferrari, J. Forshaw, A. Frey, G. Giudice, R. Godbole, M. Gruwe, G. Guignard, T. Han, S. Heinemeyer, C. Heusch, J. Hewett, S. Jadach, P. Jarron, C. Kenney, Z. Kirca, M. Klasen, K. Kong, M. Kramer, S. Kraml, G. Landsberg, J. L. Diaz-Cruz, K. Matchev, G. Moortgat-Pick, M. Muhlleitner, O. Nachtmann, F. Nagel, K. Olive, G. Pancheri, L. Pape, S. Parker, M. Piccolo, W. Porod, E. Recepoglu, P. Richardson, T. Riemann, T. Rizzo, M. Ronan, C. Royon, L. Salmi, D. Schulte, R. Settles, T. Sjostrand, M. Spira, S. Sultansoy, V. Telnov,

D. Treille, M. Velasco, C. Verzegnassi, G. Weiglein, J. Weng, T. Wengler, A. Werthenbach, G. Wilson, I. Wilson, and F. Zimmermann, "Physics at the CLIC Multi-TeV Linear Collider," 2004. 102

[116] L. Vacavant and I. Hinchliffe, "Signals of models with large extra dimensions in ATLAS," *J. Phys.* **G27** (2001) 1839–1850. 103, 112, 137

[117] J. Ruppert, C. Rahmede, and M. Bleicher, "Determination of the fundamental scale of gravity and the number of space-time dimensions from high energetic particle interactions," *Phys. Lett.* **B608** (2005) 240–243, `arXiv:hep-ph/0501028`. 103

[118] J. Hewett and M. Spiropulu, "Particle Physics Probes Of Extra Spacetime Dimensions," *Annual Review of Nuclear and Particle Science* **52** (2002) 397. 104, 105, 107, 149

[119] S. Mele, "Experimental constraints on extra dimensions," *European Physical Journal C* **33** (2004) 919–923. `doi:10.1140/epjcd/s2003-03-1106-9`. 104, 106

[120] A. Kehagias and K. Sfetsos, "Deviations from the $1/r^2$ Newton law due to extra dimensions," *Phys. Lett.* **B472** (2000) 39–44, `arXiv:hep-ph/9905417`. 104

[121] **EOT-WASH Group** Collaboration, E. G. Adelberger, "Sub-millimeter tests of the gravitational inverse square law," `arXiv:hep-ex/0202008`. 104

[122] C. D. Hoyle et al., "Sub-millimeter tests of the gravitational inverse-square law," *Phys. Rev.* **D70** (2004) 042004, `arXiv:hep-ph/0405262`. 104

[123] C. Hanhart, D. R. Phillips, S. Reddy, and M. J. Savage, "Extra dimensions, SN1987a, and nucleon nucleon scattering data," *Nucl. Phys.* **B595** (2001) 335–359, `arXiv:nucl-th/0007016`. 105

[124] S. Hannestad and G. Raffelt, "New supernova limit on large extra dimensions," *Phys. Rev. Lett.* **87** (2001) 051301, `arXiv:hep-ph/0103201`. 105

[125] L. J. Hall and D. R. Smith, "Cosmological constraints on theories with large extra dimensions," *Phys. Rev.* **D60** (1999) 085008, `arXiv:hep-ph/9904267`. 105

[126] S. Hannestad, "Strong constraint on large extra dimensions from cosmology," *Phys. Rev.* **D64** (2001) 023515, `arXiv:hep-ph/0102290`. 105

[127] S. Hannestad and G. G. Raffelt, "Stringent neutron-star limits on large extra dimensions," *Phys. Rev. Lett.* **88** (2002) 071301, `arXiv:hep-ph/0110067`. 105

[128] M. Fairbairn, "Cosmological constraints on large extra dimensions," *Phys. Lett.* **B508** (2001) 335–339, `arXiv:hep-ph/0101131`. 105

[129] **The DELPHI Collaboration** Collaboration, P. Abreu et al., "Measurement of the Mass and Width of the W Boson in e^+e^- Collisions at $\sqrt{s} = 189$ GeV," *Phys. Lett.* **B511** (2001) 159. `doi:10.1016/S0370-2693(01)00453-1`. 105

[130] "The L3 experiment." Located at `http://l3.web.cern.ch/l3`. 105

[131] "The OPAL experiment." Located at `http://cern.ch/opal`. 105

[132] G. Landsberg, "Extra dimensions and more..," `arXiv:hep-ex/0105039`. 105

[133] **DØ** Collaboration, V. M. Abazov et al., "Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using kinematic characteristics of lepton + jets events," *Phys. Lett.* **B626** (2005) 45–54, `arXiv:hep-ex/0504043`. `doi:10.1016/j.physletb.2005.08.104`. 105

[134] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework," *Nucl. Inst. and Meth. in Phys. Res. A* **389** (1997), no. 1-2, 81–86. See also `http://root.cern.ch/`. 112

[135] D. R. Stump, "A new generation of CTEQ parton distribution functions with uncertainty analysis." Prepared for 31st International Conference on High Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 24-31 Jul 2002. 112

[136] **CDF** Collaboration, L. Duflot and M. Kado, "Probing extra spacetime dimensions at the Tevatron," `arXiv:hep-ex/0211060`. 115

[137] V. Drollinger, "The Beam Halo Generator Manual with Validation Plots," *CMS Internal NOTE* **2005/04** (2005). 116

[138] V. Drollinger, "Simulation of Beam Halo and Cosmic Muons," *CMS Note* **2005/012** (2005). 116

[139] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 1," *CERN/LHCC* **2006-00** (2006). CMS TDR 8.1. 121

[140] **CMS** Collaboration, B. Clerbaux, T. Mahmoud, C. Collard, M.-C. Lemaire, and V. Litvin, "TeV electron and photon saturation studies," *CMS NOTE* **2006/004** (2006). 123

Joanna Weng

- **May 1977**: Born in Warsaw.

- **1984-1997**: Attended primary school and Paulsen Gymnasium in Berlin.
  June 1997: "Abitur".

- **1997-1999**: Studies of Physics at the Technical University Berlin (TU).
  May 1997: "Vordiplom" .

- **1999-2000**: Studies of Physics with ERASMUS scholarship at the University
  Louis Pasteur (ULP), Strassbourg, France.

- **2000-2003**: Studies of Physics at the Karlsruhe University (TH).
  February 2003: "Diplom".

- **February 2003-September 2003**: Doctoral student at the IEKP[1], University of Karlsruhe, and member of the HIK/GES[2] at the Forschungszentrum Karlsruhe (Doctoral student position at the Forschungszentrum).
  Research Topics: Data analysis, simulation of the CMS detector, studies of the physical potential of the LHC, GRID computing.

- **September 2003-November 2006**: Doctoral student at the IEKP, University of Karlsruhe, with doctoral student grant and permanent positioning at CERN[3] Geneva, Switzerland (CERN doctoral student program).
  Research Topics: Parameterisation of electromagnetic showers, search for large extra dimensions (GRID based analysis), development and deployment of CMS detector simulation software.

---

[1]Institut für Experimentelle Kernphysik
[2]Institute for Grid Computing and e-science, now IWR (Institute for scientific computing)
[3]European Organisation for Nuclear Research

Acknowledgements

First of all, I am most grateful to Prof. Dr. Günter Quast and Prof. Dr. Albert De Roeck, who proposed me this exciting topic and were supervising my thesis. They allowed me to work freely and independently and gave me the unique opportunity to spend most of the time at CERN. The last 3 years here have been - quite simply - the most interesting and exciting in my life ! Real thanks as well to Maria Spiropulu for all the help in my analysis and other fields of life. And for a lot of fun !

This work would not have been possible without the assistance of my CMS colleagues at CERN. Many thanks for all the help and fruitfull discussions to Maya Stravrianakou, Filip Moortgat, Georgia Karapostoli, Anne-Sylvie Giolo Nicollerat, Monica Vasquez-Acosta, Patrick Janot, Sunanda Banerjee, Oliver Buchmueller and Stephan Wynhoff.
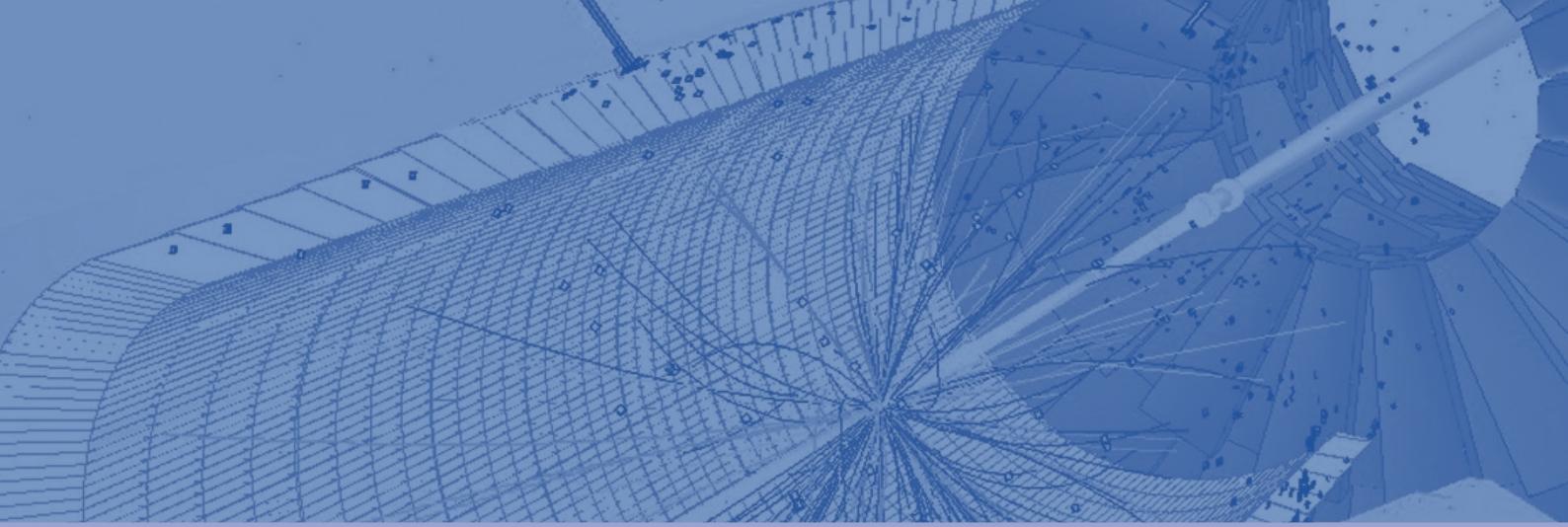
I am also grateful to the Geant4 team for their support, especially to Hans-Peter Wellisch, John Apostolakis, Gabrielle Cosmo and Alexander Howard.
In the context of my generator studies, I would like to thank Frank Krauss, Konstantin Matchev, Joe Lykken and especially Tanju Gleisberg.

Many thanks as well to my Karlsruhe colleagues: Ulrich Felzmann, Christian Weiser, Alexander Schmidt and Ulrich Kerzel for their help at work and for the cool time we had in Geneva ! Thanks to Klaus Rabbertz, Christian Sander, Ulrich Kerzel, Alexander Howard, Wim de Boer, Maria Spiropulu, Altan Cakir and Alexander Schmidt for reading and correcting this thesis. Special thanks to Christophe Saout for his contribution to the analysis part and a lot of technical and editorial help.

I am also thankful to my dear friends in Geneva, Katharina & Gregor Mair, for their mental support and all the fun.

Finally, I would like to express my deepest gratitude to my family, for all their support, comprehension and care. I love you.

For the start-up of the future proton-proton collider LHC at CERN in 2008 a multitude of simulated events are needed in order to understand the detectors and perform physics analyses at a centre-of-mass energy of 14 TeV. The solution presented here, namely the usage of shower parameterisation, can speed-up this simulation considerably without sacrificing too much precision. A package based on this concept, GFlash, was implemented in the general simulation framework of Geant4. It has been successfully integrated in the CMS detector simulation and has proven to work with an impressive speed and acceptable accuracy.

In the second part, the discovery potential of large extra dimensions (ADD model) with the CMS detector in the photon + missing transverse energy final state was studied. A normalisation method is proposed to measure the main background from data. The analysis shows that via this channel a 5 sigma discovery can be even made in the first year.